GUBERNATORIAL AGENDAS AND FISCAL OUTCOMES

by

MIKHAIL IVONCHYK

(Under the Direction of Tima Moldogaziev)

ABSTRACT

This dissertation uses computational text analysis to study gubernatorial policy agendas and analyze how they translate into state and city fiscal outcomes. Governors are commonly perceived as state leaders responsible for setting state policy agenda. State chief executives, however, are generally disadvantaged in the legislative process vis-à-vis legislators. Extant literature suggests that governors can use several formal and informal tools to overcome this disadvantage and achieve their policy goals. Empirical evidence on the power of American governors in state policymaking, however, remains inconclusive. It is also unclear from the current literature to what extent, if any, city officials may respond to gubernatorial policy priorities. This dissertation employs unsupervised machine learning algorithms to extract the main policy themes from the transcripts of gubernatorial speeches. Next, the dissertation explores the relationship between gubernatorial policy agenda and state and city fiscal outcomes. Empirical findings indicate that gubernatorial policy agenda may be a powerful force in some state policy areas, but not all. The results also suggest that in some policy areas cities align their allocation of fiscal resources with gubernatorial policy goals.

INDEX WORDS: agenda, governor, fiscal policy, computational text analysis, Latent Dirichlet Allocation, unsupervised machine learning, intergovernmental relations

GUBERNATORIAL AGENDAS AND FISCAL OUTCOMES

by

MIKHAIL IVONCHYK

BPA, Academy of Public Management, Belarus, 2007

LLM, Belarusian State University, 2011

LLM, University of Georgia, 2012

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial

Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2019

GUBERNATORIAL AGENDAS AND FISCAL OUTCOMES

by

Mikhail Ivonchyk

Major Professor: Tima Moldogaziev

Committee: Laurence J. O'Toole
                  L. Jason Anastasopoulos
                  Katherine Willoughby

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2019

DEDICATION

I dedicate this dissertation to my family and especially to my wife, Katia, whose patience, support and encouragement made this journey possible.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION AND LITERATURE REVIEW

This study uses an unsupervised machine learning algorithm to identify the main policy items that have been on the agenda of American governors over the last decade. Next, it tests how these policy items translate into state and local fiscal policies. By applying computational text analysis to gubernatorial speeches, this work complements the existing literature with new empirical evidence on the ability of state chief executives to influence fiscal behavior of state and local governments.

The main aim of the first chapter is to extract and quantify the policy themes discussed in gubernatorial speeches over the last decade. The speeches used are high-profile State of the State addresses delivered by all governors throughout the nation. These events provide a unique opportunity to speak to both houses of the legislatures and lay out gubernatorial policy agenda for the upcoming legislative sessions. The content of the speeches is often used as a gauge of gubernatorial policy agenda, but the analyses have been limited to nonrandom samples, short time periods and have relied on unclear classification criteria. The first chapter explains the main logic and formal aspects of Latent Dirichlet Allocation probabilistic topic model. Then, the topic model is applied to a collection of all State of the State addresses delivered between 2007 and 2017 to infer the main policy themes discussed by governors. The expectations are that the policy items identified by the topic model will conform to the manual results from the previous research. This chapter is intended to illustrate how modern advancements in computer science can be utilized to quantify large collections of text documents in a much more efficient manner that may be less vulnerable to human error and subjective judgement.

The second chapter uses the data generated in the first chapter as a measure of gubernatorial policy agenda and tests the ability of state chief executives to influence state spending choices. American governors are commonly perceived as leaders symbolizing the state with substantial influence on state policymaking, even though state constitutions strip governors of any direct legislative powers. The only direct legislative authority of state chief executives comes at the end of the lawmaking process, when governors can sign, ignore or veto a bill put forward by legislators.

Governors, nonetheless, possess an array of formal and informal tools that improve their bargaining stance in the lawmaking process. Being the most visible of individual state officials, a governor can boost support for the reelection of individual lawmakers through public speeches and appearances, or publicly criticize them to reduce their popularity. Most governors possess various veto powers allowing them to stop or change legislators' pet projects. Governors also make important appointments and, in most states, possess various powers to reorganize state agencies through executive orders (Council of State Governments, 2017). They generally have more access to the media than any other state official does. Using these formal and informal methods, state chief executives may be able to stimulate support among the lawmakers for their policy initiatives.

Previous literature has tackled this question from different angles. There have been multiple surveys asking state administrators at different levels about the power of governors to influence their activities (Abney & Lauth, 1983; Abney & Lauth, 1987, 1998; Dometrius & Wright, 2010; Goodman, 2007; Ryu, Bowling, CHO, & Wright, 2008; Thompson, 1987), with contradictory findings. The results may be compromised by the fact that these studies often relied on perceptual measures. Other studies show that the power balance between a governor and the legislature is not a zero-sum game, where one institution gains influence at the expense of the other (Dilger, Krause, & Moffett, 1995). Gubernatorial ability to influence may be enhanced by their

institutional powers in the budgetary process (Barrilleaux & Berkman, 2003; Breunig & Koski, 2009), with increasing unilateral control over fiscal policy decisions (Krause & Melusky, 2012) and experience and staff support (Ferguson, 2003). Economic slowdowns, divided government and being in the last year of service may decrease the rate of gubernatorial success in the policymaking process (Ferguson, 2003; Kousser & Phillips, 2012).

Whereas the powers of the state chief executive have been tested in the past, our knowledge remains fragmented and inconsistent. Most of the previous studies relied on nonrandom samples of states which limits external validity, or used cross-sectional datasets, which weakens the internal validity of their findings. Thus, the goal of the second chapter is to advance our knowledge by analyzing a panel dataset on all states from 2007 to 2015. It is expected that gubernatorial policy priorities will be significantly correlated with state spending patterns. The relationship, however, may be conditional on the political environment, institutional power and socio-economic conditions.

The third chapter tests whether there is any relationship between gubernatorial policy agenda and city fiscal outcomes. It is known from the previous literature that binding state institutions significantly affect local fiscal behavior (T. Brown, 2000; Joyce & Mullins, 1991; Mullins & Joyce, 1996; Preston & Ichniowski, 1991; Shadbegian, 1998, 1999). The main purpose of this chapter is to test whether municipal governments react to softest and non-binding signals, such as gubernatorial policy priorities. Governors have ample opportunities to share their policy goals with a broad audience and quite often successful implementation might depend on whether local officials go along with gubernatorial policy initiatives. Whereas policy statements are not binding for lower level officials, they send a clear signal about the issues that matter to the state chief executives. Local officials will likely pay heed to those statements as they operate in the

shadow of the fact that the informal signals may be backed up by formal action, be it new legislation, changes in local aid, or an important appointment. If the proposition is true, it is expected that there will be a significant correlation between gubernatorial policy priorities and city spending patterns, which reflect local officials' policy choices. This hypothesis is tested in a longitudinal study of the largest cities from 48 states. Previous literature is largely silent on the ability of state chief executives to influence local behavior through public speeches. This study seeks to fill this gap and test the relationship between gubernatorial policy goals expressed in public speeches and local fiscal outcomes.

CHAPTER 1

TEXT AS DATA: A STUDY OF GUBERNATORIAL POLICY AGENDA WITH

MACHINE LEARNING

Introduction

The volume of available information is constantly growing at an exponential rate (Szalay & Gray, 2006). Almost all of it nowadays is in digital format (Cukier & Mayer-Schoenberger, 2013, p. 29). The main implication for researchers is that there is an instant access to an ever-growing variety of data. Among these are different types of text documents, such as public speeches, legislative deliberations, administrative hearings and various reports generated by all levels of government. This potentially rich source of information may provide public administration and policy scholars opportunities for new discoveries that have not been possible before.

Modern collections, however, may consist of thousands of documents and millions of words, which along with opportunities introduces new challenges. The traditional approach to text analysis involving manual coding becomes practically unfeasible. Individual scholars rarely can manually quantify even small collections of documents, and larger sets inevitably require expensive external assistance. As a result, researchers without extensive funding must be content with small subsets or rely on existing, previously analyzed datasets. Consequently, to reap the benefits of an increasing availability of large volumes of text documents, researchers need new information processing methods that will help overcome the high costs of analyzing text and enable individual scholars to automatically extract meaningful data.

To this end, this paper aims to describe and illustrate an application of a machine learning algorithm that will be useful in expanding the pool of currently analyzed texts. The focus here is on probabilistic topic modeling, which represents "a suite of algorithms that aims to discover and annotate large archives of documents with thematic information" (Blei, 2012, p. 78). Many scientific disciplines as well as private enterprises have been using machine learning algorithms to deal with large volumes of text data, but the method has not received much attention in the study of public policy and administration. This is paradoxical given that public organizations produce a great deal of easily-accessible documents. Language is used to convey policy goals, to comment on administrative regulations, to explain and justify government decisions. A systematic analysis of these texts can yield new insights for public administration and policy scholars. The goal of this study is to show how our ability to deal with text data can be augmented with automated content analysis methods.

The choice of a suitable machine learning algorithm depends on specific research purposes. The method evaluated here is the Latent Dirichlet Allocation topic model (LDA) developed by Blei, Ng, and Jordan (2003). LDA has been increasingly popular[1] and has proven to be reliable for different types of texts, including tweets (Weng, Lim, Jiang, & He, 2010), deliberations of Federal Open Market Committee (Hansen et al., 2014), abstracts of journal articles (Griffiths & Steyvers, 2004), New York Times articles (Zhao et al., 2011) and medical studies (Wu, Liu, Zheng, Zhao, & Xu, 2012). If the goal is to infer the thematic content of a large set of documents, this method

---

[1] The original article introducing LDA was cited 8,000 times between 2003 and 2014 (Hansen, McMahon, & Prat, 2014, p. 4), and over 14,000 times since 2014.

seems to be one of the most promising ones. Without reading the documents, it is possible to find the topics they discuss as well as the share of each document devoted to each topic.

The next section explains the main intuition behind LDA. The third section illustrates how LDA can be used to infer the main policy topics from transcripts of the State of the State addresses delivered by all American governors between 2007 and 2017. The content of these speeches is commonly used in public policy and administration research as a gauge of the gubernatorial policy agenda. An average speech lasts about half an hour, and this collection contains over two million words. It would be a nontrivial effort for a researcher to manually code them, so extant studies have relied on nonrandom samples or focused on a single year. Both solutions undermine the validity of empirical findings. Besides being an important indicator of an executive policy agenda, the addresses are also chosen because it is possible to know from manual analysis the main topics running through these documents (Willoughby, 2008, 2009, 2016, 2017). Thus, the focus on the State of the State addresses provides a rare opportunity to assess LDA output against manual results.

The results of the current study indicate that LDA output compares very favorably against manual coding. LDA has been able to automatically identify the main themes discussed by governors and indicate the proportion of each address devoted to each theme. Thus, LDA appears to be an efficient way to quantify, analyze and visualize large collections of documents with limited prior knowledge of their content. Formally, LDA does not require any knowledge about the thematic structure of documents. But as this and other studies (Chang, Gerrish, Wang, Boyd-Graber, & Blei, 2009; Grimmer & Stewart, 2013) illustrate, human validation of automated content analysis is indispensable for meaningful results. Overall, the findings suggest potentially wide application of this method in the study of public administration and policy.

The Intuition Behind Latent Dirichlet Allocation Topic Model

This section lays out the main logic of the Latent Dirichlet Allocation topic model based on its original description (Blei, 2012; Blei et al., 2003). A formal description of the model is provided in Appendix A. The intuition behind LDA is that each document in a collection may contain multiple topics. This document collection is often referred to as a corpus. LDA exploits the fact that there is a tendency to use similar words when talking about the same topics. This co-occurrence of words is used to infer the main word clusters from a corpus. LDA represents each cluster as a collection of words ordered by their probability of appearing in this cluster. Each cluster contains all unique words found in the corpus, but the probability of appearing for each word differs between clusters. All words' probabilities in each cluster add up to one, so more formally a cluster is defined as a probability distribution over a fixed vocabulary. LDA is bag-of-words model because the word order is not taken into account.

The most probable words in each cluster, say top five or top ten, can be used to annotate the cluster with thematic meaning. For example, if the top five words in an LDA cluster are "teacher," "school," "graduation," "tests," "dropout," it becomes clear that this cluster is related to the topic of education and it is possible to manually label this cluster as "education." A thematically unrelated word "airplane," for example, may also be in this cluster, but it will have much lower probability of appearing in this cluster. Since these clusters of words provide meaningful information about the thematic content of documents, they are often referred to as topics. Each document in the corpus may exhibit any combination of such topics. LDA represents each document as a distribution over topics. This means that it estimates the relative share of each document devoted to each topic. The sum of all shares adds up to one. This relative share of each topic can be used to distinguish the documents based on their thematic content.

LDA is an unsupervised machine learning algorithm. This means that there is no need for a set of annotated documents to train the model. LDA uses the observed documents to automatically infer the underlying topic structure that most likely generated the observed collection of words. This inferred hidden topic structure generally resembles the thematic structure of the documents (Blei, 2012). But although LDA does not require any prior knowledge of the content of the documents, it is necessary to specify the number of unique topics expected to be found in the analyzed corpus. The primary guide for researchers in selecting the appropriate number of topics should be theoretical expectations, expert knowledge and substantive meaning of the results (Blei, 2012; Chang et al., 2009). The number of topics can be adjusted based on how well the LDA results fare in terms of these criteria.

Additionally, researchers can estimate several models with different numbers of topics and estimate their perplexity score. Perplexity is a theoretical measure of how well a topic model with given parameters predicts observed data.[2] The lower the perplexity score, the better fit to the data. Researcher may select the number of topics at which the marginal perplexity score stops decreasing (Blei & Lafferty, 2007). It is worth mentioning, however, that parsimony is important. Experiments show that topics become more fine-grained and less useful for humans with a larger number of estimated topics (Chang et al., 2009). Hence, the perplexity score should complement, rather than replace, other criteria mentioned earlier.

In this study, an LDA topic model is applied to State of the State addresses delivered by all American governors from 2007 to 2017. Researchers have manually analyzed the content of

---

[2] Please refer to Blei et al. (2003) for a formal definition of the perplexity score.

these speeches and found that they discussed 15 distinct topics (Willoughby, 2008, 2017). Based on these findings, the LDA topic model is set to find 15 unique topics in the corpus. The following section describes the analyzed corpus and the main steps commonly performed to prepare the text for analysis.

LDA Application Example: State of the State Addresses

LDA is illustrated in this study by applying the model to discover the main policy topics discussed in gubernatorial speeches. The corpus comprises all State of the States Addresses delivered by US state governors from 2007 through 2017. The collection includes a total of 519 documents.[3] The intention is to demonstrate how machine learning algorithms can be used to extract meaningful data from text and in this case to automatically create a panel dataset quantifying the gubernatorial policy agenda that varies across states and over time.

Admittedly, governors receive a lot of attention from the media and have multiple opportunities to share their agenda with a wide audience. There are several reasons for choosing the State of the State Addresses as a source of their agenda. These are high profile events taking place in all states, most of the time in January, and in a few states in February or March. The speech signifies the start of the legislative session and provides a unique opportunity for all governors to lay out their policy goals for the upcoming year in front of both chambers of the legislature.[4] It is a highly-anticipated event often broadcasted live on local news channels. Given its significance, the process of drafting the speech begins weeks in advance (Kousser & Phillips, 2012, p. 74). The

---

[3] The corpus contains fewer than 550 documents because the speech is delivered biennially (usually in odd years) in Arkansas, Montana, Nevada, North Carolina, North Dakota and Texas. Except for Arkansas, these states have biennial legislative sessions.
[4] Nebraska is the only exception here with a unicameral, nonpartisan legislature.

address is limited in time, so governors have to pick carefully the most important policy initiatives for the year (Herzik, 1991; Light, 1999). These speeches generally contain a range of policy statements and allow a much more nuanced measurement of gubernatorial policy goals in each year than any other public speech. The address is intended to influence public opinion as well as legislative process (Rosenthal, 1990; Van Assendelft, 1997). It is a superior gauge of executive policy goals (Ferguson, 2003) and a good predictor of subsequent gubernatorial actions (Jackson & Kingdon, 1992; Segal & Cover, 1989). As a result, the content of the speeches has been used to measure the substantive content of a gubernatorial policy agenda in multiple studies (e.g. Crew, 1992; DiLeo, 2001; Ferguson, 2003; Herzik, 1991; Kousser & Phillips, 2012; Morehouse, 1998).

Other gubernatorial messages are generally shorter and much narrower in scope. The annual or biennial executive budget message, for instance, may also shed some light on the executive policy preferences, but its length varies from barely a page, as in California and North Carolina, to multiple pages. Apparently, this statement has a different value to state chief executives, in some cases only nominal, and serves different goals. Most of the time the budget messages focus on specific issues pertaining to state fiscal performance rather than general policy goals of the administration. It usually provides a summary of the recommended budget, describes revenue and expenditure trends, lays out the main assumptions built in the proposed budget and highlights major areas of concern for state fiscal health. Unlike the widely-publicized State of the State addresses, budget messages rarely receive as much attention of the constituency and are not intended as an outline of public policy agenda. In that respect, the State of the State address is a superior source of executive policy preferences. As it has been aptly summed up, "lacking detailed interview or survey data from governors, their major public addresses provide the best insight available into their preferences, values, and ideology" (Coffey, 2005, p. 90).

Earlier studies commonly analyzed the content of text documents by hand with human coders quantifying text into data. Documents were coded based on the number of sentences devoted to a topic (Heidbreder, 2012; Heidbreder & Scheurer, 2013) or the topic structure was simply assumed based on the reading of the document (Ferguson, 2003; Ferguson & Barth, 2002). This is a very time-consuming approach that is also prone to errors. The use of LDA in this study allows me to quantify gubernatorial speeches in a more efficient and objective manner. Assuming that governors devote a larger share of the addresses to the policy issues that are more important, then the probability distribution of documents over topics generated by LDA can be used as a measure of this topic's importance to the governor in a given year. A governor spending 30 percent of the speech on education, for example, is assumed to be more concerned with this policy area than a governor spending only 5 or 10 percent.

Preparing the Corpus

The data input for LDA is a document-term matrix. The rows in this matrix correspond to the documents and the columns to the terms. The entry m_{ij} indicates how frequently the j_{th} term occurred in the i_{th} document. The number of rows is equal to the number of analyzed documents (in this case 519) and the number of columns to the size of unique vocabulary in the corpus. Generally, the text data needs to be pre-processed for the analysis, which determines the number of columns in the matrix. Essential preprocessing steps include converting the text to lower case, removing numbers, punctuation and any other unwanted characters (e.g. " : ", " ; " and other). This collapses identical words into one column and leaves out unusable symbols.

Another important step is to remove so called *stop-words*, which are frequently used, but not informative terms. The most obvious examples include prepositions, some adverbs, articles and pronouns as well as different types of shortenings (e.g. "we've") that do not give the meaning

of the document. Removing the stop-words reduces the size of the matrix and makes the topic inference more effective. Additional words removed in this study were "state", "year", and any other words containing the names of the states, such as "alabama," "alabamians" etc. These terms do not convey any useful information about the topic structure of the speeches but are among the most frequently used in the corpus.

Another important step is called *stemming*, during which the root, or stem of a word is identified, and morphological forms of words are translated to its stem. For example, the words "experienced," "experiences," "experience" are reduced to "experienc." This example also illustrates that in some cases the resulting stem is not a word or root of a word. The main purpose of this step is to "remove various suffixes, to reduce the number of words, to have accurately matching stems and to save time and memory space" (Vijayarani, Ilamathi, & Nithya, 2015, p. 10). The LDA results are discussed in the next section.

LDA results: The gubernatorial policy agenda

Descriptive Statistics of the Text

The original corpus includes a total of 2,288,260 terms,[5] an average of 4,400 words per speech. Assuming the average rate of public speech is about 150 words per minute (Rodero Antón, 2012), the average address should take about 30 minutes. The length of the address in the corpus varies from slightly over 800 words to almost 13,000 words. The shortest speech was delivered by governor Jan Brewer in Arizona 2011 two days after the Tucson shooting, which left 6 dead. Governor Matt Bevin delivered the longest speech of 12,700 words in Kentucky in 2016, when he

---

[5] This total includes all words, numbers, abbreviations and any other distinct terms in the corpus.

laid out his plan to trim $650 million from state expenditures without cutting taxes and called upon both chambers of the legislature to "pass this budget."

The total number of words after pre-processing is 805,399, or about 1,552 words per average speech. Pre-processing thus reduced the original size of the text by about 65 percent. The vocabulary before pre-processing included over 31,100 unique terms, almost 900 of which were different numbers mentioned in the speeches. Four of the unique entries in the corpus vocabulary, for example, included the words "abandon", "abandoned," "abandoning," "abandonment." The most used word in the original text was the stop-word "the" with 106,000 occurrences. The processed vocabulary contains 19,740 terms, or 63 percent of the original set. The term "abandon" now takes one column in the document-term matrix.

As the first glance at the State of the State Addresses content over the last 11 years, it would be useful to analyze the frequency of used words. The word cloud in Figure 1.1 shows the 50 most frequent terms in the corpus. Leaving aside the stop words and state names, the most frequent word in the speeches is "job." It shows up about 7,400 times in the corpus of 519 documents, that's about 14 times per speech. Governor Jennifer Granholm of Michigan holds the record here mentioning "job" 67 times in her 2009 address, whereas only three addresses did not mention this word. This unequivocally shows that job creation and low unemployment was among the most salient issues for state chief executives. It is also evident from the word cloud that along with jobs, education, healthcare, budgeting, economy and taxes fare prominently on the executive agenda.

Figure 1.1. Word cloud of the fifty most frequent words in the corpus.

World clouds and other statistics on word occurrence can be easily obtained from the document-term matrix as a first step in exploring the content of any collections of documents. Even without machine learning algorithms or manual reading, it is easy to get a general idea about the content of the documents. In this case, one could generate clouds by state, year, political parties, individual governors or any other identifier. With some experience, it will take very little time to obtain these statistics, which surely will be quicker than even skimming through 500 documents. In the following subsection, the discussion turns to the LDA probability distributions, which should give us a more detailed picture of the thematic content of the documents.

Probability Distribution of Topics over Words and Documents Over Topics

The LDA topic model represents each of the estimated 15 topics as a probability distribution over words, and each of the 519 State of the State addressees as a probability

distribution over these 15 topics. The topic distribution over words is presented in the Table 1.1.

Words are ordered by their probability of appearing in each topic.

Table 1.1. Probability distribution of 15 topics over words.

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8
Economic Development	Local Government	Healthcare	Public Service	Facility Building	Taxes	Child care	Capital Projects
community	reform	health	serve	school	tax	child	County
develop	economy	care	family	student	income	country	legislature
economy	community	child	child	bill	business	senate	Coal
public	city	family	office	local	reform	care	Project
education	local	economy	world	tax	rate	home	Federal
future	public	invest	honor	course	family	love	Budget
innovate	law	cost	America	home	property	remember	Water
resource	perform	community	bless	build	job	represent	legislate
success	house	insure	important	create	spend	world	encourage
challenge	actual	future	women	level	lower	county	Mine

Topic 9	Topic 10	Topic 11	Topic 12	Topic 13	Topic 14	Topic 15
Transportation	Jobs	Economy	Budget	Education	Public Safety	Energy/Natural Resources
transport	job	economy	budget	school	drug	energy
road	business	budget	revenue	student	abuse	oil
invest	create	public	service	education	law	gas
water	company	cut	spend	teacher	addict	resource
infrastructure	economy	job	cost	child	crime	economy
future	invest	child	depart	college	offend	develop
bridge	worker	school	federal	job	treatment	renew
project	train	choice	plan	graduate	prison	project
improv	hard	challenge	employee	invest	child	wind
growth	manufacture	politics	current	learn	family	power

Table 1.1 lists the 10 most frequent words in descending order for each of the 15 estimated

topics. Each column shows how likely we are to observe each word conditional on the topic. The

sum of all probabilities within each individual column is equal to 1, but the word probabilities

across columns will not add up to 1. This topic distribution over words allows us to assign meaningful labels to each topic (labels are in the second row in bold).

