# Min(d)ing the President

# Min(d)ing the President: A text analytic approach to measuring tax news

Adam Jassem*[1], Lenard Lieb[2], Rui Jorge Almeida[1,3], Nalan Baştürk[1], and Stephan Smeekes[1]

[1]Department of Quantitative Economics, Maastricht University
[2]Department of Macro, International and Labour Economics, Maastricht University
[3]Department of Data Analytics and Digitilisation, Maastricht University

April 8, 2021

**Abstract**

We propose a novel text-analytic approach for incorporating textual information into structural economic models and apply this to study the effects of tax news. We first develop a novel semi-supervised two-step topic model that automatically extracts specific information regarding future tax policy changes from text. We also propose an approach for transforming such textual information into an economically meaningful time series to be included in a structural econometric model as variable of interest or instrument. We apply our method to study the effects of fiscal foresight, in particular the informational content in speeches of the U.S. president about future tax reforms, and find that our semi-supervised topic model can successfully extract information about the direction of tax changes. The extracted information predicts (exogenous) future tax changes and contains signals that are not present in previously considered (narrative) measures of (exogenous) tax changes. We find that tax news triggers a significant yet delayed response in output.

*Keywords:* News, fiscal foresight, tax shocks, identification, text mining, topic model, Latent Dirichlet Allocation.
*JEL Codes:* C11, E62, H30, Z13

*Corresponding author. Department of Quantitative Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Email: a.jassem@maastrichtuniversity.nl

# 1 Introduction

In this paper we propose a novel approach for incorporating textual information into structural economic models with three main contributions. First, we develop a two-step semi-supervised topic model that allows for the automatic extraction of specific information relevant to future policy changes from textual data. Second, we provide methodology for transforming such textual information into an economically meaningful time series to be included in a structural econometric model as variable of interest or instrument. Third, we apply our method to study fiscal foresight by extracting information in speeches of the U.S. president about future tax reforms. With our approach we extract information that is highly predictive for future (exogenous) tax changes, and show that economic agents react to these early signals.

News about future policy changes is likely to have effects today. When receiving new information, economic agents' forward-looking behavior implies that economic decisions are also contemporaneously affected. There is a growing empirical literature supporting these findings.[1] Taking foresight into account is particularly important when analyzing the effects of fiscal policy: economic agents typically receive clear signals about future tax reforms long before a particular bill is enacted. Yang (2007) shows that, in the U.S., almost all implemented tax changes were preceded by legislative lags ranging between a quarter and three years. An illustrative example of how fiscal foresight impacts economic behavior is the Tax Reform Act of 1986. This Act stipulated an increase of the effective maximum tax rate on capital gains from 20% to 28%, to be implemented in the following year. As a result, tax revenues from capital gains jumped by 90% before the act entered into force (Auerbach and Slemrod, 1997). Assessing the effect of this particular tax reform is flawed if fiscal foresight is not taken into account, and one may easily confuse directions of causality.

As a general critique, Leeper et al. (2013) argue that standard empirical approaches such as structural vector autoregressions (SVARs) often fail to properly account for fiscal foresight, as the variables typically included do not span the full information set available to economic agents. The estimated effects of tax changes using such models may as a result be biased and misleading; tax multipliers might

[1] For example (Jaimovich and Rebelo, 2009; Beaudry and Portier, 2014; Forni et al., 2017) investigate news shocks related to future productivity, Ramey (2011) considers fiscal news, monetary policy news shocks are analyzed in Nakamura and Steinsson (2018); see also Ramey (2016) for a comprehensive overview of the recent literature.

even be of the wrong sign. Several studies therefore instead use a narrative approach to trace the arrival of information on future tax policy changes from the legislative process. Romer and Romer (2009a, 2010) (henceforth RR) identify the key motivation behind all legislated post-war tax changes in the U.S. and determine their impact on government revenues. Using various sources of narrative records such as presidential speeches and Congressional records they identify those tax changes that are not systematically related to changes in output, and classify them as *exogenous*. RR additionally define a series of tax *news* as the present value of tax changes discounted back to the date of enactment (in contrast to the exogenous tax changes which are assigned to the implementation date), with the aim of capturing anticipation effects. Mertens and Ravn (2012, 2014) further split the RR series of exogenous tax changes into two components and use these to identify anticipated and unanticipated tax shocks. They classify a tax change in the RR series to be unanticipated if the specific tax law took effect, i.e. was implemented, less than 90 days after the corresponding bill was enacted.

In our view neither RR's tax news nor the Mertens and Ravn (2012, 2014) approach are successful in fully capturing anticipation effects. This is because they do not get the timing of arrival of new information right. In addition, narrative approaches that identify tax shocks solely from tax changes that were eventually implemented do not take into account that a news signal might be noisy and that policy plans, and alongside them economic agents' expectations, are subject to revision over time. Before a tax bill is signed into law, information on the exact design of the tax reform is imprecise. Moreover, some proposals of tax changes initially put forward by the administration never come to fruition at all, while nonetheless influencing expectations.

While similar to RR, our approach is conceptually different in two important aspects. First, while analyzing similar auxiliary sources of data, our goal is not to search for the motivation behind each tax change but rather to trace the arrival of information regarding future tax reforms. In the United States, the president is the main driving force behind tax policy legislation (Yang, 2007; Romer and Romer, 2010). Presidential speeches are therefore a highly informative source for signals about future tax reforms. We review the US president's speeches and communications and quantify how prominently tax reforms, as well as their direction (cuts vs. hikes), are featured on the political agenda at a given point in time. To implement this

idea we build on the following hypothesis: if a policy maker repeatedly emphasises the importance of future tax changes, the public should expect that these changes will likely be implemented in the near future. The result is a measure of *prevalence* capturing the importance of tax cuts and hikes on the administration's policy agenda.