Several metrics can be used to evaluate the fit of an LDA topic model, such as perplexity score. However, experimental studies show that topic models which perform better on these metrics may in fact provide less semantically meaningful topics (Chang et al., 2009). The authors note that models are trading higher likelihood for lower interpretability and suggest that human judgement of topic coherence, usefulness and interpretability is a better way to evaluate the results. Hence, they recommend evaluating topic semantic coherence by looking for *intruders*, or spurious words that are out of place and do not belong with others.

Most of the topics seem to pass this test. Topic 3, for example, includes the words expected to be in a discussion of healthcare policy and the same goes for topics 1, 2, 3, 4, 6, 9, 10, 11, 12, 13, 14, 15. Topics 5, 7 and topic 8, on the other hand, seem to be less coherent. Intruders would be hard to identify here. Judging solely by the most probable words, one might conclude that each of them touches upon several policy subjects. In order to provide meaningful labels to these topics, the LDA results were supplemented with manual analysis of the documents with a greater proportion devoted to these topics. The documents were manually searched for areas where the words from the topic were clustered and the text read to identify the main theme discussed in the cluster. Ultimately, by supplementing the LDA output with manual analysis, the topic distribution over words provides a list of meaningful and district topics mentioned in State of the State addresses from 2007 through 2017.

The next table summarizes the probability distribution of documents over the 15 topics. This LDA output is a matrix with rows representing the documents and columns representing the topics. Each entry indicates the share of each document devoted to each topic in percent. The rows

add up to 1, but the columns may not add to 1. In substantive terms, the shares may serve as an approximation of a topic's importance in each speech. Facing time limitations on the speech, governors will likely spend more time on the topics that are more important. Table 1.2 provides a summary of descriptive statistics of the document distribution over topics.

Table 1.2. Descriptive statistics of document distribution over topic (in percent).

Topic	Mean	SD	Min	Max
Topic 12 – Budget	10.8	9.1	0.4	60.2
Topic 11 – Economy	10.1	7.7	0.3	54.4
Topic 1 – Economic Development	9.1	8.4	0.5	46.6
Topic 3 – Healthcare	8.8	7.3	0.5	39.5
Topic 13 – Education	8.7	6.8	0.3	35.6
Topic 10 – Jobs	7.8	7.2	0.3	45.3
Topic 4 – Public Service	7.6	6.9	0.3	55.6
Topic 6 – Taxes	6.3	6.4	0.2	48.4
Topic 7 – Child care	5.3	6.7	0.2	57.5
Topic 9 – Transportation	5.1	5.4	0.3	42.9
Topic 15 – Energy/Natural Resources	4.8	5.6	0.2	41.4
Topic 5 – Facility Building	4.5	6.5	0.3	60.4
Topic 2 – Local Government	3.8	5.7	0.2	47.5
Topic 8 – Capital projects	3.7	5.8	0.2	47.5
Topic 14 – Public Safety	3.6	5.5	0.2	76.8

The minimum and maximum values indicate that attention to each topic varies from very low to more than a half of the speech. Table 1.2 suggests that on average Budget and Economy topics received most attention over the 11 years of study. The time period includes the Great Recession, so more attention to these issues is expected. The next most important issues are economic development, health care, education and jobs. All of these also are among the most frequently discussed issues according to manual analysis by Willoughby (2016). The summary statistics show variation of policy agenda. Governor Peter Shumlin of Vermont, for example, in 2016 spent over 70 percent of his speech discussing the issue of public safety and reiterated his commitment to make Vermont safer and take additional actions to deal with a drug addiction crisis. Meanwhile, some governors do not mention public safety. The same applies to other topics. The

highest Pearson correlation coefficient between topic proportions is -0.3, indicating that multicollinearity is not a problem.

With these granular data, it is possible to label the topics depending on the policy subsystems they likely belong to. Such aggregation may be useful if one wants to see what policy area, rather than an individual topic, is discussed most. Table 1.3 below summarizes the average share of the speeches devoted to different policy subsystems represented by several topics. According to LDA, the most discussed area was economy and fiscal policy (44 percent), followed by social policy (27 percent) and infrastructure investment (13 percent). Such a focus of attention is expected given the economic challenges states have gone through.

Table 1.3. Document distribution over policy subsystem.

Policy Subsystem	Topic	Mean
Economy / Fiscal Policy	topic1 – Economic Development topic6 – Taxes topic10 – Jobs topic11 – Economy topic12 – Budget	43.9
Social Policy	topic3 – Healthcare topic7 – Child care topic13 – Education topic14 – Public Safety	26.5
Infrastructure investment	topic5 – Facility Building topic8 – Capital projects topic9 – Transportation	13.4
Public Service	topic4 – Public Service	7.6
Energy & Environment	topic15 – Energy/Natural Resources	4.9
Intergovernmental relations	topic2 – Local Government	3.7

Alternatively, it possible to break down the summary statistics by region to get a better view of policy agenda variation throughout the nation. As Table 1.4 shows, the most salient issues

differ by region as measured by the average percent of the State of the State addresses devoted to the topics. The most common issues discussed by the governors in Midwestern states are related to job creation, education and health care. The Northeastern governors were mostly concerned with the economy, budgetary issues and health care. Those in the South region devoted most attention to budgetary issues, education and public service. Chief executives in the West region talked most about economic development, with budget and economy following closely behind.

Table 1.4. Summary statistics of document distribution over topics by region.

Policy Subsystem	Topic	Midwest			Northeast			South			West		
		Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Economy/ Fiscal Policy	topic1	7.9	0.5	28.1	7.3	0.5	34.1	7.8	0.6	30.4	**13.0**	0.9	46.6
	topic6	7.7	0.5	48.4	9.4	0.7	45.3	5.2	0.2	36.1	3.9	0.5	10.9
	topic10	**11.1**	0.5	38.6	5.9	0.3	25.0	8.5	0.4	45.3	5.2	0.4	15.8
	topic11	7.7	0.3	32.1	**13.7**	1.0	54.4	9.3	1.0	32.8	**10.4**	0.4	35.7
	topic12	8.2	0.4	27.5	**11.7**	0.6	60.2	**13.6**	0.5	44.6	**9.3**	0.6	32.6
Social Policy	topic3	**8.4**	0.5	36.4	**10.8**	0.8	39.5	8.4	0.6	26.9	8.3	0.8	38.5
	topic7	5.6	0.4	51.3	3.7	0.2	12.8	6.0	0.3	57.5	5.4	0.5	31.1
	topic13	**9.0**	0.3	28.6	6.4	0.6	28.2	**10.6**	0.8	35.6	8.0	0.5	30.2
	topic14	2.1	0.2	10.9	5.5	0.3	76.8	3.7	0.2	19.9	3.1	0.3	24.1
Infrastr. Investment	topic5	7.5	0.5	60.4	2.6	0.5	10.3	3.6	0.3	11.9	4.2	0.4	11.0
	topic8	2.5	0.3	9.3	1.7	0.2	5.2	3.3	0.3	18.2	6.8	0.4	47.5
	topic9	6.5	0.3	42.9	4.7	0.3	21.9	4.4	0.3	23.4	5.0	0.3	19.4
Public Service	topic4	6.9	0.3	33.5	5.4	0.3	28.9	**9.6**	0.8	39.9	7.4	0.5	55.6
Energy & Environ.	topic15	4.7	0.2	31.5	4.7	0.5	28.6	3.3	0.2	18.4	7.2	0.5	41.4
Intergov. relations	topic2	4.1	0.2	29.3	6.4	0.3	47.5	2.7	0.3	12.5	2.7	0.2	12.0
N (states)		12			9			16			13		

Figure 1.2 below illustrates regional variation of the gubernatorial agenda across time. The increase in executive concern with the economy during the recent recession is evident in all regions with minor differences. Social policy concerns may be on the rise in the Northeast and South regions but have not changed much in the West and Midwest. The time trend suggests increasing attention to infrastructure investment in the western regions. Executive attention to energy and

environment followed a similar trend in all regions; an increase during the recession followed by a decline.
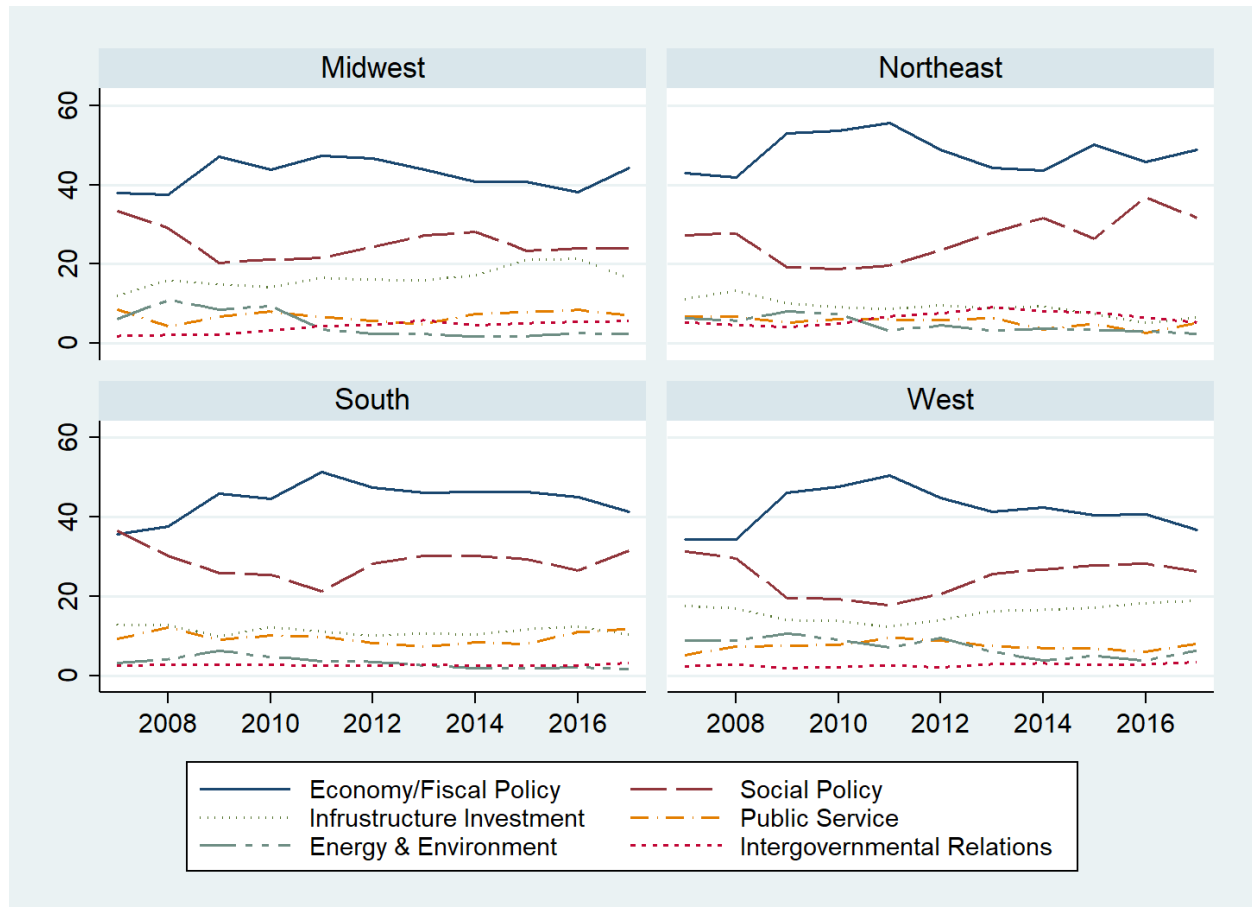
Figure 1.2. Document distribution over topic by region over time.

Since the documents of interest are political statements, it would be interesting to explore the distribution by political ideologies. The summary is broken down by party in Table 1.5. Again, the LDA results indicate a variation in gubernatorial attention to policy topics by political ideology. Democrats were most concerned with health care, followed by economy and budget. Republicans, on the other hand, gave most attention to budgetary issues, followed by education and economic development. The four observations with independent governors come from Alaska's Bill Walker

and two from Rhode Island's Lincoln Chafee.[6] Economy and budgets were top topisc for the

independent governors also, but they focused more than Democrats or Republicans on economic

development. These, of course, are only descriptive statistics, but they nonetheless capture

variation of policy priorities by political party.

Table 1.5. Summary statistics of document distribution over topics by party.

Party	Democrat				Republican				Independent			
Topic label (topic number)	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Economic Develop. (1)	8.7	7.9	0.5	36.8	**9.4**	8.6	0.5	46.6	9.1	3.6	2.8	12.5
Local Government (2)	4.4	6.9	0.3	47.5	3.2	4.2	0.2	29.3	3.6	2.7	1.4	8.8
Healthcare (3)	**12.1**	8.1	0.8	39.5	6.2	5.3	0.5	36.4	4.1	2.3	1.4	7.1
Public Service (4)	6.4	5.4	0.3	33.5	8.6	7.7	0.3	55.6	7.9	6.4	1.4	16.7
Facility Building (5)	3.6	2.3	0.4	11.9	5.4	8.2	0.3	60.4	3.2	1.6	0.9	5
Taxes (6)	4.8	4.5	0.2	48.4	7.3	7.2	0.5	45.3	7.4	8.4	2	24
Child care (7)	4.9	5.7	0.2	57.5	5.8	7.5	0.3	51.3	3.4	1.5	1.1	5
Capital projects (8)	3.8	6	0.3	47.5	3.7	5.3	0.2	43.9	2.5	1.6	0.8	4.6
Transportation (9)	5.5	5.3	0.3	29.8	4.8	5.5	0.3	42.9	5.3	3.5	2.5	11.5
Jobs (10)	8.3	7.6	0.3	38.6	7.5	6.9	0.4	45.3	3.1	2.8	0.4	7.9
Economy (11)	**11.8**	7.8	1	54.4	8.5	7.1	0.3	38.8	**10.7**	2.9	6.3	14
Budget (12)	**9.3**	8.9	0.4	60.2	**12**	9.1	0.5	44.6	**15.2**	7.4	3.4	25.9
Education (13)	7.8	6.1	0.5	35.5	**9.6**	7.2	0.3	35.6	7.3	11.2	1.8	30.2
Public Safety (14)	3.4	6.2	0.2	76.8	3.5	4.6	0.2	40.5	2.5	2.1	1	6.4
Energy/Natural Res. (15)	5.2	4.9	0.2	28.6	4.5	6.1	0.2	41.4	**14.6**	10.3	2	26.8
N	244				300				6			

Overall, the LDA output presents a reasonable picture of the policy items and their

importance to the governors throughout the nation. The topic distribution over words provides a

list of topics mentioned, whereas the document distribution over topics estimates the relative

importance of each item in each speech-year. Previous research analyzed the documents manually

to infer gubernatorial policy concerns. It would be interesting to see how LDA results compare to

---

[6] Governor Charlie Crist from Florida was elected as a Republican and is coded as Republican in this dataset.

manual analysis as a measure of gubernatorial agenda. To further evaluate the adequacy of the LDA output, it is possible to assess its content and construct validity as a measure of gubernatorial policy agenda.

Validity of the LDA results as a Measure of a Gubernatorial Policy Agenda

Content validity refers to "the extent to which a specific set of items reflects a content domain" (DeVellis, 2003, p. 49). To demonstrate content validity, "one must be able to identify clearly the components of the total domain and the show that the [measure] adequately represents these components" (Singleton & Straits, 2010, p. 139). In this study, the labels from the topic distribution over words would ideally match the main items that are on the actual gubernatorial agenda. Once again, the manual analysis results become a very useful reference point (Willoughby, 2008, 2009, 2016, 2017). Table 1.6 below lists the LDA topics alongside the manually identified policy topics.

Table 1.6. Topic comparison between manual analysis and LDA topic model.

	Topics identified manually (Willoughby, 2016)	LDA topics (this study)
1	Education	Topic 13
2.1	Economic development	Topic 1
2.2	Jobs	Topic 10
3	Health care	Topic 3
4	Safety/corrections	Topic 14
5	Tax/revenue initiative	Topic 6
6	Natural resources/energy	Topic 15
7	Transportation/roads/bridges	Topic 9
8	Performance/accountability	Topic 4
9	Surplus/deficit/rainy day funds/reserves	Topic 12
10	Local government	Topic 2
11	Pensions/OPEBs	
12	Transparency	
13	Ethics reform	
14	Debt reduction	
15	Borders/illegal immigrants	
Note: Topic 2.1 and 2.2. were treated as one topic (i.e. Economic development/jobs) by Willoughby (2016), whereas LDA identified economic development and jobs as two separate topics.		

The topics in this table are ordered by how frequently they are mentioned in the State of State addresses. Education, for example, is the most frequently mentioned topic with over 90 percent of governors discussing it every year (Willoughby, 2016). Borders and illegal immigration, on the other hand, is the least discussed topic (only 5 to 10 percent of governors mention it in the addresses) (Willoughby, 2016). It appears the LDA topic model discovered most of the topics discussed in the speeches. LDA discovered all the frequently mentioned topics, whereas the least frequent ones are not picked up well. Pensions and transparency, for example, were mentioned in about 20 percent of the speeches delivered in 2016, and borders/illegal immigrants in about 5 percent of the speeches. Overall, the list of topics identified by the unsupervised machine learning technique corresponds to the manual classification well. The LDA output adequately captures the main policy items on gubernatorial agenda, which renders support to its content validity as a measure of gubernatorial policy agenda.

Construct validity is concerned with "the extent to which a measure 'behaves' the way the construct it purports to measure should behave with regard to established measures of other constructs" (DeVellis, 2003, p. 53). In this case, the accuracy of a prediction, per se, is of less interest than what a relationship reveals about the meaning of the concept being measured (Singleton & Straits, 2010, p. 141). Evidence of construct validity consists of any empirical data that support the claim that a given operational definition measures a certain concept (Singleton & Straits, 2010, p. 142).

The construct validity of the LDA output as a measure of policy agenda depends on whether the distribution of documents over topics represents a meaningful pattern and its correlation with other, theoretically related measures. For instance, the expectation may be that the

share of the speeches devoted to the topic of natural resources will on average be larger in the states that are more dependent on natural resources for revenue. The analysis reveals that about 87 percent of governors dedicated less than 10 percent of their speeches to this topic. Meanwhile, the top 10 percent in terms of attention to the natural resources topic include such states as Alaska, Montana, North Dakota, West Virginia and Wyoming. All these states have historically relied on the natural resources for revenue, be it oil, natural gas, coal, metal, timber or other. Additionally, as seen in Figure 1.3, the attention to natural resources seems to increase during macroeconomic downturns. This may capture growing gubernatorial concern with declining demand for commodities and lower revenue from severance taxes. Hence, increased attention to this topic in these states and across time discovered by LDA makes good sense and the data provide some support to the construct validity of the measure.

It would also be reasonable to expect the governors to talk more about the economy in general during macroeconomic downturns. If the LDA results represent meaningful patterns, governors should spend more time on the economy topic during the Great Recession. Figure 1.3 provides support to this proposition. A sharp increase in gubernatorial attention to the economic issues during the Great Recession is followed by a subsequent decline to pre-recession levels. This further corroborates the notion that LDA measures are correlated with the constructs they should be related to.
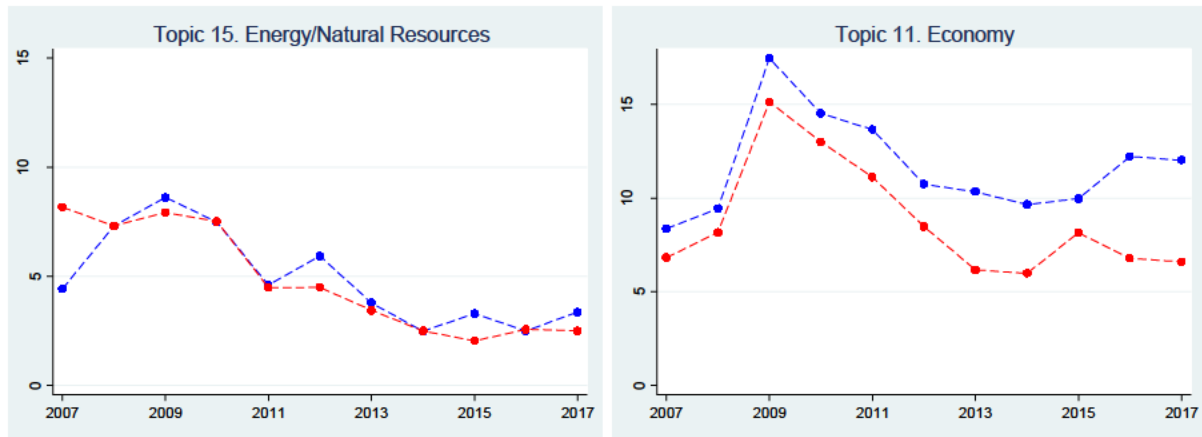
Figure 1.3. Percent devoted to Economy (topic11) and Energy/Natural Resources (topic15) by political party (Republicans in red, Democrats in blue).

Similarly, Figure 1.4 indicates that health care policy (topic 3) was widely discussed prior to the ACA adoption. The attention to this topic decreases with ACA adoption, but slightly increases again when problems with implementation become more obvious in 2013 and 2014. LDA also finds that Democratic governors seem to be more concerned with healthcare than Republicans, whereas Republican governors on average spend more time discussing taxes than Democrats. Both observations seem to fit commonly held perceptions about policy agenda of the parties and render further support to the construct validity of the LDA results.
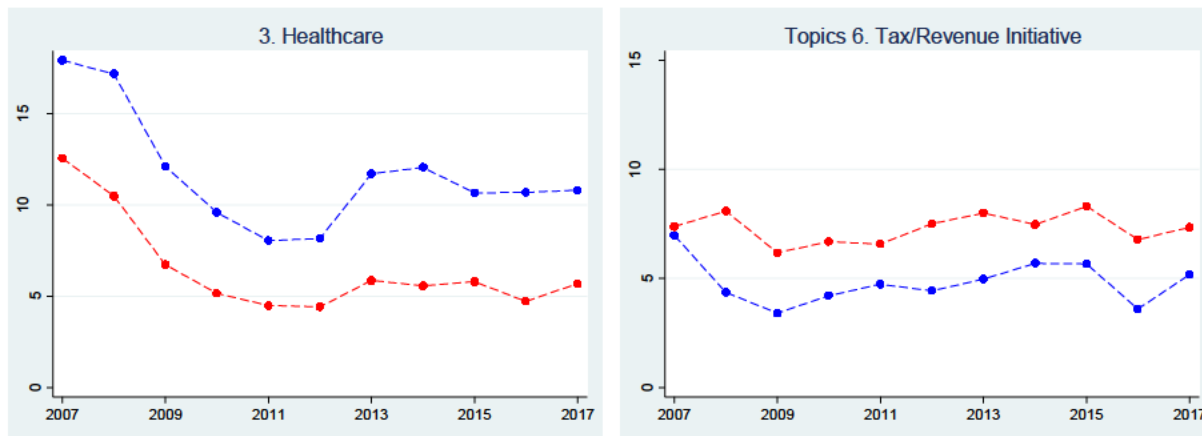
Figure 1.4. Percent of speeches devoted to Healthcare (topic3) and Tax/Revenue (topic 6) by political party (Republicans in red, Democrats in blue).