Second, to construct this prevalence measure from a considerable number of documents (we consider *all* of the President's public statements since 1949), we introduce a novel semi-supervised text-analytic approach. Using a standard, unsupervised Latent Dirichlet Allocation (LDA) topic model we would be able to quantify the prevalence of specific political issues, which we henceforth refer to as topics. However, while we can distinguish speeches about tax reforms from speeches on other topics, further differentiating between tax hikes and tax cuts is not possible. We thus propose an alternative strategy to feed additional information into the LDA topic model to construct more informative priors for the tax hike/cut topics. To this aim, we combine lexical knowledge of a priori selected terms related to the direction of tax changes, with the results from an unsupervised model. It is important to stress, that our dictionary of selected terms only 'nudges' the model towards the terms we deem to be important: the topic estimation remains data-driven and robust to misspecification.

We find that our tax prevalence series predict future federal tax reforms regardless of how we measure tax changes; cyclically adjusted revenue changes, as well as all RR narrative measures are Granger-caused by our prevalence series. In contrast, we do not find evidence for predictability in the opposite direction. Interestingly, our series also Granger-cause implicit tax rates, which are often used to proxy tax news in the US. Again, we do not find evidence for predictability in the opposite direction. These findings indeed suggest that our tax prevalence series capture the timing of the arrival of tax news more accurately. While we do not explicitly address the issue of identifying *exogenous* tax changes, we do not find evidence that our tax prevalence series are driven by business cycle conditions in the (recent) past. Further, we use our constructed prevalence series to draw meaningful conclusions. We propose how the present value of a potential policy action, in monetary terms, could be constructed from the two prevalence measures. This measure we label as *noisy tax news*[2] and investigate its effects on output. In contrast to Leeper et al. (2013) or Mertens and

[2]We use the term *noisy tax news* to differentiate our measure from other suggested measures on perfectly anticipated tax changes (or shocks). By constructing an *ex-ante* news signal we explicitly allow for such a measure to pick up noise (or *ex-post* misleading information). The term *noisy tax news* is inspired by the discussion in Forni et al. (2017).

Ravn (2012) who analyze the effects of anticipated tax changes, we do not find that noisy news about future tax cuts lead to a decline in output. We find that output remains largely unaffected by noisy tax news for at least a year, but then starts to increase. We carry out an extensive robustness analysis to examine whether other factors are influencing the results and if our findings are sensitive to different model specifications. We find, however, that shape and magnitude of the output response to a change in noisy tax news are remarkably robust.[3]

The remainder of this paper is organized as follows. Section 2 illustrates the type of textual information we aim to quantify and gives an intuitive overview of our approach to do so. In Section 3 we outline the semi-supervised topic model in more detail and discuss how we estimate tax policy signals. Identified topics, including the tax prevalence measures, are presented in Section 4. In Section 5 we discuss how to construct a measure of noisy tax news from the estimated series on tax prevalence. We analyze the impact of noisy tax news on economic activity in Section 6. Finally, Section 7 provides concluding remarks and identifies some potentially interesting avenues for future research. Additional methodological details and further empirical results are collected in several appendices.

# 2 Quantifying tax policy signals

In the United States, the president is the main driving force behind tax policy legislation (Yang, 2007; Romer and Romer, 2010), making presidential speeches a highly informative source for signals about future tax reforms. Below we illustrate the information flow and the various types of signals we aim to quantify. All four statements were made by Ronald Reagan in the months leading up to enactment of the Economic Recovery Act of 1981.

Throughout his term, the president outlines tax reforms he wants to pursue, albeit in general terms:

> "It is time to reawaken this industrial giant, to get government back within its means, and to lighten our punitive tax burden."
>
> - *Inaugural Address* January 20, 1981

[3]The data and code to replicate our results can be found at https://doi.org/10.34894/JHVJQS.

When the president believes the tax law should be changed he recommends that to the House of Representatives:

> "At the same time, however, we cannot delay in implementing an economic program aimed at both reducing tax rates to stimulate productivity and reducing the growth in government spending to reduce unemployment and inflation. On February 18th, I will present in detail an economic program to Congress embodying the features I've just stated."
>
> - *Address to the Nation on the Economy*, February 05, 1981

In that announcement, as well as in the following months, the president may offer details about particular measures included in the upcoming tax law changes:

> "I shall ask for a 10-percent reduction across the board in personal income tax rates for each of the next 3 years. Proposals will also be submitted for accelerated depreciation allowances for business to provide necessary capital so as to create jobs."
>
> - *Address to the Nation on the Economy*, February 05, 1981

Once the bill is passed through the Congress, the president provides final remarks as he signs it and ends the legislative process:

> "These bills that I'm about to sign—not every page—this is the budget bill, and this is the tax program—but I think they represent a turnaround of almost a half a century of a course this country's been on and mark an end to the excessive growth in government bureaucracy, government spending, government taxing."
>
> - *Remarks on Signing the Economic Recovery Tax Act of 1981 and the Omnibus Budget Reconciliation Act of 1981*, August 13, 1981

Quantifying individual statements in terms of monetary value is difficult as they often lack sufficient details about the proposed reform. While the Congressional Budget Office (CBO) might publish an estimate of its tax revenue impact, this is done only in the last stages of the legislative process, once the law is drafted. Moreover, the president's statements are often subject to revisions: statements might be retracted,

announced changes might not come to fruition or they might include different measures than initially planned. The goal of our approach is to capture those signals as well.

Our solution is to instead quantify how prominently each direction of tax reforms (cuts vs. hikes) is featured on the political agenda of the president. We do that by measuring the proportion of presidential speeches in a given quarter that is devoted to tax cuts and tax hikes respectively.

The idea of measuring how often a particular issue is mentioned in a body of texts is not new. Most prominently, Baker et al. (2016) develop an Economic Policy Uncertainty (EPU) index by quantifying the proportion of news article referencing various types of uncertainty. The main challenge of such an approach is determining which texts, or which parts of a specific text, are relevant. The simplest approach is to check whether (or how many of) the words in a document belong to a pre-defined set of terms (a so-called lexicon). For example, Baker et al. (2016) use a rule-based extension of this approach, which consists of checking if a combination of certain terms appears in a given text document. Lexicon-based approaches are particularly useful in sentiment analysis, since readily available sentiment lexicons can often be used in a variety of applications. Shapiro et al. (2020) combine multiple lexicons to produce a robust analysis of economic news sentiment and its effects. In the case of our application however a lexicon-based approach would require that we exactly specify the words and phrases which the president uses to reference tax reforms. This introduces a risk of misspecification, especially if the relevant terms can appear in different contexts.

Topic modelling solves this problem by jointly considering the whole text instead of looking for particular terms. We can compare the terms used in a document with the term distributions used to discuss a certain topic to determine to what degree the text is about it. Crucially, in contrast to lexicon- or rule-based approaches, a topic model "learns" those distributions from the data in an unsupervised fashion. Given a collection of documents and a specified number of topics, we find topics which best explain the way the words are used in the texts. The method we chose is the Latent Dirichlet Allocation (LDA) model proposed by Blei et al. (2003) which is discussed further in the following section. It allows us to identify the way in which the president talks about tax policy changes, expressed as a probability distribution over the vocabulary, and determine which texts discuss them.

A similar approach has been used for example by Larsen and Thorsrud (2019). Their LDA-based analysis of news outlets shows the impact that various types of news have on financial markets. We are also not the first ones in the economic literature who use topic models to analyze the statements of policymakers. LDA has been widely used to analyze the economic impact of the communication by central banks. For instance Hansen and McMahon (2016); Hansen et al. (2018) analyze the minutes of the deliberations of the *Federal Open Market Committee* to identify topics of discussion and measure their impact on the economy. In a study similar to ours, Dybowski and Adämmer (2018) also use a LDA model to identify the part of presidential speeches devoted to taxation in general. Using a lexicon-based approach they further measure how optimistic or pessimistic the identified tax communication is and investigate whether the effect of the communicated tax changes on economic activity depends on the tone. Crucially, the effect they capture is by design that of changes in perceived uncertainty. Since those are only very loosely connected with the communicated tax changes, their approach is unfortunately not well suited for analyzing the effects of tax news. The above examples are part of a fast growing body of literature using text-mining in economic research; for a recent survey see Gentzkow et al. (2019).

Our approach is meant to explicitly distinguish between signals about tax hikes and tax cuts. Because standard LDA estimates the topics in an unsupervised fashion, there is little control over their composition. Intuitively, the discussion about both tax increases and decreases relies on a relatively similar, tax-oriented subset of vocabulary. As a result, when using the standard approach, the two are "grouped" together into a single topic, preventing us from determining the direction of the discussed changes. By reading through (some of) the speeches that we know are tax related we can gather precise information about terms which differentiate the two types of signals. Including this additional (prior) information in the model requires however a departure from the conventional LDA approach. In our two-step approach we combine the so obtained information with the results from the unsupervised approach to construct informed priors for the topics. By using those priors in a LDA model we are able to differentiate between the content devoted to discussing tax hikes and tax cuts. We aggregate the per-document results for each quarter to obtain a measure of the relative prominence of each political issue on the presidents agenda which we refer to as prevalence. In Section 5 we show that the prevalence measures of the two tax topics contain

information about future tax changes.

# 3 A topic model for measuring tax news

In this section we outline our methodology to extract information about tax policy from presidential statements using topic modelling. Intuitively, our model assumes that when the president discusses different topics, he uses a different distribution over words (his vocabulary). Identifying these distributions, and their occurrence in each statement, therefore allows us to identify the topics the president talks about. We hypothesize that one or more of these topics can be linked to instances when the president talks about tax policy. Indeed, as we show in Section 4, after estimating an unsupervised LDA model, one of the estimated topics can be labeled as the *tax* topic. However, as shown in Section 5.1, such a tax topic contains information too imprecise to be useful for predictive or structural analysis. We therefore aim to explicitly differentiate the discussion about tax policy into *tax increase* and *tax decrease* topics. For this purpose we introduce a two-step topic modelling approach. Before we discuss the topic model in detail, we first describe the data and the steps we take in pre-processing.

## 3.1 Text data and pre-processing

Our analysis is based on *the Public Papers of the Presidents*, a compilation of all documents originating from the president. We obtain the raw texts from the *American Presidency Project* (APP).[4] We analyze 59,214 texts spanning from 1949-01-20 to 2017-01-19.

The raw text of the speeches needs to be pre-processed to quantify relevant features of the text data, facilitating further statistical analysis. This process is described in detail in Appendix C.1. First, the documents in the dataset range from short remarks consisting of several sentences to long speeches, such as the *State of the Union Address*, which cover a variety of otherwise unrelated issues. We therefore split the texts into a total of 1,119,200 individual paragraphs of roughly the same length. For the remainder of the analysis we treat those as separate *documents*.