Overall, this discussion reveals that the LDA results provide a strong estimation of the underlying topic structure of gubernatorial speeches and infer meaningful policy topics from the corpus. The results compare favorably against human coding and have good content and construct validity as a measure of gubernatorial policy goals. LDA appears to be an efficient way to quantify, visualize and annotate large collections of documents. It provides a clear idea about the substantive issues that have been on the executive agenda over the last decade as well as how the attention to these issues changed over time.

Discussion of LDA Topic Model Results and Utility

The aim of this chapter has been to describe a machine learning algorithm for automated text analysis and illustrate its application to more than 500 gubernatorial speeches. The LDA allows one to automatically infer the main themes running through a large collection of documents as well as the share of each document devoted to each theme. Previous research relied on manual coding of gubernatorial speeches, and this study demonstrates how the same task can be performed more efficiently with the help of machine learning algorithms.

One of the conclusions to be drawn from this study is that an LDA topic model provides meaningful data and can significantly enhance our ability to quantify, analyze and visualize large collections of documents. It is especially noteworthy as the original task of the LDA was just to differentiate between documents based on their thematic content (Blei, 2012; Blei et al., 2003). As this study shows, LDA can do more than simple differentiation. Applied to a corpus of 519 gubernatorial State of the State addresses, LDA has generated meaningful data on the main themes running thorough the speeches along with proportions of each speech dedicated to each topic. Taken together, the two distributions represent a good operationalization of gubernatorial policy agendas. It is a more dynamic and nuanced measure than static party affiliations and might be less

prone to subjective judgement that is part and parcel of manual coding. LDA results fare well in terms of content and construct validity. In this case, LDA generated a panel dataset documenting gubernatorial policy agendas over 11 years for all governors. The agendas vary within and between states. This tool enables individual researchers to analyze all states across longer period rather than rely on nonrandom samples or one year of data.

Another important conclusion of this study is that although LDA can be an efficient way to quantify large collections of text documents, prior knowledge of the document content may be necessary. The original purpose of document differentiation by theme can indeed be fulfilled without any content knowledge. However, if the goal is to infer semantically meaningful topics rather than co-occurring clusters of words, researchers should have some prior expectations about potential themes mentioned in the corpus. Previous studies of the State of the State addresses, for example, have been particularly instrumental in validating LDA results in this study. They guided the choice of the number of latent topics, which would be difficult to determine if the perplexity score was the only reference point (Chang et al., 2009). They provided a standard against which the estimated topics could be measured. Had there not been prior research, it would have been much more challenging to validate LDA results for our purposes. It appears that a mixed method approach where automated text analysis techniques are complemented with manual analysis would be the most appropriate. Given the large quantities of documents, this approach can still be much less time consuming.

This leads to another implication highlighted in the previous literature (Grimmer & Stewart, 2013), which is that automated content analysis methods cannot completely replace careful reading of text. Human validation of the results is essential. It was easier to do here due to availability of the results from manual analysis, but how can researchers validate their findings in

situations where prior research is unavailable? It would be reasonable to select a random sample of documents, analyze them manually and use that information to assess LDA results, which ideally would closely resemble the actual thematic structure of the documents. The model can be adjusted by manipulating the number of topics as well as by excluding noninformative frequent words. In this study, *state*, *year*, *will* and state names were among the most frequent, but least useful for discovering the thematic content. Exclusion of these words improved the substantive meaning of the topics and provided more useful results. The performance of the LDA topic model cannot be guaranteed for all types of texts, and adjustments will be necessary. Theory and knowledge of the content are indispensable in the process.

On a final note, LDA and other machine learning algorithms have been gaining popularity in all segments of social activity. Public administration and policy scholars should embrace the trend and adopt the new methods to enhance our abilities to work with large quantities of unclassified data. Automated text analysis opens up vast opportunities for systematic analysis and extraction of numeric data from large collections of text documents. LDA has been successfully applied to collections exceeding a million of documents (Jelodar, Wang, Yuan, & Feng, 2017). This tool can be useful for public sector practitioners in sorting almost any collections of electronic documents by their thematic content. Without having to read them, one could sort the document by topic and go directly to those that contain the topic of interest. With public bureaucracy generating large quantities of text documents, this tool can save time and maybe money.

CHAPTER 2

GUBERNATORIAL POLICY PRIORITIES AND STATE FISCAL OUTCOMES

Introduction

The main goal of this study is to test empirically the ability of American governors to influence state fiscal choices. The governors are expected to set policy goals and lead the state forward in a way available to hardly any other individual state official (Rosenthal, 2012). State chief executives, however, lack direct legislative authority and are therefore substantially disadvantaged in the realm of lawmaking vis-à-vis legislators. The theoretical framework developed in the extant literature and adopted in this study postulates that governors nonetheless possess an array of formal and informal tools that may help overcome this institutional disadvantage. They can publicly praise or criticize individual lawmakers, appoint political allies, and threaten to veto or promise to sign policy bills in exchange for legislative support - thereby exerting influence on state lawmaking.

Previous studies testing these theoretical propositions often employed rather crude measures as proxies for gubernatorial preferences, such as party affiliations (Alt & Lowry, 1994; Barth & Ferguson, 2002; Dye, 1966, 1984; Garand, 1988; Hofferbert, 1966; Kousser, 2002; M. A. Smith, 1997; Winters, 1976) or size of recommended budgets (Kousser & Phillips, 2009). They also used nonrandom samples of the states (Kousser & Phillips, 2012). As a result, many of the findings throughout the extant literature are inconsistent and have limited external validity.

This study seeks to fill several gaps. First, it uses a panel data set on all 50 states covering a nine-year span from 2007 to 2015, which should be an improvement in terms of internal and

external validity. Second, unlike previous studies, this research infers gubernatorial policy priorities from the transcripts of high profile State of the State addresses using unsupervised machine learning. This technique permits a more nuanced measurement of gubernatorial policy goals, which vary significantly across time and between governors. Third, this study uses more granular data by focusing on state spending by main function. This approach should improve our understanding of gubernatorial influence on the distribution of public resources by specific policy areas, rather than the total size of state budgets. The regression results presented in this chapter suggest that although governors are indeed powerful players in state policymaking, their influence is not equal in all policy areas. This research shows that state chief executives are more likely to exert significant influence on state budgetary choices in the area of infrastructure and public safety. State budgetary choices in other policy areas are not significantly associated with the executive policy agenda.

The paper is organized as follows. The theoretical framework is laid out in the next section. It is followed by a review of the relevant literature. The research design in section four describes the data, variable operationalization and estimation methodology. Section five presents empirical findings. Finally, section six presents a discussion of the empirical findings and section seven draws pertinent conclusions.

Theory of Interbranch Bargaining

American governors are ideologically flexible (Willoughby, 2014, p. 176), but "whatever the policy preferences of the chief executives, they generally want more power to impose those preferences on the budget" (Rubin, 2016, p. 11). This endeavor becomes increasingly challenging in contemporary public budgeting processes characterized by intense partisanship and gridlocks to the point where government shutdowns have become more common and last longer (Rubin, 2016,

p. xix). Given this environment combined with their lack of direct legislative authority, how might governors pursue their policy goals? Under what conditions might they be more successful?

This research adopts a theoretical framework of rational choice developed in the previous studies of gubernatorial powers (Gray, Hanson, & Kousser, 2017; Kousser & Phillips, 2012; Kousser & Phillips, 2009; Sigelman & Dometrius, 1988). The theory describes how the state budgeting process should unfold. The core assumption of this framework is that state chief executives and legislators are two types of rational actors seeking to achieve their diverse policy goals, but neither has complete control. The process proceeds through bargaining and negotiations. Whether they want to pass a budget or a policy bill, governors depend on the legislators who, in turn, may be interested in side payments and favors that governors can deliver in return for cooperation.

The office of state chief executive provides an array of formal and informal tools that can be utilized to a governor's advantage. All governorships are well-paid, full-time jobs, whereas many state lawmakers are formally part-time, and some legislative sessions may formally last only a short period of time. Many legislatures are also not as well staffed as state chief executives. Governors thus may have more analytical support and time to develop an effective bargaining strategy. They also are the most prominent individual state officials with ample opportunities to address wide audiences. Their prominence, of course, may vary from state to state depending on how many other state executive branch officials are elected directly by voters. Citizens in various states elect their lieutenant governors, secretaries of state, attorneys general, auditors, commissioners of insurance, treasurers and other officials (Council of State Governments, 2017). Nonetheless, experienced governors remind newcomers that this "bully pulpit" is one of the most important resources to achieve their priorities (National Governors Association, 2015). "When a

32

governor speaks, whether to the press, the legislature, on the road or through responses to constituent inquiries, focus is immediate" (National Governors Association, 2015, p. 77). The inaugural address, State of the State address, and budget message can be "incredibly useful and powerful communication tools" (National Governors Association, 2015, p. 77).

State chief executives can use this visibility to cultivate support for reelection campaigns and fundraising efforts of individual lawmakers. Seasoned governors note that this is done in several ways, from inviting legislators to parties at the executive mansion to joint appearances in electoral districts (National Governors Association, 2010, p. 25). Some also gave legislators departmental and agency bills to sponsor so that they could claim credit (Rosenthal, 2012, p. 128). Most of the states authorize state chief executives to reorganize state agencies through executive orders (Council of State Governments, 2017). Governors can use their reorganization and appointment powers to give jobs to legislators' political allies, especially at the beginning of an administration and after reelection. Uncooperative lawmakers and their campaigns, on the other hand, can be publicly attacked and criticized. Governors have the flexibility of choosing to sign certain bills that are important to lawmakers, or to veto them. By these means state chief executive can make legislators' "lives happier and their careers brighter" (Rosenthal, 2012, p. 124), or vice versa.

The ability to deliver favors and credibility of their threats will depend on gubernatorial political capital, which is a function of their popularity, veto powers and time in office. Lower popularity, holding all else equal, will make joint appearances with governors less beneficial or even detrimental. Public criticism is also less likely to have an intended effect when governors have low approval ratings. Consequently, less popular chief executives may have fewer favors to offer and will be less influential in the budgeting game.

Veto power is an essential institutional factor. North Carolina was the last state to grant its governors a veto power in 1996 (Kousser & Phillips, 2012, p. 37). Today 44 governors possess the line item veto power, some on all bills, and others on appropriations only (Council of State Governments, 2017). Veto powers vary substantially from state to state, but they generally enable state chief executives to cut funding for specific pet projects of the lawmakers without risking the entire bill and should give them more bargaining power (Abney & Lauth, 1985, 2002). Veto can be overridden, and the credibility of the veto threat will also depend on institutional rules and partisan composition of the legislature. Some states require simple majorities to overrule the veto, some 3/5 or even 2/3 of the legislature. The higher the threshold, the more credible the veto threat.

Previous studies also found that "executive success wanes over time. This supports the traditional 'bank-account' theory of chief executive clout" (Ferguson, 2003, p. 172). Legislators may foresee long relationships and be more considerate of gubernatorial priorities during the first term as governors will have plenty of time to pay back for (un)cooperation. Hence, governors may be more successful when it is their first time in office and less successful when serving their last consecutive term.

This theoretical framework describes the main strategies state governors can utilize to overcome the institutional disadvantage in lawmaking vis-à-vis legislators. With these tools at their disposal, a gubernatorial policy agenda can be carried out through policy-focused legislative or budgetary bills. Admittedly, a policy bill incorporating executive priorities would have a more lasting effect on state polices. It is expected, however, that state chief executives have less leverage over policy bills and might be more successful with their policy agenda in the budgeting game for several important reasons. The next subsection outlines this aspect in more detail.

Budgeting Game versus Policy Initiative

Fiscal, political and legal realities require both branches to come to the negotiation table (Kousser & Phillips, 2009). Only eleven states have mechanisms for passing a temporary budget in the situation that a particular year's budget is late (Rubin, 2009, p. 90), but even these measures cannot become a permanent replacement for a new budget and cut deeply into approval ratings of both branches (Kousser & Phillips, 2009, p. 57). In most states a budget delayed past the start of the fiscal year triggers an automatic shutdown (Kousser & Phillips, 2012, p. 31). Thus, in case of budgetary negotiations, both branches bear the cost of a delayed agreement and have considerable incentive to compromise. In addition to political costs, there are substantial financial costs because late budgets significantly increase the cost of state borrowing (Andersen, Lassen, & Nielsen, 2014). The municipal bond market views late budgets as a sign of bad fiscal governance, thus raising concerns in the investment community about stable repayments of outstanding obligations. When repayments become less certain, investors require a corresponding compensation for higher risks. Consequently, neither side can afford inaction in the budgeting realm.

The consequence of failing to come to an agreement and pass a governor's policy bill, on the other hand, means that state policy simply retains the status quo. In the typical case there are no noticeable interruptions in government operations. Governors are left with take-it-or-leave-it option in this realm or, at least, a weaker negotiating position (Kousser & Phillips, 2012, p. 10). In addition, negotiations and bargaining for a policy bill may continue over several years. In such a

situation not being able to succeed in the first legislative session does not necessarily mean failure for a state administration.[7]

Another advantage to governors stems from the fact that most states use some form of executive budgeting, where governors have constitutional responsibility for formulating the budgets proposals (Rosenthal, 2012, p. 95). In this model, the governor's office provides budget instructions to the agencies, collects their proposals, trims them as necessary and forwards the whole budget as a recommendation to the legislature for consideration (Rubin, 2016, p. 128). By virtue of the fact that governors create the first draft, their policy priorities form the basis for the allocation of state resources and set the initial framework for budget deliberations. Depending on fiscal constraints, governors get to make the initial decision as to what policy initiatives and programs receive more, less or no funding.

Beyond setting the initial parameters on who gets what, executive budgeting enables governors to include pork projects of legislators in their proposals (or in supplemental appropriations) in order to cultivate legislative support for the whole package (Rubin, 2016, pp. 104-105). Taken together, these factors improve the gubernatorial bargaining stance in the budgeting process compared to negotiations over policy bills. Indeed, the National Governors Association advises the governors to use budgets strategically as an important vehicle for achieving key gubernatorial objectives (2015, p. 65). The relative balance of budgeting powers between the two branches varies, but most states lie between the extremes of executive or legislative dominance with an average governor having more authority than legislators (Rubin,

---

[7] It should be noted that logrolling, by connecting or separating issues, can make the policymaking process easier.

2016, p. 126). The literature review in the next section provides some evidence about the impact governors with more authority have on state policies.

Thus, this theory of interbranch bargaining suggests that generally governors possess several formal and informal tools to overcome this institutional disadvantage in the lawmaking process. Fiscal, political and legal realities also imply that governors might be more successful with their policy agenda in the budgeting process. Given this framework, the expectation is that gubernatorial policy agenda will have a significant effect on state budgetary outcomes. Hence, the main hypothesis of this study for empirical testing is:

Hypothesis: There is a statistically significant relationship between gubernatorial policy priorities and state budgetary choices.

Gubernatorial Agenda and Temporal Ordering

Governors are flexible in terms of their policy agendas so their party labels may not be particularly meaningful (Kousser & Phillips, 2012, pp. 97, 144; Rubin, 2016, p. 11). They have a wide range of options when it comes to their agenda items (Rosenthal, 2012, p. 88). The origins of policy items and the exact moment of agenda formation is impossible to trace. Most governors state that the promises and commitments made during electoral campaigns are the primary source of their first year's policy goals, and that their agenda in general is largely a function of personal experience, values and political philosophies (Rosenthal, 2012, pp. 100-105).

Whatever the source, "it is popularly believed that agenda completion [for the governor] occurs with the State of the State address" (Rosenthal, 2012, p. 113). This is the highest-profile event of the year for the chief executives who get a chance to speak to both chambers of state legislatures. The event takes place at the outset of the legislative sessions in all states. State officials expect the governors to realize their policy leadership by articulating the issues that need attention

and by setting the tone for the upcoming legislative session. The process of drafting the speech starts weeks in advance, and, due to time limitations, it focuses on major items (Kousser & Phillips, 2012, pp. 74-75). Many governors believe their successes are measured by how many of the State of the State items get passed (Rosenthal, 2012, p. 91). As such, the speech represents a serious commitment and they need to finalize their priority list by the time of the speech. Around that time, governors are expected to complete and present their budgets to the legislatures. Therefore, the establishment of clear priorities as a prerequisite to budgeting and delivering the speech is paramount (National Governors Association, 2015, p. 65).

Consequently, for the purposes of this work, it is assumed that gubernatorial agenda formulation precedes the submission of the recommended budget and the address. The policy priorities are then built into the recommended budget and divulged to the public in the State of the State addresses. The executive agenda is an antecedent of both the speech and the budget. This also suggests that the State of the State should be a reasonably good gauge of the executive agenda in any given year. Indeed, studies show the topics most frequently mentioned in the addresses (Willoughby, 2008, 2009, 2016, 2017) match almost one to one to the actual priorities highlighted by the governors in their responses to a survey (Rosenthal, 2012, pp. 98-99).

The main aim of this study is to test whether these priorities from the State of the State addresses have any effect on the budgetary choices as adopted by the legislatures. The submission of the recommended budget is only the first step. It has no effect on the actual distribution of resources unless approved by the lawmakers. And this is where the real test of executive dexterity with sticks and carrots begins. If the main hypothesis holds, there should be a significant correlation between gubernatorial policy priorities and adopted budgetary choices. This would indicate an effective use of formal and informal tools to pursue their agenda. There is some

evidence from the previous literature that the gubernatorial agenda may have a significant effect on the size of state budgets (e.g. Barrilleaux & Berkman, 2003; Kousser & Phillips, 2012; Kousser & Phillips, 2009). One of the contributions of this study is to test whether individual executive priorities are correlated with budgeting choices by functional areas rather than with the total size of state budgets.

A gubernatorial agenda may comprise any number of policy items. However, analysis of their State of the State addresses reveals that throughout the last decade most governors have been primarily concerned with the following items: education, health care, public safety, natural resources and infrastructure (Willoughby, 2008, 2009, 2016, 2017). Hence, these items are selected for this study. This choice ensures that the study is focused on issues that vary in terms of importance but nonetheless are applicable to all governors throughout the nation. This list also corresponds closely to the main functional areas of state governments as classified by the United States Census Bureau (US Census, 2018).

Since the State of the State address is limited in time, it is reasonable to assume that governors will devote a larger share of the speech to the items that are more important. It is also reasonable to assume that governors will be more concerned with securing funding for the items that are more important. Of course, some policy choices may take little or no resources and require minimal discussion. But, on balance, it is assumed that governors want to be remembered for their achievements, and almost any effort to make a change in the society takes resources. Besides their personal aspirations, once the public commitment is made, chief executives often will be judged by how many of the goals from the address have been achieved (Rosenthal, 2012, p. 91). Arguably, some policy goals need little resources and/ or a shorter public statement, whereas others may need substantial funding

Hence, if the theoretical framework provides a realistic description of reality, individual policy priorities may be correlated with the level of state spending by function. The more important an item is to the governors, the more likely they are to seek more funding. More specifically, this proposition leads to the following hypotheses:

- **Hypothesis a**: There is a significant association between gubernatorial policy priorities in education and the level of state spending on education.

- **Hypothesis b**: There is a significant association between gubernatorial policy priorities in health care and the level of state spending on health care.

- **Hypothesis c**: There is a significant association between gubernatorial policy priorities in infrastructure and the level of state spending on infrastructure.

- **Hypothesis d**: There is a significant association between gubernatorial policy priorities in public safety and the level of state spending on public safety.

- **Hypothesis e**: There is a significant association between gubernatorial policy priorities in natural resources and the level of state spending on natural resources.

A priori, an expectation here is that it is likely the level of spending for some of the categories will be determined by law, be it by statutory provision or judicial precedent, rather than by the annual budgetary process. For example, the judiciary has effectively required more funding for several state prisons ruling that overcrowding was unconstitutional. The Abbot cases in New Jersey haven affected state education expenditures for at least two decades now. Similar cases have had an impact on education spending in Kentucky, Washington and other states. The level of state spending in these examples was defined by court decisions leaving less space for bargaining and policy priorities of government officials. Much of healthcare spending is devoted to Medicaid, where eligibility is stipulated by law and fluctuates with socio-economic conditions, not policy

preferences. However, no two states are the same, and the main question here is whether an average American governor can exert significant influence on state budgetary choices, and if so, in what areas.

This study seeks to shed more light on what state chief executives ask for in their public agendas and what they get. The main goal is to expand our understanding of gubernatorial powers in state policymaking by providing new empirical insights and contributing to the existing literature.

Literature review

The research question explored in this study is not new, and the power of state governors has been analyzed in previous research. Comparative case studies date back as far as the 1930s (Lipson, 1939). More recently, scholars have employed different techniques to test gubernatorial influence in state policymaking processes. Some early surveys of state department heads indicated that they perceived legislatures, not governors, as having more impact on programs and objectives of state government departments (Abney & Lauth, 1983). Subsequent surveys of executive and legislative budget officers first found that the appropriation process was dominated by governors (Abney & Lauth, 1987), but later suggested a perceived decline of gubernatorial dominance and a relative parity between the executive and legislative powers in the appropriation process (Abney & Lauth, 1998). These findings led the authors to conclude that executive dominance present since the early part of the 20th century had been declining by the end of the century. Other scholars also found a more affirmative role of state legislatures vis-à-vis the governors in general fund budget matters during that period (Thompson, 1987).

Dometrius and Wright (2010), on the other hand, surveyed state agency leaders around the same time and found no consistent decline in gubernatorial influence on agency budgets. Instead,

41

their results suggest that on average governors slightly increased their influence vis-à-vis legislatures from the 1980s through the 1990s (Dometrius & Wright, 2010, pp. 783-784). Ryu et al. (2008) corroborate these findings with their survey of state administrators, who stated that governors dominate the budgeting process. Another survey of legislative and executive budget analysts in Western states in 2000-2001 also found governors, not legislatures, to be the main beneficiaries of any influence shifts (Goodman, 2007). There is also some evidence that power between the two branches is not a zero-sum game, where one institution is gaining influence at the expense of the other (Dilger et al., 1995). This divergence of findings with respect to what branch dominates the legislative process may be due to different operationalization of key variables and in some cases because of reliance on perceptive measures. Many studies also use nonrandom samples of the states.

More detailed analysis reveals that gubernatorial effectiveness is a function of several factors, and that both institutional and environment considerations matter (Dilger et al., 1995). Governors with more institutional powers in the budgetary process have been found to increase the type of spending that confers statewide benefits and aids their political ambitions (Barrilleaux & Berkman, 2003) and may increase the extent to which state budgets are punctuated irrespective of legislative constraints and party control (Breunig & Koski, 2009). There is some evidence that state budgets grow at higher rates with increasing governors' unilateral control over fiscal policymaking, which implies that more balanced budgetary powers between two branches may be necessary to restrain excessive spending (Krause & Melusky, 2012). Other studies (Alm & Evers, 1991; Holtz-Eakin, 1988) show that executive veto powers may have a negative, but weak impact on state overall level of spending and its composition. Generally, gubernatorial success with legislative bills can decline in their last year of service, during economic slowdowns and when

dealing with divided government, but might increase with more experience, staff support and higher overall bills' passage rate in the legislature (Ferguson, 2003). Researchers also highlighted the mediating effect of legislative professionalism on the power of chief executives (Kousser & Phillips, 2009).

The study by Kousser and Phillips (2012) perhaps comes closest to this research. The authors tested the influence of the gubernatorial policy agenda on budgetary outcomes and policy bills in a nonrandom sample of 28 states over two legislative sessions – 2001 and 2006. The policy agenda items were manually inferred from the State of the State Addresses. Their findings suggest that governors are more successful with budgetary than policy proposals due to their more powerful formal stance in the budgeting negotiations. The main limitation of this study is that the external validity of the findings is limited.