Next, we prepare a matrix of *word counts* which is used as the input to our

[4] www.presidency.ucsb.edu, retrieved on 2019-03-25.

algorithm. Informally, for each text we count how often every word is used in it. To define the possible 'words' we take two steps. First, we clean the text and exclude function words such as "a" or "and" and rare words. In this step we also transform words to their 'root' form, e.g. "taxes" and "tax" are both counted as "tax" and "implements" and "implemented" both count as "implement". In the second step we identify combinations of two subsequent words that occur frequently together (so-called *bigrams*), which are then also counted as a 'word'. This is important for our analysis as combinations of words such as "tax cut" may contain very different information than either of the words would contain separately. To avoid confusion with the actual words used in the speeches, henceforth we refer to each of the 'words' we count as *terms*, which then refers to either individual (pre-processed) words or the identified bigrams.

## 3.2 LDA topic model

In this subsection we present the Latent Dirichlet Allocation (LDA) topic model. For this purpose we first define the following elements of the pre-processed data:

- The **vocabulary** $\mathcal{V} = \{v_1, \ldots, v_V\}$ is the set of all **terms** $v_i$ that appear in the data, where $V = |\mathcal{V}|$ is the total number of terms in our dataset. The vocabulary consists of $V = 50,851$ individual words and bigrams.

- The **corpus** $\mathcal{W}$ is the collection of all **documents** $\boldsymbol{w}_d$, where $d = 1, \ldots, D$. In our analysis the corpus consists of $D = 1,119,200$ individual paragraphs.

- Each **document** $\boldsymbol{w}_d$ is defined as a vector of $N_d$ **tokens**, which are random variables each taking one value from the vocabulary $\mathcal{V}$.

$$\boldsymbol{w}_d = (w_{d,1}, \ldots, w_{d,N_d})^T, \qquad w_{d,n} \in \mathcal{V}, \quad n = 1, \ldots, N_d, \quad d = 1, \ldots, D.$$

  Intuitively, instead of thinking about a document as a continuous string of text, we consider it as a sequence of $N_d$ slots, for which a particular realization is chosen from the set of possible values $\mathcal{V}$ when the document is created.

Our approach is based on the LDA topic model proposed by Blei et al. (2003). LDA is a relatively straightforward, unsupervised approach that often leads to easily interpretable results, and thus has become one of the most popular choices for topic

modelling (Jelodar et al., 2019). In LDA the process of creating a document is modelled as a series of independent draws from a particular distribution over the terms in the vocabulary. The key idea behind topic modelling is that the particular distribution changes depending on the *topic*, i.e., what the text is about. For example, we expect the president to use certain terms with different frequency when talking about education versus for example military build-up. As such, with $K$ topics,[5] the distribution with which terms are used overall, is a mixture of the $K$ distributions per topic, which we label as the *topic-term distributions*. Furthermore, each document can consist of multiple topics. The proportions of each topic's occurrence in a single document is labelled the *mixing proportion* of that document, which is of key interest for our analysis, as informally it addresses how much each topic (such as tax) is discussed in the document. The mixing proportions therefore allow us to identify documents which relate predominantly to tax policy. It is crucial to allow documents to discuss multiple topics in varying proportions. For instance, the president will generally not discuss tax in isolation, but in conjunction with other topics such as the need to balance the budget, stimulating the economy or creating the funding for particular investments such as military expenditures in times of war.

Hence, the creation of a document – and the tokens inside that document – can be seen as a two-step process. First, the creator (in our case the president), decides on the proportions of each topic to be discussed in that document (the mixing proportion). Then, given that proportion, each token in the document is randomly assigned to one topic. Next, given this *topic assignment*, a term from the vocabulary is drawn from the distribution corresponding to the assigned topic. We now formalize this as follows.

- The $n$-th token in document $d$, $w_{d,n}$, has a **topic assignment** $z_{d,n} \in \{1, \ldots, K\}$. The vector of topic assignments for the $d$-th document is denoted as $\boldsymbol{z}_d = (z_{d,1}, \ldots, z_{d,N_d})$, and the collection for the whole corpus as $\mathcal{Z} = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_D\}$.

- Within one document, all tokens have the same probability of 'belonging' to a certain topic, specified by that **document's mixing proportions**. For document $\boldsymbol{w}_d$, those proportions are defined as a $K$-dimensional vector $\boldsymbol{\theta}_d =$

[5] In the model that follows, the number of topics $K$ is assumed to be known. Of course, in practice this has to be estimated. The particular choice of $K$ for our dataset is described later.

$(\theta_{d,1}, \ldots, \theta_{d,K})^T$ such that

$$\mathbb{P}(z_{d,n} = k|\boldsymbol{\theta}_d) = \theta_{d,k}, \quad n = 1, \ldots, N_d, \tag{1}$$

where $\theta_{d,k} \geq 0$ and $\sum_{k=1}^{K} \theta_{d,k} = 1$ for all $d = 1, \ldots, D$. We denote those proportions jointly as $\boldsymbol{\Theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_D)^T$, a matrix of size $D \times K$ where the $d$-th row is the vector of the $d$-th document.

- All tokens assigned to topic $k$ share the same **topic-term distribution**. This distribution is parametrized by a $V$-dimensional vector $\boldsymbol{\phi}_k = (\phi_{k,1}, \ldots, \phi_{k,V})^T$ such that

$$\mathbb{P}(w_{d,n} = v_i|z_{d,n} = k) = \phi_{k,i}, \quad n = 1, \ldots, N_d,\ d = 1, \ldots, D,\ i = 1, \ldots, V.$$

  where $\phi_{k,i} \geq 0$ and $\sum_{i=1}^{V} \phi_{k,i} = 1$ for all $k = 1, \ldots, K$. We denote those distributions jointly as $\boldsymbol{\Phi} = (\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_K)^T$, a matrix of size $K \times V$ where the $k$-th row is the vector of topic $k$.

We estimate the topic-term distributions and the mixing proportions in a Bayesian procedure by means of the posterior parameter distributions using Gibbs sampling. We denote them as $\hat{\boldsymbol{\Phi}}$ and $\hat{\boldsymbol{\Theta}}$ respectively. In order to do so, LDA puts Dirichlet priors on both the mixing proportions $\boldsymbol{\Theta}$ and the topics' term distributions $\boldsymbol{\Phi}$. Appendix B.1 provides an intuitive description of the properties of the Dirichlet distribution $\boldsymbol{\theta}_d \sim \text{Dir}(\boldsymbol{\alpha}), \quad d = 1, \ldots, D$ and the implications of its use on estimation. In particular, we assume that $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_D$ share the same Dirichlet distribution with $K$-dimensional parameter vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K)^T$. Following the recommendations of Wallach et al. (2009), we do not treat $\boldsymbol{\alpha}$ as a fixed hyper-parameter but place another level of (Gamma) priors on its elements, such that $\boldsymbol{\alpha}$ is estimated along with the lower-level parameters.

Dirichlet priors are also placed on the topic-term distributions. How we do that exactly depends on the first or second step of our model, and is the source of our methodological innovation in the second step. Generally, each topic-term distribution has a Dirichlet prior $\boldsymbol{\phi}_k \sim \text{Dir}(\boldsymbol{\eta}_k), \quad k = 1, \ldots, K$ with $V$-dimensional parameter vector $\boldsymbol{\eta}_k = (\eta_{k,1}, \ldots, \eta_{k,V})^T$. In the first step we use the conventional LDA setup where no 'expert knowledge' about topic composition is formulated. In particular, all topics share the same symmetric Dirichlet prior, such that $\eta_{k,i} = \eta$ for all $k = 1, \ldots, K$

and $i = 1, \ldots, V$ - the prior probability of every term in every topic is equal. This parameter $\eta$ is then also estimated by assuming a Gamma prior (Wallach et al., 2009). These uninformative priors imply that the topic-term distributions is fully driven by the data.

Crucially however, and as discussed before, a pure data-driven, unsupervised approach is unsuited to differentiate between *tax increase* and *tax decrease* topics, for which purpose we add a novel second step LDA estimation. Based on our first-step estimates $\hat{\mathbf{\Phi}}$, we first identify a *general tax* topic and use it to develop informed priors for the two topics of interest for the second step of LDA estimation. The exact process is described in the following section. The result is a set of parameter vectors $\boldsymbol{\eta}_k, \quad k = 1, \ldots, K$ which specify separate, informed, prior for each topic, where we postulate that particular terms occur either more frequently, or less frequently, in specific topics. These priors are then used in another Gibbs sampling to estimate the posterior means of the parameters, which are denoted as $\hat{\hat{\mathbf{\Phi}}}$ and $\hat{\hat{\mathbf{\Theta}}}$. The full details of the estimation for both steps are provided in Appendix B.3

The approach described above relies on the assumption that the number of topics $K$ is known, but in practice this number needs to be estimated. An incorrect specification of the number of topics decreases the overall quality of the model and the interpretability of the topics. While the choice of $K$ can be guided by metrics based on the likelihood of observing the corpus given the model, we base this choice on the interpretability of our results instead. Specifically, our approach requires that discussion about tax policy forms a distinct topic. Intuitively, specifying too few topics is likely to lead to a topic that includes also other issues, not necessarily related to tax policy. On the other hand, setting the number of topics too high introduces unnecessary computational complexity and can result in the tax topic being split, for example into personal and corporate tax policy discussion. As a result we choose the number of topics to be $K = 25$ in the first step. In the second step, because we split the *general tax* policy topic into the *tax increase* and *tax decrease* topics, we use a total of 26 topics. In section 4 we substantiate our choice by inspecting the estimated topic-term distributions and mixing proportions.

### 3.3 Two-step 'semi-supervised' prior construction

By using different priors for the topic-term distributions we can include 'expert knowledge' as a priori beliefs about topic composition to 'steer' the topics into the direction we want them to go, in our case topics about tax increase and tax decrease. Here we describe how we achieve this in our two-step procedure. The main challenge is now to correctly translate a priori beliefs about occurrence of particular terms into meaningful prior parameters $\boldsymbol{\eta}_i$ of the Dirichlet distribution. To illustrate, note that for any topic-term distribution $\boldsymbol{\phi}_k \sim \text{Dir}(\boldsymbol{\eta}_k)$ the Dirichlet prior implies that $\mathbb{E}[\phi_{k,i}|\boldsymbol{\eta}_k] = \frac{\eta_{k,i}}{\sum_{j=1}^{V} \eta_{k,j}}$. Hence, in order to obtain sensible priors we need to formulate beliefs about the relative occurrence of *all* terms. For example, we may believe that terms belonging to a certain set $L$, a so-called *lexicon*, are likely to be mostly associated with a particular topic of interest. Such a belief could be expressed by choosing a topic $k$ and setting $\eta_{k,i} = a\mu_L$ if $v_i \in L$ and $\eta_{k,i} = a\mu_{\neg L}$ if $v_i \notin L$ where $\mu_L > \mu_{\neg L}$. However, the shape of the resulting a priori distribution is not in line with the way terms occur in text in general, as empirical term distributions tend to roughly follow a power law (Sato and Nakagawa, 2010). Setting such a prior would therefore lead to a highly distorted topic distribution; instead we must make sure the prior parameters are in line with the empirical properties of the text.

To address this issue, we first learn about the general shape of the topic-term distributions from the data in the first-step unsupervised estimation and then *modify* these distributions using lexicons of terms relevant to the topics of interest. From the first-step unsupervised estimation we are able to identify a single topic that clearly encompasses all the discussion about tax changes, , which we refer to as the *general tax* topic.[6] We denote the corresponding estimated topic-term distribution as $\hat{\boldsymbol{\phi}}_{k^*}$.