To summarize, whereas there have been multiple attempts to understand the executive influence in state policymaking, most of the extant research has been either cross-sectional or used only a few years' worth of data. The internal and external validity of the available evidence is thus unclear. This study should be an advance because policy priorities are documented over nine recent years for all 50 states using computational text analysis. Unlike cross-sectional studies, the results should provide useful information in terms of longitudinal analysis and should be generalizable to all U.S. states.

Research design

Data

In order to test the hypotheses developed in the theoretical section, this study utilizes a balanced panel dataset on all fifty states from 2007 to 2015 (N=450, n=50, t=9).[8] The main variables along with data sources are described below. All dollar amounts were adjusted for inflation to 2017 constant values using the Consumer Price Index for all urban consumers from the Bureau of Labor Statistics. The descriptive statistics are presented in Table 2.1.

Table 2.1. Descriptive statistics.

Variables	Mean	SD	Min	Max
Policy Topics (% of the speech)				
Health Care	8.8	7.3	0.5	39.5
Transportation	5.1	5.4	0.3	42.9
Education	8.8	6.8	0.3	35.6
Public Safety/Law	3.5	5.3	0.2	76.8
Energy/Natural Resources	4.9	5.7	0.2	41.4
Dependent Variables (dollar amount)				
Education	12,500	14,200	967	95,600
Health	2,681	3,512	80	21,400
Highways	2,411	2,378	229	15,000
Public Safety	1,367	1,744	81	12,300
Natural Resources	476	697	43	7,045
Controls				
GDP (per capita in $1,000)	47.3	8.9	31.2	73.5
GDP change (% change)	0.3	2.5	-9.2	19.3
Unemployment (% labor force)	6.3	2.2	2.2	13.7
Unemployment change	0.0	1.3	-2.2	5.7
End Balance (est. current FY, bln)	0.5	1.2	-4.3	13.5
Population (Ln)	6,263.7	6,943.1	534.9	39,355.2
Population change (%)	0.7	0.6	-0.7	3.1
Population white (%)	77.7	12.6	24.2	96.1
Population over 65 (%)	14.2	2.0	6.9	20.1
Governor Democrat (yes=1, else=0)	0.4	0.5	0.0	1.0
Governor Independent (yes=1, else=0)	0.0	0.1	0.0	1.0
Governor Male (yes=1, else=0)	0.9	0.3	0.0	1.0
Governor Female (yes=1, else=0)	0.1	0.3	0.0	1.0
First Term (yes=1, else=0)	0.5	0.5	0.0	1.0

---

[8] The time span is limited by the scarcity of speeches before 2007 and absence of some of the more recent financial data after 2015.

Last Term (yes=1, else=0)	0.3	0.4	0.0	1.0
Popularity (%)	56.2	7.7	36.1	77.6
Legislature Democrat (yes=1, else=0)	0.4	0.5	0.0	1.0
Legislature Split (yes=1, else=0)	0.2	0.4	0.0	1.0
Legislature Republican (yes=1, else=0)	0.5	0.5	0.0	1.0
Governor-Legislature Same party (yes=1)	0.6	0.5	0.0	1.0
State Liberalness (% vote in presidential elections)	48.0	10.0	21.9	71.9
Agenda Diversity (HHI)	0.14	0.05	0.08	0.60

Note: Dependent variables are in $1,000,000s and are not log-transformed in this table for illustrative purposes but were logged for regression analysis. Topics are summarized for 2007 – 2015 period.

Dependent Variables

The outcome variables were extracted from the U.S. Census Annual Survey of State Government Finances. As has been noted, the main goal of this study is to test the relationship between gubernatorial policy priorities and state budgetary choice in the main policy areas, such as education, health care, public safety, natural resources and infrastructure. To accomplish this, the U.S. Census data on state spending by functional areas is used as dependent variables, namely expenditures on education, health, hospitals, highways, police, corrections and natural resources.[9] Heath and hospital expenditures are both related to health care policy and are combined into one category. Likewise, police and correction expenditures are combined as both are related to public safety. This gives five dependent variables each representing the total level of state expenditures by distinct policy area.[10] All dependent variables were log-transformed to alleviate positive skew.

On average, these functions take up more than 55 percent of state general government expenditures, ranging from 40 to 70 percent. Most of the remaining budget is devoted to public

[9] Elaborate Census definitions of these spending categories are presented in Appendix B.
[10] Spending per capita was used to check the robustness of the findings. The main findings were the same, but omnibus fit statistics (i.e. adjusted-R^2 and log-likelihood) indicate that per capita models have worse fits. Hence, models with the total level of spending, controlling for population, are reported.

welfare, which absorbs about 30 percent of an average budget and ranges from 15 to 41 percent.[11]

Other minor categories include parks and recreation (0.4 percent), government administration (3.6 percent), interest on debt (2.8 percent) and other miscellaneous "unallocable" expenditures combined (9.2 percent). Thus, the outcomes used here represent most of the state expenditures determined through the budgetary process. Given the theory laid out earlier, the main expectation is that the policy areas of higher importance to the governor will receive more funding, all else being equal.

Independent Variables

Unlike previous research, which often inferred an executive agenda from party affiliations, gubernatorial policy priorities for this study were extracted from transcripts of the State of the State Addresses using an unsupervised machine learning technique. Most of the transcripts came from the National Governors Association, but the transcripts for most of earlier years were obtained from individual state websites. Gubernatorial agendas do not always toe party lines and executive policy goals may be hard to predict without knowledge of the individuals (Rubin, 2016, p. 11). Party affiliations are also time invariant and may not capture changes in gubernatorial agendas over time. Using the State of the State addresses as an indicator of the executive policy priorities allows the policy goals to vary across states and time. This study uses the results of the Latent Dirichlet Allocation topic model generated in the first chapter.

This probabilistic topic model allows one to infer the underlying topic structure that most likely generated the observed collection of words. It represents each topic as a probability

---

[11] Public welfare spending is largely determined by formula and therefore is not expected to fluctuate with policy priorities.

distribution over words, and each document as a probability distribution over topics. This second distribution estimates the share of each address devoted to each of the policy topics. The proportions of each address devoted to each of the five policy areas used in this study, namely education, health care, transportation, public safety and natural resources, are the main predictors of interest. Each of these variables has a theoretical range from 0 to 100 percent of the speech. It is assumed that a governor devotes more attention to the policy areas that are more important. Thus, the LDA allows one to estimate the relative importance of each policy topic in each year. [12] The descriptive statistics of these variables are summarized in Table 2.1.

An important note here is that the analysis was structured so that gubernatorial speeches always precede budget deliberations. The vast majority of speeches are delivered in early January and February, with a handful delivered in March. The fiscal year begins on the 1st of July in 46 states, on the 1st of April in New York, on September 1st in Texas, on October 1st in Michigan and Alabama (NCSL Research, 2016). In all cases, the speech precedes the beginning of legislative sessions where budgets are adopted. By ensuring the data for the independent variable precedes the dependent variable, a temporal ordering is retained to help clarify causality in the tested associations.

Control variables

The theoretical framework suggests that governors' success may vary with several factors necessitating a set of the following controls. According to our theory, the number of favors and credibility of threats may vary with the popularity of the chief executive. A Governors' popularity

---

[12] It should be noted that in some states with biennial legislatures the Address is delivered every other year. The models were tested with these state-years dropped from the analysis as well as with speeches extended for two years. The main findings are not affected.

is measured as a percent of votes received during elections. It varies from 36 to 78 percent, with a mean of 56 percent. Some are relatively unpopular with only a third of voters supporting them whereas others start off with almost 80 percent of support. The disadvantage of this measure is that it remains constant throughout the term, but it is the only proxy for popularity available for all states for the period of the study.

Another important factor is gubernatorial veto powers, which can make a governor a more (or less) imposing opponent. Some governors can apply a line-item veto to all bills (20 percent), some to appropriation bills only (66 percent), and some have no line-item veto powers (Indiana, Nevada, New Hampshire, North Carolina, Rhode Island, Vermont). Those able to apply the line-veto to all bills should have the most bargaining power, and those without the line-item veto the least (the omitted comparison category). The veto override requirements also vary from a simple majority (comparison category), to 3/5, to 2/3. All of these, however, are time-invariant, and because the fixed-effects within estimation was the appropriate approach for the data, the effect of these factors was not estimated.[13]

As stated in the theoretical discussion, a governor's bargaining power may vary with the amount of time s/he has spent in office. This is accounted for with two dummy indicators. The first is coded 1 if governor is in the first year of office, otherwise it is coded 0. The other is coded 1 if the governor cannot be reelected for another term due to reaching the maximum number of consecutive terms in this state, otherwise it is coded 0 (i.e. legacy term, or "lame duck"). Those in between serve as the comparison category.

---

[13] With this estimation approach, their effect is purged along with other time-invariant state-specific characteristics.

The models include controls for governor's party affiliation (Republican is the omitted comparison category) and gender (male is the omitted category). Partisan control of the legislature may be related to gubernatorial success and is measured by what party controls both houses (Republican control is the omitted category), or whether the control is split between the two parties.[14] Another dummy is coded 1 when the governor is of the same party as the one that dominates the legislature.

The next set of controls accounts for socio-demographic factors that are known to influence state fiscal policies (population, population change, percent white, percent population over 65), all provided by U.S. Census Bureau. State macroeconomic conditions may influence the amount of resource states have and ease or complicate the bargaining process. These are controlled for with Gross Domestic Product (GDP) per capita,[15] GDP change, unemployment rate and its change relative to the last year.

Budgetary decisions are made in light of the state fiscal health. To account for that the models control for estimated end balance of the budget in millions of dollars.[16] This estimate gives decision makers an idea on how the current FY is going to end and may influence their preferences for higher or lower spending. Governors may have harder time pushing their agenda through the legislature when states are facing deficits and running low end balance. Citizen ideological leaning may be related to state distribution of resources. State liberalness in this study is proxied by the

[14] This category included nonpartisan unicameral legislature in Nebraska.
[15] The other two variables, state poverty level and personal income were highly collinear with GDP and were dropped.
[16] The estimated end balance represents the estimated difference between expenditures and revenues at the end of the current fiscal year. Due to economic or other fluctuations, the actual end balance may deviate from an adopted budget. These data were extracted from Fiscal Survey of the States of the National Association of State Budget Officers.

percent of votes cast for a Democratic candidate in the most recent presidential elections. This measure, like gubernatorial popularity, remains constant over 4 years until next election but is the best available proxy measure for the period of the study.

Another control variable is the diversity of the gubernatorial agenda. It could be important because gubernatorial success may depend on focus. A diverse agenda may dilute public and legislative attention and "may overtax the governor's limited political capital" (National Governors Association, 2015, p. 57). Focusing on a limited range of issues, perhaps from three to five, may improve the chances of success, whereas a diverse agenda may be detrimental. Agenda diversity in this study is measured using a Herfindahl-Hirschman Index (HHI), which is equal to the sum of squares of all shares of the speech. For example, if only two issues were discussed in a speech and each took exactly a half the speech, the HHI would be 0.5 (i.e. $0.5^2 + 0.5^2$). If four issues were discussed and each took a quarter of the speech, the HHI would be about 0.25 (i.e. $0.25^2 + 0.25^2 + 0.25^2 + 0.25^2$). If two issues were discussed and one took 90 percent of the speech and the other only 10 percent, the HHI would be 0.82 (i.e. $0.1^2 + 0.9^2$). Higher values of the index represent a more concentrated agenda, and lower values mean a less focused agenda.

The preceding list of control variables represent the common factors that might affect spending in all categories. Each of the functional areas may be affected by additional factors. Education spending, for example, may vary with the level of school enrollment. This variable, however, was highly collinear (0.996) with the population variable and therefore omitted. The number indigents and retirees may drive state health care expenditures. These factors should be accounted for with the common variables, such as the percent of population over 65 and percent unemployment measures. The length of public roads managed by the states is included in the model of highway spending. Public safety spending may be correlated with crime rate in the state. The

intention was to control in the public safety models for the number of violent and property crimes per 100,000 population. However, these variables tend to be collinear (0.6). Whereas the results are the same with either variable, the omnibus fit statistics were better with the property crime rate. Hence, the property crime rate per 100,000 is used in the public safety model (logged to alleviate positive skew). Natural resources spending may be correlated with the rate of state reliance on the severance taxes for revenue. Hence, severance taxes as percent of total tax revenue is included in the natural resources model.

Estimation Methodology

This is a balanced panel dataset with continuous outcome variables. A Hausman specification test indicates that the data do not support a random effects model. Hence, fixed-effects within estimator is used. The advantage of this approach is that it accounts for any unobserved time invariant individual state characteristics that might otherwise contaminate the findings. The shortcoming, as it has been mentioned, is that the effect of time invariant variables cannot be estimated. Standard errors are clustered by state in all models to correct for potential error heteroskedasticity and autocorrelation.

Empirical Findings

The empirical findings are reported in Tables 2.2 and 2.3. The results from the fixed-effects within estimation suggest that state spending by functional areas varies with economic prosperity, population size and slack resources. Political composition of state governments also is an important factor. Politically independent governors (n=4), for example, may be associated with more spending on education and natural resources, but less spending on health care compared to Republican governors. Democratic governors are associated with marginally higher spending on public safety relative to Republicans. Democratic control of both houses of state legislatures is

significantly correlated with higher spending on education and public safety. State liberalness is

positively association with highway spending.

Table 2.2. Regression results for education, healthcare and highway expenditures.

	Education		Healthcare		Highway	
Gubernatorial Policy Priority	0.000	(0.90)	-0.001	(-0.79)	**0.005****	**(3.38)**
GDP per capita	0.005	(1.49)	0.028***	(3.66)	0.021**	(3.06)
GDP change	-0.001	(-0.58)	-0.009*	(-2.20)	-0.005	(-1.36)
Unemployment	-0.005	(-0.96)	-0.006	(-0.32)	0.021	(1.24)
Unemployment change	0.002	(0.44)	0.002	(0.11)	-0.009	(-0.63)
Estimated End Balance	0.006	(1.28)	-0.003	(-0.65)	0.010	(1.42)
Population	0.973**	(3.03)	0.627	(0.60)	-0.612	(-0.69)
Population change	0.008	(0.60)	0.003	(0.12)	0.027	(1.00)
Population white	-0.005	(-1.37)	-0.001	(-0.07)	0.004	(0.38)
Population over 65	-0.016	(-0.62)	0.058	(0.85)	-0.041	(-0.92)
Governor Democrat	0.017	(1.40)	-0.005	(-0.13)	-0.013	(-0.44)
Governor Independent	0.055**	(3.33)	-0.173***	(-4.34)	0.034	(0.42)
Governor female	-0.029	(-1.08)	-0.033	(-1.28)	-0.001	(-0.02)
Governor 1st term	-0.009	(-0.72)	-0.040	(-1.25)	0.057*	(2.15)
Governor last term	-0.009	(-0.48)	-0.083*	(-2.22)	0.077[+]	(1.87)
Governor popularity	-0.000	(-0.19)	0.000	(0.03)	-0.001	(-0.52)
Democrats control both houses	0.026*	(2.02)	0.088	(1.56)	0.040	(0.85)
Legislature split control	0.005	(0.35)	0.009	(0.18)	-0.019	(-0.55)
Governor-Legislature same party	0.008	(1.11)	-0.029	(-0.70)	-0.019	(-0.63)
State liberalness	-0.002	(-0.54)	-0.002	(-0.30)	0.016*	(2.08)
Agenda diversity (HHI)	0.000	(0.00)	-0.398	(-1.29)	0.116	(0.64)
State public road length					-0.499[+]	(-1.76)
Constant	1.610	(0.32)	2.910	(0.19)	26.518[+]	(1.93)
N	450		450		450	
R^2 within	0.40		0.23		0.27	
R^2 between	0.96		0.78		0.63	
R^2 overall	0.96		0.77		0.60	
Note: t-statistics in parenthesis; [+] $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$.						

Table 2.3. Regression results for Public Safety and Natural Resources.

	Public Safety		Natural Resources	
Gubernatorial Policy Priority	**0.002****	**(2.89)**	0.001	(0.45)
GDP per capita	0.013***	(4.65)	0.010	(1.65)
GDP change	-0.001	(-0.79)	0.005	(0.96)
Unemployment	-0.017*	(-2.18)	-0.060***	(-4.04)
Unemployment change	0.024***	(3.62)	0.055***	(4.11)
Estimated End Balance	0.005*	(2.56)	-0.012**	(-2.75)
Population	-0.145	(-0.48)	-1.360+	(-1.90)
Population change	-0.006	(-0.41)	-0.005	(-0.14)
Population white	0.012*	(2.57)	0.002	(0.20)
Population over 65	0.021	(1.24)	0.009	(0.23)
Governor Democrat	0.017+	(1.69)	0.044	(1.10)
Governor Independent	0.062	(0.98)	0.379***	(9.77)
Governor female	-0.012	(-0.79)	-0.045	(-1.38)
Governor 1st term	-0.008	(-0.67)	-0.001	(-0.06)
Governor last term	-0.019	(-1.48)	-0.026	(-0.85)
Governor popularity	0.000	(0.52)	0.001	(0.27)
Democrats control both houses	0.030*	(2.14)	0.050	(1.34)
Legislature split control	-0.007	(-0.46)	0.110**	(2.88)
Governor-Legislature same party	-0.012	(-0.94)	0.055*	(2.37)
State liberalness	0.001	(0.33)	-0.004	(-0.96)
Agenda diversity (HHI)	-0.038	(-0.38)	-0.106	(-0.55)
Year=2008	-0.005	(-0.27)	0.055+	(1.88)
Year=2009	-0.015	(-0.50)	0.118+	(1.96)
Year=2010	0.043	(1.25)	0.283***	(4.00)
Year=2011	0.037	(0.96)	0.286***	(3.51)
Year=2012	0.022	(0.55)	0.195*	(2.21)
Year=2013	0.008	(0.18)	0.166+	(1.69)
Year=2014	0.010	(0.20)	0.169	(1.48)
Year=2015	-0.016	(-0.25)	0.087	(0.65)
Property crime per 100k	0.011	(0.19)		
Other tax share			0.008	(1.39)
Constant	13.953**	(2.98)	32.890**	(2.95)
N	450		450	
R^2 within	0.41		0.26	
R^2 between	0.48		0.46	
R^2 overall	0.46		0.44	
Note: t-statistics in parenthesis; * $p<0.05$, ** $p<0.01$, *** $p<0.001$.				

The main variables of interest, gubernatorial policy priorities, are significantly associated with two out of five spending categories. According to the findings, controlling for institutional, political and socio-economic factors, budgetary choices in the area of education, health care and natural resources are not affected by gubernatorial agenda. Highway spending and public safety, on the other hand, are significantly associated with gubernatorial policy priorities. The importance of transportation and public safety is positively related to state spending in these functions.

More specifically, with each additional percent of the State of the State address devoted to transportation, the total state highway spending tends to increase by 0.5 percent. For an average highway budget of $1.7 billion, that translates into an increase of $8.5 million. An average public safety budget of $830 million tends to increase by 0.2 percent, or by $1.7 million with each additional percent of the speech devoted to the topic of public safety.[17] Given that the average speech is about 4,400 words, 1 percent of the speech is about 50 words, which on average takes about 20 seconds to pronounce. Hence, 10 percent of an average speech take about 3 minutes. This implies that if a governor spends about three minutes of the address on the topic, an average increase in total highway spending may be about $85 million and in public safety about $17 million. The findings provide support only to hypotheses b and c.

Discussion

The empirical findings indicate that governors are more successful with their policy agenda in the areas of transportation and public safety. There are significant increases in state spending on these functions when these policy topics become more important to the chief executives. The

---

[17] The average values used in this section were obtained by calculating the exponential of the mean of log-transformed variables. They do not correspond to the average values of not log-transformed variables presented in the descriptive statistics table.

results support previous studies cited earlier that find significant influence of gubernatorial policy goals on state budgets. This study, however, provides a more granular analysis showing that executive influence varies by policy area. This would not have been captured had the total budget been used as the dependent variable.

It is worth considering why gubernatorial influence is discovered in some areas but not in others. Several explanations are possible. One potential reason is the difference between policy areas due to the difficulty of striking mutually acceptable deals with the legislators. Based on how focused their benefits are, the totality of public expenditures can be divided into redistributive and developmental policies (Baumgartner & Jones, 2010; P. E. Peterson, 1995). The key distinction is that redistributive polices have statewide benefits, and developmental expenditures are more focused on specific groups and/or geographic areas. Following this typology, infrastructure and public safety expenditures can be classified as developmental policies with focused benefits (Barrilleaux & Berkman, 2003). This has direct implications for our findings. Legislators may be more likely to go along with those gubernatorial proposals that allow them to attach their own pork projects. Bringing pork back home is a common way to win over a constituency. There is less personal incentive for lawmakers to vote for initiatives with dispersed, statewide effects. Preferential treatment of some districts for education funding, for example, would be harder to achieve as courts commonly rule such cases in violation of the equal protection clause. Similarly, once eligibility requirements are established for state-funded health care benefits, they become available to all qualify regardless of their legislative district.

Additionally, the debates over heath care and education policies are more ideologically contentious. The question often is not only how much to spend, but what is the most appropriate solution. Government funded health insurance and public schools are commonly contrasted with

private insurance and charter schools. Finding a common ground may be much more challenging in these areas. For these reasons, the ultimate outcome may be more cooperation between the two branches in the infrastructure and public safety area, where everyone wants better roads and equipment, and opportunities for collaboration are fewer in other policy subsystems.

As for natural resources expenditures, this dependent variable comprises spending on conservation, promotion, and the development of natural resources. Hence, it is impossible to distinguish conservation from development, which would be very helpful. Combined as they are, these expenditures may or may not provide focused benefits. This may explain why on average there is no significant association with gubernatorial priorities.

Yet another potential explanation for no significant relationship in some areas could be related to the measurement of the main variable of interest, gubernatorial policy priority. Whereas LDA provides a solid estimation of how important a topic is in each year, LDA does not reflect the tone and sentiment of the discussion. Some of the governors may highlight the challenges in the area and call for action, whereas others may simply boast about their achievements without asking for changes. This differentiation would not be captured by LDA. As a result, the estimation results may show no effect on average as different sentiments might cancel each other out. This shortcoming can be addressed in the future by supplementing the LDA topic model with some type of sentiment analysis of the text (Liu, 2015).

Conclusion

The main conclusion of this study is that governors indeed are powerful players in state policymaking as depicted in the previous literature. This study reveals, however, that their influence is not equal in all policy subsystems. The empirical findings suggest that state budgetary choices are significantly associated with the gubernatorial policy agenda only in the areas of

infrastructure and public safety. According to U.S. Census data, these functions take up about 12 percent of state budgets (infrastructure 8 and public safety 4 percent). Interestingly, scholars note that state discretionary spending that is subject to annual (or biennial) reconsideration also ranges between 5 and 15 percent (Rosenthal, 2012, p. 97). Infrastructure and public safety, of course, are not the only functions with discretionary budget authority, but these data point toward shrinking share of state funds that are subject to bargaining and negotiation described in the literature.

As more of state funds are absorbed by mandatory spending, there will be less opportunities for individual policymakers, including state chief executives, to pursue their policy agenda through the budgeting process. To some extent, this conclusion is what this study already finds. Hence, another conclusion is that empirical findings from the previous literature on gubernatorial powers may not hold in some areas of modern public budgeting. Increased partisanship, gridlock, brinkmanship, hostage taking and government shutdowns are much more common today (Fletcher & Weiner, 2011; Rubin, 2016; The Morning Call, 2017). In such an environment, it may be even harder for state chief executives to realize their roles as policy leaders than previously depicted.