We then construct the priors for the *tax increase* and *tax decrease* topics based on the assumption that both topics of interest have a similar distribution over the vocabulary, except for some key differentiating terms. Based on reading of documents related to tax policy we manually identify the terms which are predominantly used when discussing changes in a particular direction.[7] When then group them into two lexicons - $L_{\text{inc}}$ for terms related to tax increases, and $L_{\text{dec}}$ for terms related to tax

[6]This interpretation is based on the fact that this distribution assigns uniquely high probabilities to terms related to tax policy (e.g. "tax"), which is discussed in detail in Section 4.

[7]For this purpose we use all off the speeches identified by Romer and Romer (2009a) and Yang (2007) as announcements of tax changes.

decreases. The composition of the lexicons and details concerning their creation are presented in Appendix B.4.

We then 'guide' the algorithm towards the two tax topics by modifying the prior probabilities of terms which are contained in either of the lexicons. For the tax increase prior $\boldsymbol{\eta}_{\text{inc}}$, we modify $\hat{\boldsymbol{\phi}}_{k^*}$ by multiplying probabilities of terms in $L_{\text{inc}}$ by a constant $m_1 > 1$ to 'up-weight' them, and simultaneously multiplying probabilities of terms in $L_{\text{dec}}$ by $m_2 < 1$ to down-weight those. For the prior of the tax decrease topic we do the reverse. For the priors of the other topics we use their respective distributions estimated in the first step, without modifying their shape. This allows us to 'fix' the other topics while splitting up the *general tax* topic. The last step in the modification of those priors is to choose the strength of the effect of our prior, which is done by multiplying the vectors by a scalar $m_{3,k}$. Hence, we construct our prior parameters for all $i = 1, \ldots, V$ as

$$\begin{aligned}
\eta_{\text{inc},i} &= m_{3,\text{inc}} \left[ \hat{\phi}_{k^*,i} + (m_1 - 1)\hat{\phi}_{k^*,i} \mathbb{1}\left(v_i \in L_{\text{inc}}\right) + (m_2 - 1)\hat{\phi}_{k^*,i} \mathbb{1}\left(v_i \in L_{\text{dec}}\right) \right], \\
\eta_{\text{dec},i} &= m_{3,\text{dec}} \left[ \hat{\phi}_{k^*,i} + (m_1 - 1)\hat{\phi}_{k^*,i} \mathbb{1}\left(v_i \in L_{\text{dec}}\right) + (m_2 - 1)\hat{\phi}_{k^*,i} \mathbb{1}\left(v_i \in L_{\text{inc}}\right) \right], \\
\eta_{k,i} &= m_{3,k} \hat{\phi}_{k-1,i}, \qquad k = 3, \ldots, K+1
\end{aligned} \tag{2}$$

We set $m_1 = 100$ and $m_2 = 100^{-1}$, while $m_{\text{inc}}, m_{\text{dec}}, m_k$ are chosen such that the sum of the vectors is $\sum_i \eta_{\text{inc},i} = \sum_i \eta_{\text{dec},i} = \sum_i \eta_{k,i} = 10,000$.

This prior can be interpreted as postulating a prior belief equivalent to an additional observation of 10,000 tokens assigned to a given topic from the corresponding topic-term distribution; a relatively weak prior given the size of our dataset. In practice, the modified $\boldsymbol{\eta}_{\text{inc}}$ and $\boldsymbol{\eta}_{\text{dec}}$ act as 'seeds', during each iteration of the estimation 'nudging' the distributions of the relevant topics through the mechanism described in B.1. As a result, our approach is more robust to semantic misspecification than typical lexical approaches as their impact is relatively small compared to that of the observed dataset.

## 3.4 Constructing measure of tax policy signals

The final step of our text-analytic approach is to transform the estimation results from the topic model into numerical measures that reflect the prominence of tax-related discussions by the president. That is, we aim to construct a quantitative

measure of the prominence of signals about tax policy over time, which does not follow directly from our topic model. The estimated topic model allows us to estimate the mixing proportion $\hat{\hat{\boldsymbol{\theta}}}_d$ for each document $d = 1, \ldots, D$ as the posterior means of the second-step LDA. These estimates signify what proportion of its words is associated with a given topic: documents for which $\hat{\hat{\theta}}_{d,k}$ is estimated to be high discuss topic $k$ for a significant portion relative to documents for which $\hat{\hat{\theta}}_{d,k}$ is low. We now use this property to aggregate estimation results for individual documents to create a measure for the topics' *prevalence*, or popularity, and its evolution over documents registered over time. Because we are using our measure together with macroeconomic data in the empirical models discussed later, we opt for aggregating the documents to quarterly frequency.

Formally, let $T$ denote the total number of quarters, and let $T_d \in (0, T]$ denote the normalized date corresponding to the publication of document $d$. For the measure in quarter $t$, we then average over all documents published in the period $(t-1, t]$ to obtain the measure

$$\text{prevalence}_{t,k} = \frac{\sum_{d=1}^{D} \hat{\hat{\theta}}_{d,k} \mathbb{1}\left(t-1 < T_d \leq t\right)}{\sum_{d=1}^{D} \mathbb{1}\left(t-1 < T_d \leq t\right)}, \qquad t = 1, \ldots, T, \quad k = 1, \ldots, K.$$

To illustrate the proposed prevalence measure, consider two extreme cases:

- If all tokens in all documents in the time period belong to topic $k$, we have $\hat{\hat{\theta}}_{d,k} = 1$, $\hat{\hat{\theta}}_{d,l} = 0$ for $l \neq k$ for all documents $d$. In that case $\text{prevalence}_{t,k} = 1$, while $\text{prevalence}_{t,l} = 0$ for all other topics $l \neq k$.

- If all tokens in all documents are uninformative, so equally belonging to $k = 1, \ldots, K$ topics, that is $\hat{\hat{\theta}}_{d,k} = 1/K$ for all $k = 1, \ldots, K$, then the prevalence of all topics is $1/K$, implying that all topics are appearing equally likely for the time period considered.

This results in a measure indicating how prominently the president talks about tax increase and tax decrease over time. We stress that this measure should not be read as *tax news*. There is no reason why (i) the identified mentioning of tax cuts (or hikes) relates to *future* tax policy, and even if it does, (ii) it is interpreted as news by economic agents. We will address both issues explicitly in the remainder of the paper.

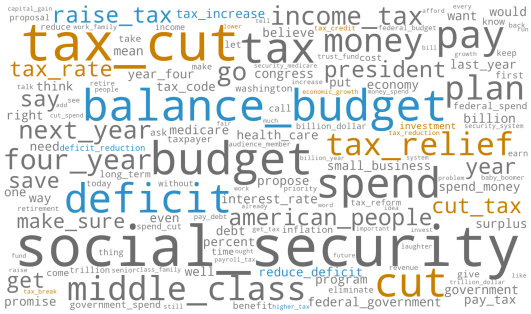
Arguably, the myriad of choices made in pre-processing, topic model specification and measure creation may make our final measures seem arbitrary. Indeed, while we made those decisions generally in accordance with standards used in the literature, alternative choices appear equally plausible to justify. Therefore, rather than arguing that our choices are the optimal ones, we empirically investigate if our resulting estimates have the properties we attribute to a measure of tax policy signals. In the next section we assess the topics estimated through our two-step LDA approach and evaluate the ability of our model to properly classify the tax content of the documents.

# 4 Identified policy topics

In this section we investigate in how far the fitted topic model captures our concepts of policy topics, with special attention to the tax (increase and decrease) topics. We first evaluate the two steps of constructing tax topics in detail. Next, we also briefly consider the other topics in order to understand how well the topic model captures the general essence of the speeches.

## 4.1 Tax topics

Our two-step LDA topic model relies on the identification of a single (general) tax topic in the first, unsupervised step. It turns out that in our analysis we find a clear tax topic; the most occurring terms in this topic are presented in Figure 1a. Based on this topic, combined with our lexicons, the second-stage guided LDA then produces the *tax increase* and *tax decrease* topics presented in Figures 1b and 1c, respectively. The terms in our lexicons driving the prior for tax increase (decrease) are highlighted in blue (orange). Although both distributions have many terms in common with the general tax topic estimated in the first step, there are crucial differences that extend beyond the terms whose prior probability was modified. Those include such terms as 'social security' which is used predominantly when discussing tax increases, and 'small business' which is referenced often when discussing tax decreases. Lastly, we can see that the topics still feature, to some extent, terms which we determined relate to tax changes in the other direction (e.g. 'cut tax' still appears in the 'tax increase' topic). This is an example of how the data overrides the priors. Apparently, even when discussing increasing taxes, the presidents sometimes make a reference to tax




(a) General tax topic (from step 1)


(b) Tax increase topic


(c) Tax decrease topic

Figure 1: Wordclouds of the estimated tax topics. The size of a term represents the probability attributed to it in topic's distribution. The terms included in the lexicon for tax increase (decrease) are highlighted in blue (orange).

cuts.

In both steps we verify that the identified tax topics capture all of the content devoted to tax policy changes. We do that by inspecting the probabilities assigned to tax-related terms by the non-tax topics. For example, in the first step we find that the topic devoted to state and local policies has the second highest probability of using the term "tax", which is still over 100 times smaller than the probability assigned to that term by the tax topics. Overall, in both steps the identified tax topics account for over 99% of the usage of the term "'tax".

For our approach, the estimated mixing proportions of each document are more important than the topic-term distributions. For our method to make sense, we specifically need to verify that the estimated mixing proportions for the two tax topics do indeed reflect our interpretations of the texts. Given the size of our dataset, with 1,119,200 individual paragraphs, such analysis is possible only for a subset of documents. We first consider the paragraphs from speeches that Romer and Romer (2010) and Yang (2007) list as announcements of tax policy changes. Figure 2 shows

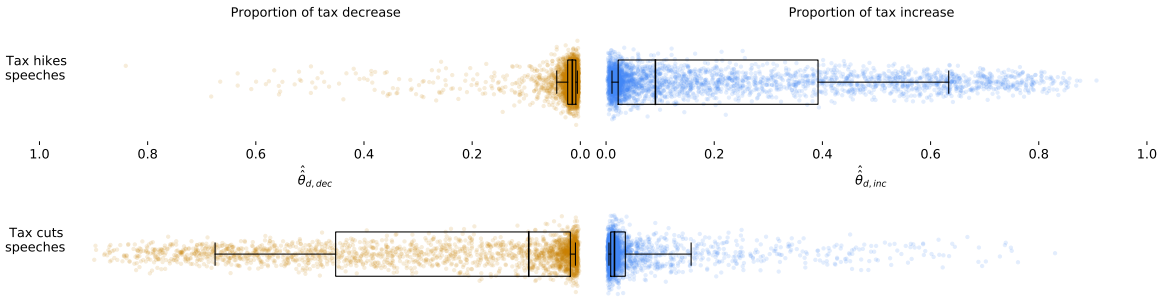



Figure 2: Mixing proportions of *tax increase* and *tax decrease* topics for paragraphs in speeches announcing tax changes. Upper (lower) row shows documents coming from speeches that announce tax hikes (cuts). Boxes show the 1st, 2nd and 3rd quartile. The whiskers show 1st and 9th decile

that, as expected, documents coming from speeches announcing tax hikes tend to have much higher mixing proportion for the *tax increase* topic, and the same is true for those announcing tax cuts and the *tax decrease* topic. In case of some documents we see that the opposite mixing proportion is high. This can be at least partially explained by the fact that some legislated changes include measures in both directions, lowering some taxes while raising others. Finally, in both groups a lot of documents do not relate to tax changes at all, since some of the speeches considered (e.g. *State of the Union Addresses*) concern more issues than just taxation.