A more detailed picture would require data by agency, department or individual programs, for example. This study makes a step forward by using the data by function, but future studies should test the theory with more granular data. Whereas this research finds no relationship with education, health care and natural resources, there may be changes within the function with the totals not affected. It would also be beneficial to see where the funds governors secure are spent. Another opportunity for future research is to test directly the extent to which the share of the State of the State Address is related to gubernatorial policy choices. The argument adopted here from the previous literature suggests that more important issues will likely receive more attention in the speech. This argument may or may not hold uniformly under different circumstances. New

empirical insights on this question would be beneficial to the researches using the State of the State

Address as a gauge of gubernatorial agenda.

CHAPTER 3

GUBERNATORIAL POLICY PRIORITIES AND CITY FISCAL OUTCOMES:

STATE-LOCAL POLICY MISMATCH

Introduction

This study aims to explore whether and under what conditions cities align the allocation of

fiscal resources with gubernatorial policy priorities. A great deal of research has been dedicated to

understanding the power of state chief executives in state policymaking. Governors are commonly

perceived as state policy leaders and often receive more attention than any other single elected

official in the state. The public routinely hold a state chief executives responsible for government

performance. Although they lack direct legislative authority, extant research suggests

gubernatorial offices have a significant effect on the distribution of state resources reflected in the

state budgets approved by legislatures as well as on overall state policies (Kousser & Phillips,

2012; Kousser & Phillips, 2009). Because of the "devolution revolution," governors have arguably

become even more politically influential over the last several decades (Beyle, 2004; DiLeo, 2001;

Van Assendelft, 1997; Weber & Brace, 1999). Thus, although their powers may vary from state to

state (Abney & Lauth, 1997, 2002; Lauth, 2016), the literature portrays state chief executives as

powerful actors within the state policymaking process.

The power of American governors at the city level has not received as much attention. The

governors' toolkit contains an array of formal and informal mechanisms that could be used to

encourage desirable city behavior. A strong stance in state policymaking may allow them to work

with lawmakers and push through a fully-fledged reform. Such state laws do have a significant

impact on local fiscal behavior (Wolman, McManmon, Bell, & Brunori, 2010; Wood, 2011). In the case of an uncooperative legislature, governors can act unilaterally through executive orders or line-item veto to exert pressure on state and local spending outcomes (Abney & Lauth, 1985; Ferguson & Bowling, 2008; Gosling, 1986; Holtz-Eakin, 1988; Reese, 1997). In addition to these time- and resource-consuming mechanisms, they can also use less formal policy instruments often referred to as a nudge, or persuasion (Bemelmans-Videc, Rist, & Vedung, 2011; Peters, 2015). Being in the center of attention, state chief executives can use their prominent positions to encourage changes by communicating their policy goals in public speeches. Holding all else equal, these softer and unbinding stimuli may induce a preferable response from local officials.

Scholars have not been particularly successful in examining the degree to which such informal signals may be associated with city fiscal behavior. Little is known on whether city governments align their spending decisions with governors' priorities, which have been found influential in state policymaking. Meanwhile, cities interact directly with citizenry and deliver a sizable proportion of essential public goods and services. Given the governors' importance at the state level, analyzing their success at the local level becomes increasingly important for understanding contemporary intergovernmental relations.

Consequently, this study seeks to fill the gap in the extant literature. It tests empirically the theoretical proposition, deduced in the following section, that gubernatorial policy priorities can serve as inducements, or a "nudge," and may have a significant influence on city fiscal outcomes. The assumption derived from the literature (Thaler & Sunstein, 2009) is that given their status, governors may be able to induce desired city policy behavior at minimal cost. The main incentive for city officials to take heed of these informal stimuli comes from the realization that they may need something from the state chief executive in the future, so they need to make a strategic

decision with the view of potential consequences. The informal signals can also be backed by formal actions and leave localities even less flexibility. Using probabilistic topic modeling algorithms, which are statistical methods facilitating the annotation of large collections of text documents with thematic information, this research documents gubernatorial policy priorities in a systematic manner over eight years. Next, the connection between governors' priorities and city fiscal outcomes is tested in a longitudinal study of the 149 largest cities in the U.S.

The empirical findings provide several valuable insights. First, city officials do align their expenditure choices with gubernatorial policy preferences. Second, city officials' reaction to such priorities varies by policy area, suggesting that local policymaking in different areas is driven by different factors. These findings provide some support to the notion that "policy drives politics" in state-local relations. Third, supporting theoretical expectations, the findings also indicate that, in some areas, policy alignment between the two levels of government is more likely when a state chief executive and city leadership identify with the same political party. These findings reveal new ways of interaction between state and city policymaking processes and outcomes. It shows that state chief executives matter for city policymaking and provides new contributions to our understanding of intergovernmental relations.

This paper is organized as follows. The next section lays out the theoretical foundation of the research and presents the formal hypothesis for empirical testing. It is followed by a review of the relevant literature. The data, research model and estimation methodology are explained in the fourth section. Empirical findings on the relationship between gubernatorial policy priorities and city fiscal outcomes are presented and interpreted in the fifth section. The final section discusses the empirical findings and relevant implications.

Theoretical Framework: Governors and Local Policies

This section advances a theoretical framework why city governments may be expected to react to gubernatorial policy preferences. Local government policy agenda and outputs, as of any other public institution, are a function of a diverse set of factors, none of which can perfectly predict final outcomes on their own. With certain reservations and generalizations, national factors can be divided into those of state and local origin. Local socio-demographic conditions, as well as political climate and institutions, for instance, are known to influence municipal government policy choices (e.g. Carr & Karuppusamy, 2010; Feiock, Jeong, & Kim, 2003; Frederickson & Johnson, 2001; Karuppusamy & Carr, 2012; Svara, 2005).

At the state level, the important sources of influence include the leaders of the state executive branch with their formal and informal instruments, legislatures, as well as the existing state institutional framework pertaining to local governments. Because municipal governments are legal creatures of the states that owe "their origin to, and derive their powers and rights wholly from, the legislature," as Justice Dillon famously opined,[18] one may be inclined to treat them as mere subordinates following state policy goals. There are, nonetheless, notable policy differences between the two levels of government. Contrary viewpoints on enforcement of immigration laws is one of the recent examples. Many localities are reluctant to follow federal immigration rules despite the emphasis that state officials put on local compliance. Many city mayors in 2017 vowed to follow the Paris Agreement and step up their efforts to fight climate change despite the fact that

---

[18] Clinton v Cedar Rapids and the Missouri River Railroad, (24 Iowa 455; 1868).

such policy contradicts ideas supported by higher levels of government , including some states (Bromley-Trujillo, 2017; Sampathkumar, 2017).

From a governor's perspective, these mismatches can be addressed through some formal procedure, such as new legislation, statutory changes to mandate certain behavior or grants to provide pecuniary incentives. Governor Greg Abbot of Texas, for instance, dissatisfied with the way local officials were enforcing federal immigration laws, signed a bill into law that threatens local officials with jail and removal from office if they do not cooperate with federal immigration agents. The action was taken despite the fact that every major police chief in Texas, including some of the largest cities in the U.S., opposed the measure (McCallister, 2017). Texas is not the first state to address such state-local policy mismatches through a legislative bill (Archibold, 2010). Governors specifically can act unilaterally by using the line-item veto or through executive orders to encourage certain behavior (Abney & Lauth, 1985; Ferguson & Bowling, 2008; Reese, 1997), which is more likely to happen when it becomes costly for them to bargain with an intractable legislature (Cockerham & Crew, 2017).

State chief executives, nonetheless, possess less formal (and in many respects less costly) ways of encouraging policy changes at the city level. Their status enables them to define and communicate the state policy agenda, which, of course, is not a clear official listing of issues that concerned city officials can turn to for guidance. It exists only in a collective judgment or as fragments of various statements and written evidence (Peters, 2015). Governors become especially prominent in defining that list and bringing it to the public and local officials' attention. State chief executives are more suitable for that role as they occupy a unique position in state politics. They are the most prominent individual elected state officials who can claim to serve all people in a given state. Speaking with one unified voice, they have more access to the media and receive more

public attention than any other public official in the state. Governors have the privilege to deliver an annual State of the State Address to both chambers of state legislatures,[19] which usually broadcast live on local media and serve as a primary opportunity to summarize achievements and lay out their policy agenda for the upcoming year. This unique position allows state chief executives to explicitly articulate and highlight current policy goals to and for the entire state. Other state officials may have different preferences, but they do not receive as much attention and enjoy as many opportunities to communicate them to the public as do state chief executives. In other words, governors have the biggest megaphone in the state.

In the light of these considerations, it is reasonable to expect that state policy priorities as emphasized by governors will be among the most noticeable and influential cues for city policymakers. These policy signals, even when not supported by the force of law, can be a subtle instrument used to encourage certain behavior among individuals and local units. Also known as policy suasion or a "nudge" (Peters, 2015; Thaler & Sunstein, 2009), this tool is often used to promote social changes more generally through the transfer of knowledge, reasoned argument and persuasion (Bemelmans-Videc et al., 2011). Political leaders using this approach not only transfer information, which can be phrased in an affirmative or negative way, but also discuss the nature of a policy-related problem and measures that can and should be taken to prevent the problematic situation as well as the reasons why these measures ought to be taken (Bemelmans-Videc et al.,

---

[19] In a handful of states, the address is delivered biennially and in some states it bears a different name, such as State of the Commonwealth Address in Kentucky, Virginia and others. Also, Nebraska has a unicameral state legislature.

2011). Thus, governors can make use of their bully pulpit[20] as state chief executives to trigger policy changes elsewhere in society, including many local governments.

In a state-local context, local units often provide essential public services. Notable problems at the local level may eventually require governors' attention, who, as state leaders, are expected by the citizenry to effectively manage the entire state. Policy suasion is not as costly and time consuming as other policy tools, and thus could be a good fit for initial attempts to alleviate any problems through encouragement of local policy changes. The effectiveness of such policy instruments that rely on the power of persuasion may depend on the presence of an underlying implicit threat of coercive action (Meyer, 2013). It has been stated that "persuasion will normally only be more effective than punishment when the persuasion is backed up by punishment" (Braithwaite, 2002, p. 19). Being influential players in state policymaking, governors can always try to buttress their informal "nudges" with formal legal action in cases of insufficient cooperation. Local funding may be manipulated with line-item vetoes or unilateral executive orders (Ferguson & Bowling, 2008). Governors in some states, like Kentucky, for example, also appoint finance officers to oversee local financial management who often have the power to require revenue or expenditure changes that do not conform to local preferences. As the precedents show, uncooperative local officials may end up in jail for contempt (Coe, 2008, p. 763). The case exemplifies how less formal encouragement can be reinforced with formal authority.

The toolbox of formal and informal sticks and carrots available to the governors varies from state to state, but local officials in any state are likely aware of their existence and will pay

---

[20] Bully pulpit is an important public position from which a person can let other people know his or her opinions on subjects.

attention to gubernatorial policy goals. Even when the formal tools are not explicitly invoked, local officials operate in the shadow of the fact that state chief executives have that option. It becomes a repetitive bargaining game, where local officials make strategic decisions about gubernatorial policy goals with the view of potential consequences in the future. Neglect on their part may turn out to be costly down the road. Hence, although informal, gubernatorial policy signals will be given careful consideration. Based on this argument, it is expected that, holding all else equal, explicit statements of gubernatorial policy priorities will be significantly associated with city fiscal outcomes.

Public policies may be a mere statement of intent, action or even inaction (Birkland, 2011). However, almost any policy intended to make a tangible change and alleviate a substantive social problem requires resources. Public budgets serve several goals, among which are setting policy priorities and distributing ever scarce resources by policy area (Mikesell, 2016; Mitchell & Thurmaier, 2012; Schick, 1966). They are key measures of policy output, since budget changes represent choices about what government will and will not do (Rubin, 2016). They result from complex policy processes involving, among other things, informational signals from a changing environment (Jones et al., 2009, p. 857). If a policy does not get its share of the pie, it is less likely to have any real effect on a social issue. Thus, if governors' policy priorities are an influential factor in city policymaking, as explained in the discussion above, they should be correlated with the level of spending and the distribution of city resources by policy area. The more important an issue is to the governor, the more likely city policymakers can be expected to react and align the distribution of resources with the governor's policy goals. Therefore, the preceding discussion leads to the following formal hypothesis for empirical testing:

Hypothesis 1: Gubernatorial policy priorities and city level spending are significantly associated.

Public budgets comprise short-term current operations expenditures and long-term capital investments. Each of the two broad categories is likely driven by different considerations, and their relationship with gubernatorial policy goals needs to be tested separately. There is a possibility that state chief executive policy signals will be correlated only with short-term initiatives and have a weaker or no relationship with long-term capital investments and infrastructure projects. Capital projects require long-term planning and multiyear commitment. Once initiated, they cannot be easily changed or terminated. Accordingly, the following two hypotheses are put forward:

Hypothesis 1a: Gubernatorial policy priorities and local current operating expenditures are significantly associated.

Hypothesis 1b: Gubernatorial policy priorities and local capital expenditures are significantly associated.

Testing these hypotheses separately should shed some light on the relationships between gubernatorial policy agendas and aggregate city current and capital expenditures. Local reaction to the policy signals, however, may be manifested in a redistribution of resources among different functions with the total level of expenditures unaffected. This possibility necessitates additional testing by individual policy area.

Functional responsibilities of state and particularly city governments also vary considerably across the nation. Local officials may be less concerned with gubernatorial policy signals in the areas where city involvement is negligible. Conversely, gubernatorial goals may matter more in those policy areas that take a sizable share of local budgets or where local units depend more on state aid. Risks are high in these cases as costs of retaliation for policy mismatches

will be more significant. Looking at the aggregate level of city spending may not reveal these intricacies.

Most state and local resources are devoted to such functions as education (K12 and higher), social services (public welfare and healthcare), transportation and public safety (NASBO, 2017; U.S. Census Bureau, 2017). These categories combined constitute on average over 60 percent of all state expenditures, almost 70 percent of local expenditures (NASBO, 2017; U.S. Census Bureau, 2017) and represent government activities in the main public policy areas. In order to test the proposition that local response may vary by policy area, additional hypotheses are put forward:

Hypothesis 1c: Gubernatorial policy priorities are statistically significantly associated with city spending on **education**.

Hypothesis 1d: Gubernatorial policy priorities are statistically significantly associated with city spending on **social services**.

Hypothesis 1e: Gubernatorial policy priorities are statistically significantly associated with city spending on **transportation**.

Hypothesis 1f: Gubernatorial policy priorities are statistically significantly associated with city spending on **public safety**.

Central to this relationship is the political partisanship of the main actors, which may mediate the strength of the association. The correlation between gubernatorial policy agenda and local fiscal outcomes may be more likely if a city mayor and governor are of the same party. On the other hand, a Republican mayor may be more reluctant to follow policy signals of a Democratic governor, holding the level of funding and other relevant factors equal. These supposition leads to another set of hypotheses:

Hypothesis 2: The association between gubernatorial policy priorities and city spending is stronger when city mayor and governor are of the same party.

Hypothesis 2a: The association between gubernatorial policy priorities and city education spending is stronger when city mayor and governor are of the same party.

Hypothesis 2b: The association between gubernatorial policy priorities and city social services is stronger when city mayor and governor are of the same party.

Hypothesis 2c: The association between gubernatorial policy priorities and city transportation spending is stronger when city mayor and governor are of the same party.

Hypothesis 2d: The association between gubernatorial policy priorities and city public safety spending is stronger when city mayor and governor are of the same party.

Overall, there may be an important policy process pattern that has been overlooked in the previous literature. Testing of these interactions will provide new insights into intergovernmental relations and improve our understanding of local policymaking and state executive powers to influence the behavior of city governments. The following section reviews available literature on the topic and highlights existing gaps.

Literature review

Studies examining gubernatorial powers have predominantly been done in the context of state policymaking. Generally, extant research portrays state chief executives as influential actors able to pursue their policy goals under different circumstances. Legislature characteristics, as well governors' personal attributes, political party, and economic environment may affect a governor's legislative success (Bernick & Wiggins, 1991; Ferguson, 2003; Ferguson & Barth, 2002;

Heidbreder & Scheurer, 2013; Jewell & Whicker, 1994; M. Peterson, 1990). Additional studies reveal that legislative professionalism, most importantly legislative session length (Kousser & Phillips, 2009), and ideological alignment (Kousser & Phillips, 2012) are important predictors of gubernatorial influence on state policies. More detailed analysis provides empirical evidence indicating that on average legislators pass 41 percent of items on the executive agenda and provide a compromise on an additional 18 percent (Kousser & Phillips, 2012). Governors are especially successful playing the budget game, where each dollar of overall revenue or spending changes proposed by the governor translates into 70 cents in the final budget adopted by the legislature. Legislative success of their policy bills, on the other hand, is more challenging and depends heavily on governors' political capital and the ideological leaning of the legislature (Kousser & Phillips, 2012).

Although there have been multiple studies of gubernatorial powers, the literature has been mostly silent on explicitly testing their role in local policymaking. Researchers have most frequently been interested in testing the effect of state laws on local fiscal behavior. They find that state tax and expenditure limitations (TELs) are effective in reducing local revenues and spending (Preston & Ichniowski, 1991; Shadbegian, 1998), although their effectiveness depends on the nature of limitations (T. Brown, 2000; Joyce & Mullins, 1991; Mullins & Joyce, 1996; Preston & Ichniowski, 1991; Shadbegian, 1999). Other state institutions defining service provision, personal management and other aspects of local discretion are also known to influence local policies (Wolman et al., 2010; Wood, 2011).

Thus, as it stands, the literature has demonstrated that governors are influential players in state policymaking, and that state institutions matter for local fiscal behavior. Little is known, however, on whether there is a relationship between softer signals in the form of gubernatorial

policy priorities and local policymaking. The field would arguably benefit from a closer examination of the potential association as it could help to uncover new links between state and local policies and contribute broadly to our understanding of contemporary intergovernmental relations.

Research Design

Data

The main hypotheses of this study are tested using a panel dataset of 149 cities in the U.S. from 2007 to 2014. City financial data were obtained from the Lincoln Institute of Land Policy and are based on original data fro the Government Division of the U.S. Census Bureau.[21] The sample includes the 100 largest cities in the country, plus at least two cities from each state, except for Hawaii and New Jersey. [22,23] The dataset is balanced as there are no missing values for any city-years (n=1,192). City socio-demographic data were obtained from the U.S. Census and city Comprehensive Annual Financial Reports (CAFR), governor party affiliation from National Governors Association and state legislature data from National Conference of State Legislatures. Municipal form of government data and mayor party affiliation were extracted from Municipal Year Books, individual websites and CAFRs. All dollar values in the dataset were adjusted for inflation to 2014 real values. About 67.4 million, or 21 percent of the U.S. population lived in

---

[21] The data are available from the Lincoln Institute of Land Policy database at
http://datatoolkits.lincolninst.edu/subcenters/fiscally-standardized-cities/.
[22] The original dataset includes 150 cities, but Washington, DC was dropped from the analysis. The list does not always include state capitals, as these are not always the largest cities in the state (see Appendix C for a full list).
[23] It was not possible to include cities from Hawaii or New Jersey in the sample because the largest cities in these states have state-administered school districts, which makes it impossible to disentangle the revenues and expenditures that should be allocated to the cities from the rest of the state's budget. Hawaii has a statewide school district, while New Jersey's largest school districts are not included in the Census of Government Finances dataset because the state government directly funds a large share of these districts' budgets as a result of the *Abbott* court decisions.

these cities in 2014. The descriptive statistics for all variables used in the study are provided in Table 3.1.

Dependent variables

The theoretical framework predicts a potential association between gubernatorial policy priorities and local spending outcomes. Following the existing literature (Jones et al., 2009; Mikesell, 2016; Mitchell & Thurmaier, 2012; Rubin, 2016; Schick, 1966), it is assumed that local spending reflects current policy priorities of local governments. If gubernatorial policy signals are important for local policymaking, they should be significantly correlated with city spending patterns. As has been noted in the theory section, this study divides the total local budget into current and capital expenditures because these are two distinct policy areas that may be driven by different considerations. Thus, the first two main outcome variables of interest in this study are the total level of spending on current operations and the total level of capital outlays for each city-year.[24] Taken together, current operations and capital outlays, on average, account for 93 percent of total direct expenditures,[25] and range from 76 to 99.7 percent.

Theoretically, there is also a possibility that individual policy objectives may be correlated with total city expenditures on individual functions, such as education, social services, transportation and public safety. Hence, additional four outcome variables include total city spending on education, social services, transportation and public safety. All dependent variables were logged to alleviate positive skew. These four spending categories on average account for

---

[24] Per capita spending by policy area was used as a robustness check and although the models perform substantially worse, the findings are not affected in any model.

[25] The remaining 7 percent include assistance and subsidies, interest on debt and insurance benefits. Direct expenditures include all city expenditures except for utility expenditure, liquor store expenditure and employee retirement trust expenditure.

about 60 percent of total city spending in the sample. Education generally takes the largest share of city budgets, from 7 to 62 percent, with the average of 31 percent. Social services' share ranges from 0 to 59 percent of the total budget, with average spending of 8 percent. Transportation expenditures take from 0.5 to 23 percent (mean 6.5 percent), and public safety from 5 to 24 percent (mean 12 percent). The descriptive statistics indicate the considerable variation of city responsibilities. Whereas education is an important activity for most of them, city involvement in other areas may be minimal. Consequently, the importance of gubernatorial policy signals may also vary by city.

Independent variables

Gubernatorial policy priorities represent the main predictor of interest in this study. Governors have several opportunities to share their policy agenda with a wide audience. Policy priorities for this study were extracted from the State of the State Addresses, which are high profile events taking place in all states. Given the time limits on the address, governors usually choose to focus on the main policy initiatives for the upcoming year (Herzik, 1991; Light, 1999). These speeches generally contain a wide range of policy statements and allow a much more nuanced measurement of gubernatorial policy priories in each year than, for example, party labels. The address is intended to influence public opinion as well as legislative process (Rosenthal, 1990; Van Assendelft, 1997). For these reasons it is a superior gauge of executive policy goals (Ferguson, 2003) and a good predictor of subsequent gubernatorial actions (Jackson & Kingdon, 1992; Segal & Cover, 1989). The content of gubernatorial speeches is commonly used to measure governors' policy agendas (Crew, 1992; DiLeo, 2001; Ferguson, 2003; Herzik, 1991; Kousser & Phillips, 2012; Morehouse, 1998).

Other gubernatorial messages are generally shorter and much narrower in scope. The annual or biennial executive budget message, for instance, may also shed some light on executive policy preferences, but its length varies from barely a page, as in California and North Carolina, to multiple pages. Apparently, this statement has a different value to state executives, in some cases only nominal, and serves different goals. Most of the time the budget message focuses on specific issues pertaining to state fiscal performance rather than general policy goals of the administration. It usually provides a summary of the recommended budget, describes revenue and expenditure trends, lays out the main assumptions built in the proposed budget and highlights major areas of concern for state fiscal health. Unlike the widely-publicized State of the State addresses, budget messages rarely receive as much public attention and are not intended as major public policy statements. In that respect, the State of the State address is a superior source of executive policy preferences. Thus, as it has been aptly summed up, "lacking detailed interview or survey data from governors, their major public addresses provide the best insight available into their preferences, values, and ideology" (Coffey, 2005, p. 90).

Earlier studies commonly analyzed the content of text documents by hand with human coders quantifying text into data. Documents were coded based on the number of sentences devoted to a topic (Heidbreder, 2012; Heidbreder & Scheurer, 2013) or the topic structure was simply assumed based on the reading of the document (Ferguson, 2003; Ferguson & Barth, 2002). This is a time-consuming approach that is also prone to errors. This study utilizes LDA topic model results from the first chapter of the dissertation. Assuming that governors devote a greater share of their speeches to the policies that are more important, LDA modelling allows one to systematically measure the importance of a given policy area to the governor in each year. Out of 15 topics identified by the LDA in the first chapter, four topics are selected as they are directly related to

each of the spending categories of interest in this study. These topics include education (topic 13), healthcare (topic 3, related to social services), transportation (topic 9) and public safety (topic 14). Seven other topics are included to control for other potentially important topics on gubernatorial agenda. This list consists of budget (topic 12), economy (topic11), economic development (topic 1), jobs (topic 10), tax policy (topic 6), energy and natural resources (topic 15) and local government (topic 2). All these issues may send an important signal to local officials and need controlling to lower the chances of omitted variable bias.

In order to test the second set of hypotheses focusing on political alignment between state and local officials, the first four substantive policy topics are interacted with a dummy indicator. This dummy is coded 1 if the city mayor and the governor identify with the same party. This variable also equals 1 when the mayor is independent (or when the city uses a council-manager form of government) and the governor is independent. The expectation is that the coefficient on the interaction term will be statistically significant if political alignment matters for city-state policy alignment.

Controls

In addition to state chief executive policy goals, the models include a set of controls found to be important for local fiscal choices (Carr & Karuppusamy, 2010; Feiock et al., 2003; Frederickson & Johnson, 2001; Karuppusamy & Carr, 2012; Svara, 2005).

City Socio-Demographic conditions and Intergovernmental Revenue. The models account for city population (logged), which varies in the sample from 16,000 in Rutland, Vermont to over eight million residents in New York, NY. Percent population change from the last year and population density per square mile of city land area are also included. Other variables control for the percent of population over 65 years old, percent unemployed and median household income.

75

The latter three controls help to account for public service demands, such as social services for the elderly, unemployment benefits and community wealth. The percent of white population accounts for city racial composition, which remains one of the most enduring demographic forces in American public policies (Meier, 2016).

The amount of total intergovernmental aid was highly correlated (0.91) with city population, hence the percent of total revenue from IGR is used instead. This variable is intended to account for city dependence on external funding and discretion in management of its resources.

City Political Forces. It is also important to account for the political ideology of local leaders in this study.[26] Two dummy indicators control for the city mayor being a Democrat or an Independent, with Republican mayors as the omitted comparison category. Cities with a council-manager form of government were coded as independent. Although some council members may have certain ideological preferences, this form of government is generally considered less politicized with appointed, rather than elected, chief executive officers who serve as apolitical administrators.

State level controls. The macroeconomic environment within a state can affect the local tax base and lead to changes in city revenues and expenditures. Thus, the models account for total Gross Domestic Product (GDP) per capita and percent change in state GDP per capita relative to the previous year.

State-imposed tax and expenditure limitations (TELs) may influence local fiscal behavior in various ways depending on their nature (e.g. T. Brown, 2000; Joyce & Mullins, 1991; Mullins

---

[26] The initial intention was to control for the form of city government, but there is little temporal variation in this measure, which makes it impossible to include in the fixed-effects within estimation models used in this study.

& Joyce, 1996; Mullins & Wallin, 2004; Preston & Ichniowski, 1991; Shadbegian, 1998; Shadbegian, 1999). Several dummy indicators account for the following types of limitations: property tax levy limit, property tax rate limit, revenue and expenditure limit, requirement of full disclosure, property assessment limit, or no limit. Cities with a property tax rate limit (about 77 percent of all observations) is the omitted comparison category.

Governors' partisanship is similarly important for policy alignment. Two dummy variables were included to control for a Democratic and for an Independent governor, with Republican governors as the baseline.

Partisan composition of state legislatures is controlled for with two dummies, Democratic control of both chambers (Yes = 1, else = 0) and split legislature control (Yes = 1, else = 0). Legislatures under Republican control are the omitted comparison category. Following previous research, the nonpartisan, single chamber legislature in Nebraska is treated as split legislative (Buschman & Sjoquist, 2017). This set of controls help to account for any local budgetary changes that may be due to state legislature policy decisions. More Democratic control, for instance, may be associated with more liberal welfare policies for the entire state (Avery & Peffley, 2005; R. D. Brown, 1995; Fellowes & Rowe, 2004).

Year Fixed effects. Seven year fixed effects were included in the model with 2007 being the baseline to account for any time effects that may affect the outcome variables. These controls are especially important to account for the potential influence of the Great Recession on local fiscal policies.

Table 3.1. Descriptive Statistics.

Variable	Mean	SD	Min	Max
Gubernatorial Policy Priorities				
Budget	10.9	9.0	0.4	60.2
Economy	10.5	7.6	0.3	38.8
Healthcare	9.4	7.7	0.7	39.5
Economic Development	8.9	8.1	0.5	46.6
Education	8.6	6.9	0.3	35.6
Jobs	8.0	7.3	0.3	38.6
Tax policy	6.1	5.9	0.5	48.4
Energy/Natural Resources	5.6	6.2	0.2	41.4
Transportation	4.9	5.0	0.3	42.9
Local Government	3.5	5.2	0.2	47.5
Public Safety	2.8	3.6	0.2	42.7
Dependent Variables				
Current operations, total (ln)	20.9	1.1	18.0	25.0
Capital Outlay, total (ln)	19.1	1.3	14.5	23.7
Education expenditures, total (ln)	19.9	1.0	17.6	24.1
Social Services expenditures, total (ln)	18.1	2.0	6.9	23.5
Transportation expenditures, total (ln)	18.2	1.2	14.6	22.1
Public Safety expenditures, total (ln)	19.0	1.1	15.9	22.9
Control Variables				
Population total (ln)	433,662.9	790,147.2	16,116.0	8,436,047.0
Population change (%)	0.7	2.1	-53.4	16.8
Population density (per sq. mile)	3,970.4	3,371.7	164.3	27,874.9
Median Household Income (ln)	44,826.4	11,127.9	23,600.0	106,921.0
Population over 65 (%)	11.6	2.5	5.0	24.1
Unemployed (%)	8.3	3.5	2.0	28.1
White (%)	63.8	18.3	10.5	97.0
IGR, percent of total revenue	38.3	10.8	16.2	71.6
Governor Democrat	0.5	0.5	0.0	1.0
Governor Independent	0.0	0.1	0.0	1.0
Governor Republican	0.5	0.5	0.0	1.0
Mayor Democrat	0.6	0.5	0.0	1.0
Mayor Independent	0.1	0.3	0.0	1.0
Mayor Republican	0.3	0.4	0.0	1.0
Mayor-Council form	0.6	0.5	0.0	1.0
Mayor-Governor same party	0.4	0.5	0.0	1.0
Legislature Democratic	0.4	0.5	0.0	1.0
Legislature Split	0.2	0.4	0.0	1.0
Governor-Legislature same party	0.6	0.5	0.0	1.0
Mayor-Governor-Legislature same party	0.3	0.5	0.0	1.0
State GDP per capita ($1,000)	47.4	8.4	31.2	73.5
State GDP change (%)	0.1	2.7	-9.2	19.3
Assessment limit	0.5	0.5	0.0	1.0
Full disclosure	0.3	0.5	0.0	1.0
Levy Limit	0.6	0.5	0.0	1.0
Revenue/Expenditure Limit	0.2	0.4	0.0	1.0
No limit	0.0	0.2	0.0	1.0

Rate Limit	0.8	0.4	0.0	1.0

Note: Gubernatorial policy priorities represent the percent of the State the State Addresses devoted to each policy area. Total city population and median income are not logged in the table for demonstrative purposes. Year dummies omitted for brevity.

Estimation Methodology

Given the continuous nature of the dependent variables, linear models are estimated. The Hausman specification tests pointed to a potential unobserved heterogeneity bias indicating that the data do not support a random effects model. Therefore, a fixed-effects within estimator model is used, which seeks to explain temporal variance within cities and accounts for time-invariant unobserved heterogeneity, which may bias the random effects estimators (Halaby, 2004). Since this study seeks to analyze the variables that vary over time, fixed-effects within estimation is appropriate for addressing the main research question. Standard errors are clustered by state in all models to deal with potential problems stemming from heteroskedasticity or autocorrelation. There is no formal test of the level at which to cluster (Cameron & Miller, 2015, p. 17). "The consensus is to be conservative and avoid bias and use bigger and more aggregate clusters when possible, up to and including the point at which there is concern about having too few clusters" (Cameron & Miller, 2015, p. 17). But how many clusters is enough? It is generally recommended to have at least 42 clusters for the standard cluster adjustment to be reliable (Angrist & Pischke, 2008). Hence, clustering by state is preferred to clustering by city.[27,28] Empirical results are described in the following section.

---

[27] Alternative model specifications with standard errors clustered by city were tested with substantive results not affected.
[28] The Variance Inflation Factor (VIF) check for multicollinearity revealed that the dummy indicator for 2009 was the variable with the highest VIF (3.13). VIF in the models with the interaction terms was the highest for topic 15 (4.84), which suggests there are no concerns about multicollinearity (Gujarati, 2007).

Empirical Findings

The empirical results are summarized in Tables 3.2, 3.3 and 3.4. The relationship between gubernatorial policy agenda and the total level of city spending are discussed first (Table 3.2, Model 1 and 2). Overall, the current expenditure model performs well and explains over 90 percent of variation between cities and over 30 percent within cities. The empirical findings indicate that controlling for city socio-demographics, political factors and state macroeconomic conditions, several gubernatorial policy priorities are significantly associated with city current spending (Table 3.2, Model 1).

According to the results, the importance of job creation and investment to the governor is significantly and positively associated with city operating expenditures (p-value<0.01). With each additional percent of the speech focused on this topic, total city operating expenditures tend to increase by 0.1 percent. An identical relationship is found between governors' discussion of healthcare and city operating expenditures. The other policy topic, energy and natural resources, is marginally (p-value=0.09) and negatively associated with the city operating expenditures. The estimated coefficient indicates a 0.1 percent decrease in total current expenditures with 1 percent more attention to the energy topic. Given that the average operating budget in the sample is $1.142 billion, a 0.1 percent change in the total operation budget translates into $1.142 million difference, or around 1 million dollars.

As for the control variables, larger and wealthier communities, as measured by total population, median household income and state GDP per capita, tend to have higher levels of expenditures, most likely due to greater availability of resources. Faster economic growth, as measured by percent change in GDP per capita relative to the previous year, is associated with lower current spending possibly due to lower demand for government assistance. Interestingly,

cities located in the states led by independent governors tend to have higher levels of current spending.

Table 3.2. Gubernatorial Policy Agenda association with Current Spending and Capital Outlays.

	Model 1 (Current)		Model 2 (Capital)	
Gubernatorial Policy Priorities	*β*	*t*	*β*	*t*
Education	0.00	(0.07)	0.00	(0.22)
Infrastructure	0.00	(1.18)	-0.00	(-0.82)
Tax policy	0.00	(-0.67)	0.00	(0.23)
Economic development	0.00	(-0.22)	0.00	(0.69)
Job Creation	**0.001****	**(2.92)**	-0.00	(-0.59)
Local gov. reforms	0.00	(0.69)	0.00	(0.59)
Energy, nat. resources	**-0.002$^+$**	**(-1.78)**	0.00	(0.15)
Public safety	0.00	(0.49)	0.00	(0.92)
Budgetary challenges	0.00	(1.49)	0.00	(0.85)
Healthcare	**0.001****	**(2.69)**	0.004$^+$	(1.87)
Economy	0.00	(0.81)	**0.007***	**(2.58)**
City Level Controls				
Population total (ln)	0.607***	(3.69)	-0.715	(-0.93)
Population change (%)	0.00	(-0.89)	0.017**	(3.17)
Population density (per sq mile)	0.00	(1.16)	0.000*	(2.32)
Median Household Income (ln)	0.176***	(3.86)	0.661**	(2.88)
Population over 65 (%)	0.00	(-0.59)	0.00	(0.29)
Unemployed (%)	0.00	(-0.03)	0.01	(1.18)
White (%)	0.00	(-0.86)	0.00	(1.15)
IGR, percent of total revenue	0.00	(0.10)	0.00	(0.87)
Mayor Democrat	-0.00	(-0.96)	0.06	(1.37)
Mayor Independent	0.00	(0.31)	-0.07	(-1.04)
Mayor-Governor same party	0.01	(1.54)	0.00	(0.02)
State Level Controls				
Governor Democrat	0.00	(0.49)	-0.03	(-0.58)
Governor Independent	0.033**	(3.02)	0.166	(1.04)
Legislature Democratic	0.00	(0.01)	-0.07	(-0.94)
Legislature Split	0.022$^+$	(1.73)	0.03	(0.56)
State GDP per capita ($1,000s)	0.006***	(3.57)	0.01	(1.46)
State GDP change (%)	-0.002***	(-3.52)	-0.01	(-0.9)
Assessment limit	0.044$^+$	(1.72)	-0.272*	(-2.50)
Full disclosure	0.013	(0.34)	-0.012	(-0.05)
Levy Limit	0.00	(0.17)	0.340***	(3.73)
Revenue/Expenditure Limit	0.075***	(5.06)	-0.178*	(-2.68)
Year Dummies				
Year 2008	0.032***		-0.003	(-0.10)
Year 2009	0.026*		-0.042	(0.77)
Year 2010	0.017		-0.16	(-0.21)
Year 2011	0.00		-0.90	(-1.15)
Year 2012	-0.018		-0.11	(-1.41)
Year 2013	-0.012		-0.125	(-1.49)
Year 2014	-0.003		-0.113	(-1.22)
Constant	11.045***	(5.38)	18.580$^+$	(1.98)
N	1,192		1,192	
R^2 within	0.35		0.13	
R^2 between	0.93		0.23	

R^2 overall	0.93	0.21

Note: t-statistics in parenthesis; $^+$ p<0.10, * p<0.05, ** p<0.01, *** p<0.001.
N=1,192, n=149, T=8.

The next model tests the relationship between gubernatorial policy priorities and the total level of municipal capital outlays (Table 3.2, Model 2). Compared to the current expenditures model, the capital outlays model overall explains less variation in the outcome variable thus suggesting that this type of spending is a function of a somewhat different set of factors. Capital outlays seem to be driven to a greater extent by individual city characteristics rather than common factors included in the model. Overall the model explains 21 percent of variation, and within R^2 is 0.13.

Model 2 indicates that of all the policy topics, only economic challenges are significantly associated with city capital outlays. One percent of the speech devoted to this topic is associated with a 0.7 percent increase in city capital investment. The association is statistically significant when controlling for annual fluctuations, IGR, macroeconomic environment, city socio-demographics and political factors. The average capital budget in the sample is $197,300,000 so 0.1 percent of the budget is $197,300. Thus, 0.7 percent change in the capital budget translates into a difference of about $1.4 million. Taken together, Models 1 and 2 provide some empirical support for hypothesis 1a and 1b.

The next four models test the relationship between gubernatorial policy priorities and city spending on education, social services, transportation and public safety (Tables 3.3 and 3.4). In each of these models relevant gubernatorial policy priority is interacted with a dummy indicator coded one if governor and mayor are of the same party, and zero otherwise. For example, the dependent variable in Model 3 is the total level of city spending on education. The dummy indicator for mayor and governor belonging to the same party is interacted in this model with the

share of the speech devoted to the topic of education. This dummy is interacted with the healthcare topic[29] in Model 4, the infrastructure topic in Model 5 and the public safety topic in Model 6.

Table 3.3. City Spending by Policy Area and Mayor-Governor Interaction Effects.

Dependent Variable	Model 3 Education		Model 4 Social Services	
	β	t	β	t
Interaction term	0.002*	(2.22)	0.007	(1.73)[+]
Education	0.00	(-1.67)	0.00	(-1.19)
Infrastructures	0.00	(1.56)	0.01	(1.48)
Tax policy	0.00	(-0.47)	0.00	(0.07)
Economic development	0.00	(0.67)	0.00	(-0.42)
Job Creation	0.00	(0.87)	0.00	(1.34)
Local gov. reforms	0.00	(0.16)	0.01	(1.10)
Energy, nat. resources	0.00	(-1.59)	0.00	(-0.17)
Public safety	0.003**	(2.90)	0.00	(0.06)
Budgetary challenges	0.00	(1.70)	-0.005	(-1.80)[+]
Healthcare	0.002***	(4.99)	0.00	(-0.98)
Economy	0.00	(1.57)	0.00	(0.89)
City Level Controls				
Population total (ln)	1.097***	(4.91)	3.039**	(3.24)
Popul. change (%)	-0.005**	(-3.31)	-0.01	(-1.21)
Popul. density (per sq mile)	0.000**	(3.15)	0.00	(0.24)
Median Household Inc. (ln)	0.242***	(3.55)	-0.18	(-0.44)
Population over 65 (%)	-0.01	(-1.04)	-0.03	(-1.25)
Unemployed (%)	0.00	(-0.50)	0.01	(1.59)
White (%)	0.00	(-0.40)	0.00	(0.03)
IGR, % of total revenue	0.003**	(2.75)	-0.01	(-0.78)
Mayor Democrat	0.02	(1.08)	-0.14	(-1.17)
Mayor Independent	-0.02	(-0.76)	-0.02	(-0.28)
Mayor-Governor same party	-0.02	(-1.84)	0.04	(0.80)
State Level Controls				
Governor Democrat	-0.01	(-0.90)	-0.04	(-0.80)
Governor Independent	0.03	(1.56)	-0.25*	(-2.57)
State GDP ($1,000s)	0.013***	(4.34)	0.01	(1.16)
State GDP change (%)	-0.004**	(-2.99)	0.00	(0.47)
Legislature Democratic	0.02	(1.21)	-0.03	(-0.25)
Legislature Split	0.03	(1.82)	0.02	(0.19)
Assessment limit	-0.03	(-1.01)	0.22	(1.28)
Full disclosure	-0.153*	(-2.21)	-0.20	(-0.86)
Levy Limit	0.079**	(2.85)	-0.16	(-1.36)
Revenue/Expenditure Limit	0.00	(-0.24)	0.04	(0.45)
Governor-Legislature Same Party	0.00	(-0.28)	0.00	(-0.02)
Year Dummies				
Year 2008	0.01	(0.92)	-0.03	(-0.68)

---

[29] Of all topics, healthcare is the most suitable for this spending category. On average, healthcare expenditures also account for three quarters of the city social services expenditures.

Year 2009		0.00	(0.06)	-0.09	(-0.93)
Year 2010		0.00	(-0.06)	-0.15	(-1.79)
Year 2011		-0.041*	(-2.45)	-0.188*	(-2.14)
Year 2012		-0.068***	(-3.73)	-0.211*	(-2.46)
Year 2013		-0.078***	(-4.12)	-0.218*	(-2.47)
Year 2014		-0.066**	(-2.80)	-0.18	(-1.87)
Constant		2.69	(0.83)	-17.70	(-1.51)
N		1,192		1,174	
R^2 within		0.42		0.10	
R^2 between		0.92		0.65	
R^2 overall		0.91		0.62	

Note: [+] $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$. N=1,192, n=149, T=8 (social services model excludes 18 city-years with zero spending on this function). Mayor-governor dummy is interacted with education topic in model 3, and with healthcare topic in model 4.

This set of models indicates that city spending on education increases by 0.2 percent when governors devote more attention to education, but only when the mayor and governor identify with the same party (Table 3.3, Model 3). City education expenditures also tend to increase when governors highlight public safety (by 0.3 percent) and healthcare (by 0.2 percent). The relationship between city spending on social services, which include welfare and health care, and gubernatorial attention to healthcare topic is marginally significant (p-value=0.09) when mayor and governor are of the same party (Table 3.3, Model 4). With each additional percent devoted to this topic, city social services spending may increase by 0.7 percent. On the other hand, city spending on social services may decline by 0.5 percent (p-value 0.08), with more attention to budgetary challenges. No other policy priority is related to social spending.

The next model finds that city spending on transportation tend to increase by 0.7 percent with each additional percent of gubernatorial speech devoted to infrastructure (p-value=0.054) when the governor and mayor belong to the same party. On average, however, more gubernatorial attention to this topic is associated with a 0.7 percent decrease in city transportation spending (p-value 0.045). It appears that when mayor and governor are of the same party, the direct and the interaction effects cancel each other out and produce no substantial changes, whereas there are

significant declines when mayor and governor are not of the same party. Other policy topics are

not related to this type of spending (Table 3.4, Model 5).[30]

Table 3.4. City Spending by Policy Area and Mayor-Governor Interaction Effects.

Dependent Variable	Model 5 Transportation		Model 6 Public Safety	
	β	t	β	t
Interaction term	0.007	(1.98)[+]	0.00	(0.85)
Education	0.00	(1.68)	0.00	(-1.38)
Infrastructure	-0.008	(-2.06)	0.00	(0.59)
Tax policy	0.00	(0.90)	0.00	(-0.08)
Economic development	0.00	(-0.84)	0.00	(0.97)
Job Creation	0.00	(0.20)	0.00	(0.59)
Local gov. reforms	0.00	(0.98)	0.00	(0.31)
Energy, nat. resources	0.00	(1.06)	0.00	(0.21)
Public safety	0.00	(-0.29)	0.00	(-0.28)
Budgetary challenges	0.00	(1.60)	0.002*	(2.19)
Healthcare	0.00	(1.80)	0.00	(1.43)
Economy	0.00	(0.59)	0.00	(1.69)
City Level Controls				
Population total (ln)	-0.18	(-0.32)	0.527**	(2.88)
Popul. change (%)	0.010**	(2.70)	0.00	(-1.92)
Popul. density (per sq mile)	0.00	(0.93)	0.00	(0.29)
Median Household Inc. (ln)	0.467*	(2.45)	0.265***	(3.90)
Population over 65 (%)	-0.02	(-1.39)	-0.01	(-1.37)
Unemployed (%)	0.01	(1.16)	0.00	(0.48)
White (%)	0.00	(0.21)	0.00	(-0.88)
IGR, % of total revenue	-0.01	(-1.81)	0.00	(0.59)
Mayor Democrat	0.00	(-0.01)	-0.01	(-0.36)
Mayor Independent	-0.08	(-1.35)	-0.01	(-0.32)
Mayor-Governor same party	-0.05	(-1.60)	0.00	(0.03)
State Level Controls				
Governor Democrat	0.087*	(2.29)	-0.01	(-0.80)
Governor Independent	0.154**	(3.22)	0.04	(0.65)
State GDP ($1,000s)	0.01	(0.73)	0.00	(1.15)
State GDP change (%)	0.00	(-0.18)	-0.003*	(-2.48)
Legislature Democratic	0.00	(0.04)	-0.02	(-0.82)
Legislature Split	0.06	(1.41)	0.02	(0.95)
Assessment limit	-0.12	(-1.36)	0.04	(1.02)
Full disclosure	0.17	(0.70)	-0.01	(-0.20)
Levy Limit	-0.11	(-0.97)	0.117***	(3.62)
Revenue/Expenditure Limit	-0.09	(-1.26)	-0.05	(-2.00)
Governor-Legislature Same Party	0.05	(1.30)	-0.01	(-0.57)
Year Dummies				

[30] This model explains a relatively small amount of variation with within $R^2=0.09$ and overall $R^2=0.01$. This is mostly because the cities in the sample play relatively little role in that area spending on average about 6 percent and maximum 23 percent of their budgets on transportation. There seems to be little variation in this outcome variable.

Year 2008	0.03	(0.63)	0.023*	(2.30)
Year 2009	0.02	(0.29)	0.01	(0.68)
Year 2010	-0.02	(-0.31)	0.00	(-0.06)
Year 2011	-0.04	(-0.73)	-0.01	(-0.59)
Year 2012	-0.03	(-0.59)	-0.01	(-0.29)
Year 2013	0.01	(0.18)	0.01	(0.36)
Year 2014	0.03	(0.48)	0.03	(1.09)
Constant	15.203*	(2.39)	9.349***	(4.23)
N		1,192		1,192
R^2 within		0.09		0.18
R^2 between		0.01		0.90
R^2 overall		0.01		0.90

Note: [+] $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$. N=1,192, n=149, T=8. Mayor-governor dummy is interacted with transportation topic in model 5, and with public safety topic in model 6.

The public safety topic is not significantly associated with public safety spending, with or without the interaction term (Table 3.4, Model 6). This model also suggests that cities tend to significantly increase public safety spending by 0.2 percent when governors spend more time discussing economic challenges. If the theory that social and economic problems increase the crime rate is an accurate description of reality, then it makes sense for cities to scale up their law enforcement spending in anticipation of higher crime rates following economic downturns.

To summarize, the empirical findings provide support for hypothesis 1, suggesting a significant relationship between gubernatorial policy agenda and local fiscal policies. Hypotheses 1a and 1b are also supported as some of the policy priorities are significantly associated with current expenditures and capital outlays. There is no empirical support for hypothesis 1c, 1d, 1e, and 1f as gubernatorial policy goals are not associated with city spending by policy area without the interaction effect.

There is empirical support for hypothesis 2, since in some policy areas there is a statistically significant relationship between the variables of interest that is observed only when there is ideological alignment between the main actors. More specifically, hypotheses 2a and 2c are supported, hypothesis 2b is marginally supported and only 2d is not supported.

Discussion and Conclusion

The main research question addressed in this study is where and under what conditions city officials may align their spending decisions with gubernatorial policy priorities. The policy priorities were extracted from the State of the State addresses delivered from 2007 to 2014 using LDA topic models. The theoretical framework advanced in this study posits that gubernatorial speeches can be an informal tool used to persuade and induce policy changes at the state and local levels. It also suggests that the potency of this tool may be conditional upon political alignment between state and city officials. The empirical findings from the longitudinal study of 149 largest US cities support the main hypotheses and indicate that city spending patterns are associated with gubernatorial policy priorities.

Conventional wisdom suggests that state-local relations are shaped by laws and financial incentives in the form of categorical and block grants. This study, however, points to a significant association between gubernatorial policy statements and local spending choices even after controlling for intergovernmental aid, fiscal intuitions, socio-economic environment and other important factors. Thus, this research provides initial evidence that state chief executives may be able to influence city spending decisions through less formal mechanisms, such as public speeches, that have not been recognized by the literature. Although this tool does not provide full control over local policymaking, it most likely enables the governors to make a significant difference in city policies at a relatively low cost. According to the findings, the effectiveness of this tool varies by policy area and depends upon political alignment.

City current spending tends to fluctuate with governors' attention to job creation and investment (Topic 10). Manual analysis of the content of speeches revealed that within this topic governors discuss states' advancements in job creation rather than economic challenges (which

are discussed in Topic 12). Most of the time this topic includes positive news when governors highlight reduced unemployment and increased investment. In addition to the commonly available economic indicators, these statements may signal an optimistic forecast to the cities, which, in turn, act to expand their current operations with more confidence. This finding supports the notion that local officials pay heed to gubernatorial policy signals and tend to react in the expected manner.

Economic challenges (Topic 11), on the other hand, are not related to current spending, but are significantly and positively associated with capital outlays. Most likely, this is because of city efforts to follow a Keynesian logic during economic downturns highlighted by the leaders of their states. The period of the study covers the Great Recession, and the sample includes the largest cities from each state. So, it is reasonable to conclude that the cities invest in labor-intensive infrastructure projects to spur aggregate demand and counteract economic decline and growing unemployment. This effect would have been hard to discern had the current and capital expenditures not been considered separately.

Governors' attention to energy and natural resources may also be related to decreases in city operating expenditures. The states where this topic takes a sizable share of the speech rely heavily for revenue on natural resources. Consequently, more attention to this topic may signal challenges with state revenue and future uncertainty about state aid to lower levels of government. This, in turn, may induce city officials to cut their expenditures and seek efficiency improvements.

In the area of healthcare policy, for example, which has long been a salient issue in the Unites States, the findings suggest that when a mayor and governor are of the same party, cities tend to significantly increase their healthcare expenditures. Most of the time within this topic governors talk about expanding health insurance and improving health care access and quality for

state residents. Governor John Lynch of New Hampshire in 2007 devoted 40 percent of his speech to healthcare to convey the main message that the state "must act to expand our children's health insurance program this year" and "make health care more affordable and accessible to all of our citizens." Governor Jim Doyle argued for "a comprehensive strategy to reduce cost, improve quality and expand access to affordable health care coverage" in Wisconsin in 2007, when he proposed the BadgerCare Plus program. Governor Christine Gregoire in 2007 in Washington proposed covering additional 32,000 children, and Governor Matt Blunt promised "to deliver improved access to affordable health care for every Missourian" in 2007. Holding other important factors constant, including the intergovernmental revenue that could drive increased city spending, these speeches seem to have a discernible independent effect on city fiscal outcomes.

What could explain these relationships? It is possible that the issues governors discuss in their speeches get amplified by the media and receive considerable public attention. Although not obliged to act, local officials in such a situation may feel it necessary to show their willingness to actively address social problems. For a rational politician concerned with reelection it may be a better strategy than passively waiting for a higher level of governments to intervene, which their political opponents will be quick to label as indifference. Policy advocates and interest groups may also use gubernatorial signals to justify their efforts to influence local spending choices. Recognition of a problem by state leadership may be able to boost their argument. This mechanism may be applicable to different policy areas beyond healthcare.

Whereas some gubernatorial policy issues are relevant for city spending decisions, others are not. These findings provide support to the notion that the policymaking process at the city level also varies by policy area. Previous research shows that the differences may be greater across policy areas within one system than across different systems (Breunig, Koski, & Mortensen, 2010;

89

Freeman, 1985; Lowi, 1972; K. B. Smith, 2002). This study demonstrates the idea that policy drives politics in state-city relations. Findings from one policy subsystem may not be generalizable to other subsystems. The empirical findings also support the notion that local decisions pertaining to long-term and short-term spending are likely driven by different considerations. Specific policy areas, such as public safety and social services, appear to be less affected by the political environment.

Another implication to be drawn from this study is that policy alignment between state and city governments may be more likely across jurisdictions with political similarities. City fiscal outcomes in the main policy areas, such as education and social services, are significantly associated with gubernatorial policy goals only when governor and city mayor identify with the same party. The more important the policy is to the governor, the more likely city officials are to increase their spending on those functions, holding IGR and other relevant factors constant. The relationship is different in other areas. These findings indicate that the informal gubernatorial policy signals may be more effective and more likely to bring about desired outcomes when dealing with political allies.

This study also demonstrates the utility of computational text analysis, which is an efficient way to annotate a large collection of documents and generate quantitative measures of their content. It allowed this study to test empirically the propositions that have been developed in the qualitative literature. Future studies may utilize it to infer the main issues discussed in presidential speeches, legislative deliberations, budgetary messages and other increasingly available digital documents relevant to public administration and policy research. The relationship between gubernatorial speeches and local behavior revealed in this study may be found between and within other levels of governments.

It is also important to note some of the limitations of this study and opportunities for future studies. The major improvement would be to include a wider range of cities in the analysis to improve the external validity of the study. Currently, the generalizability of the findings beyond the nonrandom sample of the largest cities in 48 states is limited. Future studies should include smaller cities as well as other types of local governments from all states. Another interesting question is whether or how the federal bureaucracy and lower levels of government react to presidential speeches.

CONCLUSION

The main goal of this dissertation has been to study gubernatorial policy agendas using computational text analysis and its influence on state and local fiscal policies. The first chapter applies an LDA topic model algorithm to a collection of gubernatorial speeches to infer the main themes that run through these documents. The main findings from this chapter suggest that machine learning algorithms can be successfully used to extract meaningful topics from large collections of political statements. Automated text analysis enables researchers to quantify and visualize text data in a much more efficient manner. The analysis revealed that LDA output fares well as a measure of gubernatorial policy agenda in terms of content and construct validity. Future researchers and practitioners should consider automated methods of text analysis instead of manual coding.

Some of the limitations of the topic model are its inability to distinguish the emotional side of the speeches. It gives a clear idea about the thematic structure of the text documents but has no means of separating positive and negative messages. Meanwhile, a topic may summarize positive achievements as well as highlight existing problems. It would be an improvement to be able to distinguish the dominant sentiment of each document and topic. This limitation could be addressed with supplementary analysis. One could, for example, use existing classifications, such as the National Research Council (NRC) sentiment lexicon (Mohammad & Turney, 2013), where words are annotated with their associated positive and negative sentiment and eight different emotions, and count the proportion of words associated with negative and positive sentiment in each document. An ex ante expectation is that on average the share of positive words will tend to exceed

the share of negative words because politicians generally seek to send an optimistic message. A researcher thus could use a difference between the share of positive and negative words to measure the sentiment of the speech. A positive value of this measure would indicate the message is mostly positive, a zero would mean a balanced statement, and values below zero would suggest predominance of negative sentiment. Similar technique could be applied to each distribution of words representing topics. These distributions, however, contain the entire vocabulary from the corpus with words order by their probability of appearing in the topic. If all words are counted in each distribution, the sentiment will be the same for all topics. Thus, researchers would need to focus on the most probable words for each topic and introduce some cut off point. At a minimum, this sentiment measure could be used as a control variable to account for sentiment in regression results. The benefits of such technique should be tested in future research.

The second chapter's findings indicate that governors may not be as influential in the budgeting games as depicted in the previous literature. The results suggest that gubernatorial policy priorities are related to state expenditures only in the area infrastructure and public safety. These spending categories may have geographically concentrated benefits, which makes them a good venue for bargaining with individual lawmakers. This study differs from the previous literature in three important ways. It uses the entire population of US states across multiple years, whereas most of the previous studies relied on nonrandom samples or short time periods. This work also makes use of computational text analysis to quantify gubernatorial policy priorities for high-profile speeches, whereas previous studies relied on manual coding with unclear classification criteria. Third, the influence of gubernatorial policy priorities is tested in individual policy areas rather than on the total size of state budgets. The results in this research agree with the previous research in that state chief executives do exert significant influence on state

policymaking but additionally suggest that this influence is not equal in all policy areas. This divergence may be due to different measures and/or sample selection. Given that this study uses all states across multiple years, previous findings of gubernatorial influence based on nonrandomly selected sates or cross-sectional data may have limited generazability.

Yet another possible explanation for the divergence in findings is that extent of bargaining and negotiations in state budgeting has been changing. Contemporary budgeting is often characterized by increased partisanship, gridlock, brinkmanship and hostage taking. State government shutdowns have become more common. The bargaining and negotiations put forward in the previous literature are being replaced with staunch opposition and forced compliance. Instead of seeking common ground and compromise, the more dominant strategy seems to be to use any available tools to cause harm and the threat of harm (Rubin, 2016, pp. xvi-xxi). As a result, on average no individual actor is able to influence the outcomes of state policymaking, even the governors. Additionally, as larger shares of state budgets are determined by formula, there are fewer resources that are subject to annual reconsideration.

This study, however, has recognizable limitations. Whereas total spending by policy area may not be correlated with gubernatorial policy goals, future research should use more granular data and test if there is any relationship with department or program level expenditures. Governors may be able to influence funding allocations among various programs, start new and/or hold existing initiatives, all of which may or may not affect the total level of expenditures. For example, the share of state education expenditures going to private education may change with gubernatorial efforts. This can be a serious policy shift, but it will not be revealed if the total education expenditures are analyzed. Another improvement would be to use more observations. This could be done by analyzing more granular data as described above, or/and by using a longer period.

Empirical evidence from the third chapter highlights the importance of gubernatorial policy goals for fiscal behavior of the largest cities in the United States. The findings suggest that city officials may align their spending decisions with chief executive policy goals and increase funding for those areas that matter to governors the most. This chapter reveals new ways of interaction between the two levels of government. Previous literature largely looked at the effect of state laws on local behavior, and this study reveals that city officials pay attention even to softer, unbinding signals. Given these findings, scholars seeking to understand the city policymaking process and outcomes may benefit from considering not only state laws and regulation, but also public statements of state officials' policy goals. Accounting for these softer policy signals may provide a better understanding of local behavior. In terms of implications for practice, state chief executives may find it useful to know their policy statements may influence local behavior. Public speeches can be used to persuade or "nudge" local officials into action at a much lower cost than a legislative bill.

The validity of these findings, however, is currently limited to the largest cities. Future research should test the extent to which these results are applicable to smaller jurisdictions. As with the previous chapter, another opportunity for researchers would be to test the relationship of the gubernatorial policy agenda with city expenditures at the department and/or program level. This would give a better understanding of how exactly the increased spending found in this study is used by programs.

In summary, the application of new methods in this dissertation provides new evidence complementing and challenging existing literature. The results from computational text analysis of over 500 speeches implanted in the first chapter have proven to be a reliable and efficient way to annotate large collections of documents with thematic content. These findings are in line with

the results from previously published manual analysis of the documents. The findings from the second chapter support existing literature portraying state chef executives as influential policy actors but point toward a changing landscape of the state budgeting process. Our view of gubernatorial influence may need to change and adapt to better reflect modern policymaking in the U.S. states. The third chapter highlight new ways of interaction between state and local policy processes. Whereas common perception suggests that state-local relations are primarily shaped by laws, mandates and financial invectives, it appears state leadership may be able to influence local policy choices through public statements of policy priorities. The findings from this research should be useful for researchers seeking more efficient approaches to text analysis and those interested in subnational policymaking.

REFERENCES

Abney, G., & Lauth, T. P. (1983). The governor as chief administrator. *Public Administration Review, 43*(1), 40-49.

Abney, G., & Lauth, T. P. (1985). The Line-Item Veto in the States: An Instrument for Fiscal Restraint or an Instrument for Partisanship? *Public Administration Review, 45*(3), 372-377. doi:10.2307/3109963

Abney, G., & Lauth, T. P. (1987). Perceptions of the Impact of Governors and Legislatures in the State Appropriations Process. *The Western Political Quarterly, 40*(2), 335-342. doi:10.2307/448316

Abney, G., & Lauth, T. P. (1997). The Item Veto and Fiscal Responsibility. *The Journal of Politics, 59*(3), 882-892. doi:10.2307/2998641

Abney, G., & Lauth, T. P. (1998). The End of Executive Dominance in State Appropriations. *Public Administration Review, 58*(5), 388-394. doi:10.2307/977548

Abney, G., & Lauth, T. P. (2002). Gubernatorial Use of the Item Veto for Narrative Deletion. *Public Administration Review, 62*(4), 492-503. doi:10.1111/0033-3352.00202

Alm, J., & Evers, M. (1991). The Item Veto and State Government Expenditures. *Public Choice, 68*(1/3), 1-15. doi:10.2307/30025364

Alt, J. E., & Lowry, R. C. (1994). Divided Government, Fiscal Institutions, and Budget Deficits: Evidence from the States. *The American Political Science Review, 88*(4), 811-828. doi:10.2307/2082709

Andersen, A. L., Lassen, D. D., & Nielsen, L. H. W. (2014). The impact of late budgets on state government borrowing costs. *Journal of Public Economics, 109*, 27-35. doi:http://doi.org/10.1016/j.jpubeco.2013.10.004

Angrist, J., & Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion* (1 ed.): Princeton University Press.

Archibold, R. (2010). Arizona Enacts Stringent Law on Immigration. Retrieved from http://www.nytimes.com/2010/04/24/us/politics/24immig.html

Avery, J. M., & Peffley, M. (2005). Voter Registration Requirements, Voter Turnout, and Welfare Eligibility Policy: Class Bias Matters. *State Politics & Policy Quarterly, 5*(1), 47-67. doi:10.1177/153244000500500103

Barrilleaux, C., & Berkman, M. (2003). Do Governors Matter? Budgeting Rules and the Politics of State Policymaking. *Political Research Quarterly, 56*(4), 409-417. doi:10.1177/106591290305600403

Barth, J., & Ferguson, M. R. (2002). American governors and their constituents: The relationship between gubernatorial personality and public approval. *State Politics & Policy Quarterly, 2*(3), 268-282.

Baumgartner, F. R., & Jones, B. D. (2010). *Agendas and Instability in American Politics, Second Edition*: University of Chicago Press.

Bemelmans-Videc, M.-L., Rist, R., & Vedung, E. (2011). *Carrots, Sticks, and Sermons: Policy Instruments and Their Evaluation*. New Brunswick, New Jersey: Transaction Publishers.

Bernick, E. L., & Wiggins, C. W. (1991). Executive-legislative relations: The governor's role as chief legislator. *Gubernatorial leadership and state policy*, 73-91.

Beyle, T. (2004). Governors. In V. Gray & R. Han (Eds.), *Politics in the American States* (8 ed., pp. 194-231). Washington, DC: CQ Press.

Birkland, T. A. (2011). *An introduction to the policy process : theories, concepts, and models of public policy making*. Armonk, N.Y.: M.E. Sharpe.

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

Blei, D. M., & Lafferty, J. D. (2007). A correlated topic model of science. *The Annals of Applied Statistics*, 17-35.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research, 3*(Jan), 993-1022.

Braithwaite, J. (2002). Rewards and regulation. *Journal of law and Society, 29*(1), 12-26.

Breunig, C., & Koski, C. (2009). Punctuated budgets and governors' institutional powers. *American Politics Research, 37*(6), 1116-1138.

Breunig, C., Koski, C., & Mortensen, P. B. (2010). Stability and Punctuations in Public Spending: A Comparative Study of Budget Functions. *Journal of Public Administration Research and Theory, 20*(3), 703-722. doi:10.1093/jopart/mup028

Bromley-Trujillo, R. (2017). Despite Trump, many cities and states are fighting climate change. Including Pittsburgh. Retrieved from https://www.washingtonpost.com/news/monkey-cage/wp/2017/06/06/despite-trump-pittsburghs-working-on-slowing-climate-change-so-are-many-other-cities-and-states/?utm_term=.bc0af229c436

Brown, R. D. (1995). Party Cleavages and Welfare Effort in the American States. *American Political Science Review, 89*(1), 23-33. doi:10.2307/2083072

Brown, T. (2000). Constitutional Tax and Expenditure Limitation in Colorado: The Impact on Municipal Governments. *Public Budgeting & Finance, 20*(3), 29-50. doi:10.1111/0275-1100.00019

Buschman, R. D., & Sjoquist, D. L. (2017). An Exploration of Differential State Responses to Changes in Fiscal Conditions. *Public Budgeting & Finance, 37*(1), 47-67. doi:10.1111/pbaf.12150

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources, 50*(2), 317-372.

Carr, J. B., & Karuppusamy, S. (2010). Reassessing the Link Between City Structure and Fiscal Policy: Is the Problem Poor Measures of Governmental Structure? *The American Review of Public Administration, 40*(2), 209-228. doi:10.1177/0275074009334641

Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J. L., & Blei, D. M. (2009). *Reading tea leaves: How humans interpret topic models.* Paper presented at the Advances in neural information processing systems.

Cockerham, A. G., & Crew, R. E. (2017). Factors Affecting Governors' Decisions to Issue Executive Orders. *State and Local Government Review, 49*(1), 6 - 14. doi:doi:10.1177/0160323X17702702

Coe, C. K. (2008). Preventing Local Government Fiscal Crises: Emerging Best Practices. *Public Administration Review, 68*(4), 759-767. doi:10.1111/j.1540-6210.2008.00913.x

Coffey, D. (2005). Measuring Gubernatorial Ideology: A Content Analysis of State of the State Speeches. *State Politics & Policy Quarterly, 5*(1), 88-103. doi:doi:10.1177/153244000500500105

Council of State Governments. (2017). *The Book of the States*. Lexington, KY: Council of State Governments.

Crew, R. (1992). Understanding Gubernatorial Behavior: A Framework for Analysis. In T. Beyle (Ed.), *Governors and Hard Times*. Washington, DC: Congressional Quarterly Press.

Cukier, K., & Mayer-Schoenberger, V. (2013). The rise of big data: How it's changing the way we think about the world. *Foreign Aff., 92*, 28.

DeVellis, R. F. (2003). *Scale Development: Theory and Applications* (M. H. Seawell Ed. 2 ed. Vol. 26). USA: SAGE Publications.

DiLeo, D. (2001). To Develop or to Redistribute? An Analysis of the Content of Governors' Agendas. *State and Local Government Review, 33*(1), 52-59. doi:doi:10.1177/0160323X0103300105

Dilger, R. J., Krause, G. A., & Moffett, R. R. (1995). State legislative professionalism and gubernatorial effectiveness, 1978-1991. *Legislative Studies Quarterly*, 553-571.

Dometrius, N. C., & Wright, D. S. (2010). Governors, Legislatures, and State Budgets across Time. *Political Research Quarterly, 63*(4), 783-795.

Dye, T. (1966). *Politics, economics, and the public: Policy outcomes in the American states*: Chicago: Rand McNally.

Dye, T. (1984). Party and Policy in the States. *The Journal of Politics, 46*(4), 1097-1116.

Feiock, R. C., Jeong, M.-G., & Kim, J. (2003). Credible Commitment and Council-Manager Government: Implications for Policy Instrument Choices. *Public Administration Review, 63*(5), 616-625. doi:10.1111/1540-6210.00324

Fellowes, M. C., & Rowe, G. (2004). Politics and the New American Welfare States. *American Journal of Political Science, 48*(2), 362-373. doi:10.1111/j.0092-5853.2004.00075.x

Ferguson, M. (2003). Chief Executive Success in the Legislative Arena. *State Politics & Policy Quarterly, 3*(2), 158-182. doi:doi:10.1177/153244000300300203

Ferguson, M., & Barth, J. (2002). Governors in the Legislative Arena: The Importance of Personality in Shaping Success. *Political Psychology, 23*(4), 787-808.

Ferguson, M., & Bowling, C. J. (2008). Executive Orders and Administrative Control. *Public Administration Review, 68*, S20-S28. doi:10.1111/j.1540-6210.2008.00975.x

Fletcher, M. A., & Weiner, R. (2011). Minnesota governor, GOP lawmakers agree to end shutdown. Retrieved from https://www.washingtonpost.com/business/economy/governor-gop-lawmakers-agree-to-end-shutdown/2011/07/14/gIQAP1brEI_story.html?utm_term=.869bd340b46a

Frederickson, H. G., & Johnson, G. A. (2001). The Adapted American City: A Study of Institutional Dynamics. *Urban Affairs Review, 36*(6), 872-884. doi:10.1177/10780870122185127

Freeman, G. (1985). National styles and policy sectors: Explaining structured variation. *Journal of Public Policy, 5*(4), 467-496.

Garand, J. C. (1988). Explaining government growth in the US states. *American Political Science Review, 82*(03), 837-849.

Goodman, D. (2007). Determinants of Perceived Gubernatorial Budgetary Influence among State Executive Budget Analysts and Legislative Fiscal Analysts. *Political Research Quarterly, 60*(1), 43-54.

Gosling, J. J. (1986). Wisconsin Item-Veto Lessons. *Public Administration Review, 46*(4), 292-300. doi:10.2307/976302

Gray, V., Hanson, R., & Kousser, T. (2017). *Politics in the American States: A Comparative Analysis*: SAGE Publications.

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National academy of Sciences, 101*(suppl 1), 5228-5235.

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.

Gujarati, D. (2007). *Basic Econometrics* (4 ed.). New York, NY: McGraw-Hill Education.

Halaby, C. N. (2004). Panel models in sociological research: Theory into practice. *Annu. Rev. Sociol., 30*, 507-544.

Hansen, S., McMahon, M., & Prat, A. (2014). Transparency and deliberation within the FOMC: a computational linguistics approach. *The Quarterly Journal of Economics*.

Heidbreder, B. (2012). Agenda Setting in the States: How Politics and Policy Needs Shape Gubernatorial Agendas. *Politics & Policy, 40*(2), 296-319. doi:10.1111/j.1747-1346.2012.00345.x

Heidbreder, B., & Scheurer, K. F. (2013). Gender and the Gubernatorial Agenda. *State & Local Government Review, 45*(1), 3-13.

Herzik, E. (1991). Policy Agendas and Gubernatorial Leadership. In E. Herzik & B. Brown (Eds.), *Gubernatorial leadership and state policy*. New York, NY: Greenwood Press.

Hofferbert, R. I. (1966). The relation between public policy and some structural and environmental variables in the American states. *American Political Science Review, 60*(1), 73-82.

Holtz-Eakin, D. (1988). The line item veto and public sector budgets: Evidence from the states. *Journal of Public Economics, 36*(3), 269-292. doi:https://doi.org/10.1016/0047-2727(88)90011-4

Jackson, J. E., & Kingdon, J. W. (1992). Ideology, Interest Group Scores, and Legislative Votes. *American Journal of Political Science, 36*(3), 805-823. doi:10.2307/2111592

Jelodar, H., Wang, Y., Yuan, C., & Feng, X. (2017). Latent Dirichlet Allocation (LDA) and Topic modeling: models, applications, a survey. *arXiv preprint arXiv:1711.04305*.

Jewell, M. E., & Whicker, M. L. (1994). *Legislative leadership in the American states*: University of Michigan Press.

Jones, B., Baumgartner, F. R., Breunig, C., Wlezien, C., Soroka, S., Foucault, M., . . . Walgrave, S. (2009). A General Empirical Law of Public Budgets: A Comparative Analysis. *American Journal of Political Science, 53*(4), 855-873. doi:10.1111/j.1540-5907.2009.00405.x

Joyce, P. G., & Mullins, D. R. (1991). The Changing Fiscal Structure of the State and Local Public Sector: The Impact of Tax and Expenditure Limitations. *Public Administration Review, 51*(3), 240-253. doi:10.2307/976948

Karuppusamy, S., & Carr, J. B. (2012). Interjurisdictional Competition and Local Public Finance: Assessing the Modifying Effects of Institutional Incentives and Fiscal Constraints. *Urban Studies, 49*(7), 1549-1569. doi:10.1177/0042098011415435

Kousser, T. (2002). The politics of discretionary Medicaid spending, 1980–1993. *Journal of Health Politics, Policy and Law, 27*(4), 639-672.

Kousser, T., & Phillips, J. (2012). *The Power of American Governors: Winning on Budgets and Losing on Policy*. New York, NY: Cambridge University Press.

Kousser, T., & Phillips, J. H. (2009). Who Blinks First? Legislative Patience and Bargaining with Governors. *Legislative Studies Quarterly, 34*(1), 55-86.

Krause, G. A., & Melusky, B. F. (2012). Concentrated Powers: Unilateral Executive Authority and Fiscal Policymaking in the American States. *The Journal of Politics, 74*(1), 98-112. doi:10.1017/s0022381611001149

Lauth, T. P. (2016). The Other Six: Governors Without The Line-Item Veto. *Public Budgeting & Finance, 36*(4), 26-49. doi:10.1111/pbaf.12115

Light, P. C. (1999). *The President's Agenda: Domestic Policy Choice from Kennedy to Clinton*: Johns Hopkins University Press.

Lipson, L. (1939). *The American Governor from Figurehead to Leader* (Vol. 9): University of Chicago Press.

Liu, B. (2015). *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions* (1 ed.). New York, NY: Cambridge University Press.

Lowi, T. J. (1972). Four Systems of Policy, Politics, and Choice. *Public Administration Review, 32*(4), 298-310. doi:10.2307/974990

McCallister, D. (2017). Texas Gov. Abbott Signs Measure Targeting 'Sanctuary Cities'. Retrieved from http://www.npr.org/sections/thetwo-way/2017/05/08/527344804/texas-gov-abbott-signs-measure-targeting-sanctuary-cities

Meier, K. (2016). *The Politics of Sin: Drugs, Alcohol and Public Policy: Drugs, Alcohol and Public Policy*: Taylor & Francis.

Meyer, N. (2013). Political contestation of self-regulation in the shadow of hierarchy. *Journal of European Public Policy, 20*(5), 760-776.

Mikesell, J. (2016). *Fiscal Administration* (10 ed.). Boston, MA: Cengage Learning.

Mitchell, D., & Thurmaier, K. (2012). Currents and Undercurrents in Budgeting Theory: Exploring the Swirls, Heading upstream. In *Foundations of Public Administration Series*: PAR/ASPA.

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational Intelligence, 29*(3), 436-465.

Morehouse, S. M. (1998). *The governor as party leader: Campaigning and governing*: University of Michigan Press.

Mullins, D. R., & Joyce, P. G. (1996). Tax and expenditure limitations and state and local fiscal structure: An empirical assessment. *Public Budgeting & Finance, 16*(1), 75-101.

Mullins, D. R., & Wallin, B. A. (2004). Tax and Expenditure Limitations: Introduction and Overview. *Public Budgeting & Finance, 24*(4), 2-15. doi:10.1111/j.0275-1100.2004.00344.x

NASBO. (2017). State Expenditure Report. Retrieved from https://www.nasbo.org/mainsite/reports-data/state-expenditure-report

National Governors Association. (2010). *The Many Roles of the Governor's Chief of Staff*. Washington, D.C.: National Governors Association.

National Governors Association. (2015). *Transition and The New Governor: A Planning Guide*. Washington, D.C.: National Governors Association.

NCSL Research. (2016). State Budget Summaries. *Budget and Tax Program.* Retrieved from http://www.ncsl.org/research/fiscal-policy/fy-2018-budget-status.aspx

Peters, G. (2015). *American Public Policy: Promise and Performance*: SAGE Publications.

Peterson, M. (1990). Legislating Together Cambridge. *MA: HarvardUniver-sityPress*.

Peterson, P. E. (1995). *The Price of Federalism*. Washington, D.C.: Brookings Institution.

Preston, A. E., & Ichniowski, C. (1991). A NATIONAL PERSPECTIVE ON THE NATURE AND EFFECTS OF THE LOCAL PROPERTY TAX REVOLT, 1976-1986. *National Tax Journal, 44*(2), 123-145.

Reese, C. C. (1997). The line-item veto in practice in ten southern states. *Public Administration Review*, 510-516.

Rodero Antón, E. (2012). A comparative analysis of speech rate and perception in radio bulletins. *Text & Talk. 2012 June 12; 32 (3): 391-411*.

Rosenthal, A. (1990). *Governors and Legislators: Contending Powers*. Washington,
DC: Congressional Quarterly Press.

Rosenthal, A. (2012). *The Best Job in Politics: Exploring How Governors Succeed as Policy Leaders*. United States: CQ Press.

Rubin, I. (2009). *The Politics of Public Budgeting: Getting and Spending, Borrowing and Balancing* (6 ed.). Washington, DC: SAGE Publications.

Rubin, I. (2016). *The Politics of Public Budgeting: Getting and Spending, Borrowing and Balancing*: SAGE Publications.

Ryu, J. E., Bowling, C. J., CHO, C. L., & Wright, D. S. (2008). Exploring explanations of state agency budgets: Institutional budget actors or exogenous environment? *Public Budgeting & Finance, 28*(3), 23-47.

Sampathkumar, M. (2017). Paris Agreement: Over a thousand US CEOs and mayors say they will continue to fight climate change despite Trump. Retrieved from http://www.independent.co.uk/news/world/americas/us-politics/paris-agreement-trump-climate-change-ceos-mayors-support-despite-decision-pact-we-are-still-in-a7774511.html

Schick, A. (1966). The Road to PPB: The Stages of Budget Reform. *Public Administration Review, 26*(4), 243-258. doi:10.2307/973296

Segal, J. A., & Cover, A. D. (1989). Ideological Values and the Votes of U.S. Supreme Court Justices. *The American Political Science Review, 83*(2), 557-565. doi:10.2307/1962405

Shadbegian, R. J. (1998). Do Tax and Expenditure Limitations Affect Local Government Budgets? Evidence From Panel Data. *Public Finance Review, 26*(2), 118-136. doi:10.1177/109114219802600202

Shadbegian, R. J. (1999). The Effect of Tax and Expenditure Limitations on the Revenue Structure of Local Government, 1962–87. *National Tax Journal, 52*(2), 221-237. doi:10.2307/41789391

Sigelman, L., & Dometrius, N. C. (1988). Governors as chief administrators: The linkage between formal powers and informal influence. *American Politics Quarterly, 16*(2), 157-170.

Singleton, R., & Straits, B. C. (2010). *Approaches to Social Research* (5 ed.): Oxford University Press.

Smith, K. B. (2002). Typologies, Taxonomies, and the Benefits of Policy Classification. *Policy Studies Journal, 30*(3), 379-395. doi:10.1111/j.1541-0072.2002.tb02153.x

Smith, M. A. (1997). The nature of party governance: Connecting conceptualization and measurement. *American Journal of Political Science*, 1042-1056.

Svara, J. H. (2005). Exploring Structures and Institutions in City Government. *Public Administration Review, 65*(4), 500-506. doi:10.1111/j.1540-6210.2005.00475.x

Szalay, A., & Gray, J. (2006). 2020 Computing: Science in an exponential world. *Nature, 440*(7083), 413.

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: improving decisions about health, wealth, and happiness* (Rev. and expanded ed.). New York: Penguin Books.

The Morning Call. (2017). Pennsylvania Senate eyes House budget plan that's short of what it wanted. Retrieved from http://www.mcall.com/news/nationworld/pennsylvania/mc-nws-pennsylvania-budget-standoff-senate-20171023-story.html

Thompson, J. A. (1987). Agency Requests, Gubernatorial Support, and Budget Success in State Legislatures Revisited. *The Journal of Politics, 49*(03), 756-779. doi:doi:10.2307/2131277

U.S. Census Bureau. (2017). State & Local Government Finances. Retrieved from https://www.census.gov/govs/local/

US Census. (2018). State Government Finances Tables. *Annual Survey of State Government Finances.* Retrieved from https://www.census.gov/programs-surveys/state/data/tables.html

Van Assendelft, L. A. (1997). *Governors, agenda setting, and divided government*. Lanham, MD: University Press of America.

Vijayarani, S., Ilamathi, M. J., & Nithya, M. (2015). Preprocessing techniques for text mining-an overview. *International Journal of Computer Science & Communication Networks, 5*(1), 7-16.

Weber, R., & Brace, P. (1999). *American State and Local Politics: Directions for the 21st Century*. New York, NY: Chatham House Publishers, Seven Bridges Press.

Weng, J., Lim, E.-P., Jiang, J., & He, Q. (2010). *Twitterrank: finding topic-sensitive influential twitterers.* Paper presented at the Proceedings of the third ACM international conference on Web search and data mining.

Willoughby, K. (2008). The State of the States: Governors Keep Agendas Short. *The book of the states*, 157-164.

Willoughby, K. (2009). The State of the States: Governors Mirror New Administration Priorities. *The book of the states*, 161-168.

Willoughby, K. (2014). *Public Budgeting in Context: Structure, Law, Reform and Results*: Wiley.

Willoughby, K. (2016). The State of the State Addresses: Governors in the Hot Seat. *The book of the states*, 137-146.

Willoughby, K. (2017). The State of the State Addresses: Governors Hesitant in Promising More. *The book of the states, 49*, 151 - 158.

Winters, R. (1976). Party control and policy change. *American Journal of Political Science*, 597-636.

Wolman, H., McManmon, R., Bell, M., & Brunori, D. (2010). Comparing Local Government Autonomy Across States. In M. Bell, D. Brunori, & J. Youngman (Eds.), *Property Tax and Local Autonomy* (pp. 69-119). Cambridge, MA: Lincoln Institute of Land Policy.

Wood, C. (2011). Understanding the Consequences of Municipal Discretion. *The American Review of Public Administration, 41*(4), 411-427. doi:10.1177/0275074010381111

Wu, Y., Liu, M., Zheng, W. J., Zhao, Z., & Xu, H. (2012). Ranking gene-drug relationships in biomedical literature using latent dirichlet allocation. In *Biocomputing 2012* (pp. 422-433): World Scientific.

Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E.-P., Yan, H., & Li, X. (2011). *Comparing twitter and traditional media using topic models.* Paper presented at the European Conference on Information Retrieval.

APPENDIX A. LDA TOPIC MODEL

The following description of the Latent Dirichlet Allocation topic model is based on the original publication (Blei et al., 2003). The intuition behind LDA is that documents may contain multiple topics. The main goal of LDA is to automatically discover these topics from the observed collection of documents. This collection is often referred to as a corpus. The inferred hidden topic structures generally resemble the thematic structure of the documents (Blei, 2012). A topic is formally described as a distribution over a fixed vocabulary and each document is represented as a probability distribution over topics. More formally, the definitions of a document and a corpus are:

- **Document** – Each document a presented as a sequence of N terms $\mathbf{w} = (w_1, w_2, \dots, w_N)$.

- **Corpus** - The entire collection of D documents.

LDA assumes that each document in the corpus is a random mixture of k topics. Although LDA does not require any prior knowledge of the content of the documents, the number of unique topics k expected to be found in the corpus needs to be specified. As it has been stated, the primary guide for researchers in selecting the appropriate number of topics should be theoretical expectations, expert knowledge and substantive meaning of the results. Additionally, researchers can estimate several models with different numbers of topics and assess their perplexity. Perplexity is a theoretical measure of how well a topic model predicts observed data. The lower the perplexity score, the better fit to the data. Researcher may select the number of topics at which the marginal perplexity score stops decreasing (Blei & Lafferty, 2007). It is worth mentioning that parsimony is important. Experiments show that topics become more fine-grained and less useful for humans

with larger number of estimated topics (Chang et al., 2009). Hence, the perplexity score should complement, rather than replace, other criteria mentioned above.

The share of each document devoted to each topic is represented by θ_{ad}. Figure 1 represents a graphical illustration of the LDA model. The boxes are "plates" representing replicates. The outer plate represents document D, while the inner plate represents a repeated choice of topics and words within a document (Blei et al., 2003, p. 997). As it has been noted, the only observed variable is the collection of words in the documents. All other variables are latent, and the main computational task for LDA is to infer these latent variables based on the observed words. LDA assumes that each word $w_{d,n}$ in each document is generated from both a distribution over latent topics, which corresponds to each speech and a distribution over words which comprises the topics.
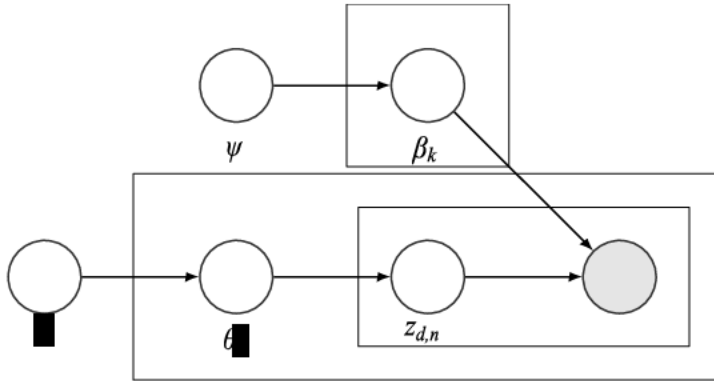

Figure 1. Graphical model representation of LDA.

In the graphical model above, the parameters are defined as:

1. $\beta_k \sim Dir(\psi)$, where $k \in \{1, \ldots, K\}$ is the distribution over words that defines each of the latent topics assumed to be found in the entire corpus of documents.

2. $\theta_d \sim Dir(\alpha)$, where $d \in \{1, \ldots, N_{st}\}$ is the distribution over topics for each document.

3. $z_{d,n}$ is topic assignment of the nth word in the dth document.

4. $w_{d,n}$ is the nth words of the dth document.

The probability distributions of topic proportions for each document $p(\theta_d \mid \alpha)$ and each topic in all documents $p(\beta_k \mid \psi)$ are distributed Dirichlet with corpus-level, or hyper-parameters α and ψ respectively. Thus, topic proportions for each document has the distribution:

$$p(\theta_d \mid \alpha) = \frac{\prod_{i=1}^{k} \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^{K} \alpha_i)} \prod_{i=1}^{k} \theta_{di}^{\alpha_i - 1}$$

Whereas each topic k across all documents as the distribution over words:

$$p(\beta_k \mid \psi) = \frac{\prod_{i=1}^{N} \Gamma(\psi_i)}{\Gamma(\sum_{i=1}^{N} \psi_i)} \prod_{i=1}^{N} \beta_{ki}^{\psi_i - 1}$$

The remaining distributions needed to specify the model including topic assignment conditional on topic distribution $p(z_{d,n} \mid \theta_d)$ and word conditional on topic assignment $p(w_{d,n} \mid z_{d,n}, \beta_k)$ are multinomial with:

$z_{d,n} \sim Multinom(\theta_d)$

$w_{d,n} \sim Multinom(\beta_k)$

Thus, the fully specified topic model over the corpus of documents is:

$$p(\theta, \mathbf{z}, \mathbf{w}, \beta \mid \psi, \alpha) = \prod_{k=1}^{15} p(\beta_k \mid \psi) \prod_{d=1}^{D} \left(p(\theta_d \mid \alpha) \prod_{n=1}^{N} p(z_{d,n} \mid \theta_d) \, p(w_{d,n} \mid z_{d,n}, \beta_k) \right)$$

The LDA algorithm seeks optimal topic assignment $z_{d,n}$ for each word in each document, and the optimal word probabilities β_k for each topic that maximizes the likelihood. This task, however, is computationally unfeasible because it would require summing over all possible topic assignment for all words in all documents. Therefore, alternative methods of the likelihood function approximation are necessary. The most commonly used method for topic modeling is

106

Gibbs sampling (Blei, 2012, p. 81) developed by Griffiths and Steyvers (2004). This approximation method is employed in this study.[31]

---

[31] LDA can be implemented in Stata (*ldagibbs* package), R (*tm* package) and Python (*lda* package).

APPENDIX B: DEFINITIONS OF U.S. CENSUS EXPENDITURE CATEGORIES

Education – Schools, colleges, and other educational institutions, and educational programs for adults, veterans, and other special classes.

Health – Outpatient health services, other than hospital care, including: public health administration; research and education; categorical health programs; treatment and immunization clinics; nursing; environmental health activities such as air and water pollution control; ambulance service if provided separately from fire protection services, and other general public health activities such as mosquito abatement.

Hospitals – Financing, construction acquisition, maintenance or operation of hospital facilities, provision of hospital care, and support of public or private hospitals.

Highways – Construction, maintenance, and operation of highways, streets, and related structures, including toll highways, bridges, tunnels, ferries, street lighting and snow and ice removal.

Police – Preservation of law and order and traffic safety. Includes police patrols and communications, crime prevention activities, detention and custody of persons awaiting trial, traffic safety, and vehicular inspection

Correction – Confinement and correction of adults and minors convicted of offenses against the law, and pardon, probation, and parole activities.

Natural resources – Conservation, promotion, and development of natural resources, such as soil, water, forests, minerals, and wildlife. Includes irrigation, drainage, flood control, forestry

and fire protection, soil reclamation, soil and water conservation, fish and game programs, and agricultural fairs.

APPENDIX C. LIST OF 149 CITIES FROM THE WORKING SAMPLE

City	State	City	State	City	State
Birmingham	Alabama	Gary	Indiana	Akron	Ohio
Mobile	Alabama	Indianapolis	Indiana	Cincinnati	Ohio
Montgomery	Alabama	Cedar Rapids	Iowa	Cleveland	Ohio
Anchorage	Alaska	Des Moines	Iowa	Columbus	Ohio
Fairbanks	Alaska	Kansas City	Kansas	Dayton	Ohio
Mesa	Arizona	Topeka	Kansas	Toledo	Ohio
Phoenix	Arizona	Wichita	Kansas	Oklahoma City	Oklahoma
Tucson	Arizona	Lexington	Kentucky	Tulsa	Oklahoma
Ft. Smith	Arkansas	Louisville	Kentucky	Eugene	Oregon
Little Rock	Arkansas	Baton Rouge	Louisiana	Portland	Oregon
Anaheim	California	New Orleans	Louisiana	Salem	Oregon
Bakersfield	California	Shreveport	Louisiana	Philadelphia	Pennsylvania
Fremont	California	Lewiston	Maine	Pittsburgh	Pennsylvania
Fresno	California	Portland	Maine	Providence	Rhode Island
Huntington Beach	California	Baltimore	Maryland	Warwick	Rhode Island
Long Beach	California	Frederick	Maryland	Charleston	South Carolina
Los Angeles	California	Boston	Massachusetts	Columbia	South Carolina
Modesto	California	Springfield	Massachusetts	Rapid City	South Dakota
Oakland	California	Worcester	Massachusetts	Sioux Falls	South Dakota
Riverside	California	Detroit	Michigan	Chattanooga	Tennessee
Sacramento	California	Flint	Michigan	Knoxville	Tennessee
San Diego	California	Grand Rapids	Michigan	Memphis	Tennessee
San Francisco	California	Warren	Michigan	Nashville	Tennessee
San Jose	California	Minneapolis	Minnesota	Arlington	Texas
Santa Ana	California	St. Paul	Minnesota	Austin	Texas
Stockton	California	Gulfport	Mississippi	Corpus Christi	Texas
Aurora	Colorado	Jackson	Mississippi	Dallas	Texas
Colorado Springs	Colorado	Kansas City	Missouri	El Paso	Texas
Denver	Colorado	St. Louis	Missouri	Ft. Worth	Texas
Bridgeport	Connecticut	Billings	Montana	Garland	Texas
Hartford	Connecticut	Missoula	Montana	Houston	Texas
New Haven	Connecticut	Lincoln	Nebraska	Lubbock	Texas
Dover	Delaware	Omaha	Nebraska	San Antonio	Texas
Wilmington	Delaware	Las Vegas	Nevada	Provo	Utah
Ft. Lauderdale	Florida	Reno	Nevada	Salt Lake City	Utah
Hialeah	Florida	Manchester	New Hampshire	Burlington	Vermont
Jacksonville	Florida	Nashua	New Hampshire	Rutland	Vermont
Miami	Florida	Albuquerque	New Mexico	Chesapeake	Virginia
Orlando	Florida	Las Cruces	New Mexico	Norfolk	Virginia
St. Petersburg	Florida	Buffalo	New York	Richmond	Virginia
Tallahassee	Florida	New York	New York	Virginia Beach	Virginia
Tampa	Florida	Rochester	New York	Seattle	Washington
Atlanta	Georgia	Syracuse	New York	Spokane	Washington
Columbus	Georgia	Yonkers	New York	Tacoma	Washington
Boise	Idaho	Charlotte	North Carolina	Charleston	West Virginia
Nampa	Idaho	Durham	North Carolina	Huntington	West Virginia
Aurora	Illinois	Greensboro	North Carolina	Madison	Wisconsin
Chicago	Illinois	Raleigh	North Carolina	Milwaukee	Wisconsin
Ft. Wayne	Indiana	Bismarck	North Dakota	Casper	Wyoming
		Fargo	North Dakota	Cheyenne	Wyoming