Out of those documents we randomly select 100 for which either $\hat{\hat{\theta}}_{d,inc} > 0.3$ or $\hat{\hat{\theta}}_{d,dec} > 0.3$ for reading and compare the estimated mixing proportions with our interpretation of their content. Additionally, we consider the estimated maximum a posteriori (MAP) topic assignment labels for particular tokens.[8] While those labels are not directly necessary for our analysis, they help visualize the clustering property of LDA. In Tables 1 and 2 we present an example for each direction of change, where tokens assigned to tax increase (tax decrease) topic are highlighted in blue (orange).[9] We present details of this part of analysis and summarize our findings for the rest of the selected speeches in Appendix D.1. We find that overall the estimated mixing proportions correctly capture the direction of the implied changes.

Looking at the estimated topic assignments for individual tokens we can see the clustering property of LDA. In many cases terms that are not clearly related to tax

[8]For a token $w_{d,n} = v_i$ the maximum a posteriori (MAP) topic assignment estimate is $z^{\star}_{d,n} = \arg\max_k \hat{\hat{\theta}}_{d,k}\hat{\hat{\phi}}_{k,i}$

[9]Words not in bold are removed during pre-processing.

That's my **commitment**. That's what's at **stake in this election**. **Change** is a **future** where we have to **reduce our deficit**, but do it in a **balanced way**. And I've **signed** a **trillion dollars' worth** of **spending cuts**; I **intend** to do more. But if we're **serious** about **reducing the deficit**, we've got to **ask the wealthiest Americans** to **go back to the tax rates** they **paid** when **Bill Clinton** was in **office**. Because, **listen**, a **budget** is about **priorities**; it's about **values**. And I'm not **going** to **kick** some **kid** off of **Head Start** so I can **get** a **tax break**. I'm not **going** to **turn Medicare** into a **voucher** just to **pay** for **another millionaire**'s **tax cut**. That's not who we are.

**President Barack Obama, 2012-11-05.**
*Context*: One of President's Obama campaign speeches in which he advocated for an increase in taxation, in particular for the wealthiest Americans.
*Future impact*: Raised the rate and introduced new surtax on capital and investment gains for households with income over $250,000 as part of financing of Medicare.
$\hat{\hat{\theta}}_{d,\text{inc}} = 0.80, \quad \hat{\hat{\theta}}_{d,\text{dec}} = 0.03$

Table 1: Example of a "tax increase" speech

The **tax reductions** I am **recommending**, **together** with this **broad upturn** of the **economy** which has **taken place** in the **first half** of this **year**, will **move** us **strongly forward** **toward a goal** this **Nation** has not **reached since** 1956, 15 **years** ago: **prosperity** with **full employment** in **peacetime**.

**President Richard Nixon, 1971-08-15.**
*Context*: Address to the Nation Outlining a New Economic Policy: "The Challenge of Peace". President Nixon announces a package of measures meant to stimulate the economy by decreasing taxation.
*Future impact*: The package was voted in later that year.
$\hat{\hat{\theta}}_{d,\text{inc}} = 0.05, \hat{\hat{\theta}}_{d,\text{dec}} = 0.44$

Table 2: Example of a "tax decrease" speech

changes are classified as referring to one of the tax topics. This is the result of the context in which they appear - because so much of the rest of the speeches relate to tax increase or tax decrease the likelihood that they relate to one of those topics is higher. Even terms that intuitively refer to changes in one direction can be classified as referring to changes in the opposite direction (e.g. 'tax break' assigned to the tax increase topic in the 2012-11-05 speech). This property of the topic model is a crucial improvement over a simple lexicon-based approach, making it much less sensitive to misspecification.

It is important to note that despite the clustering property of LDA some paradoxical classifications can occur. This happens when the terms used in the text are not sufficient to capture the meaning contained in the syntax. For instance, speeches along the lines of "we are not going to increase taxes" would likely be classified as information about a future tax hike. This is inevitable given the *bag-of-words* assumption underlying the LDA topic model. Addressing this would require methods able to recover the true intention of statements and their interpretation in a broader context. However, in our context, this is difficult and often impossible even for well-informed political observers. Thus, it cannot be reasonably expected to be achieved perfectly by any algorithm, and we believe it should therefore also not be held against our approach based on the bag-of-words assumption. Moreover, such mistakes are likely to average out over all speeches within each quarter, such that our topic prevalence measures are still accurate in capturing the changes in the president's political agenda. A related issue is that it is likely that the president at times references to past tax reforms in his speeches. Although such sentences would presumably be classified to contain information about tax increase respectively tax decrease, they would not necessarily carry information about future policy plans. We will address this point in Section 5 when constructing our measure of tax news. Overall, these issues thus do not contradict our working hypothesis: if the president repeatedly emphasises the importance of future tax changes, the public likely expects that these changes will be implemented in the near future.

Our final diagnostic check in this section concerns the aggregated quarterly prevalence measures of tax increase and decrease. These are presented in Figure 3, along with periods of legislative lag of tax hikes (blue) and cuts (orange) as identified by Yang (2007). Visual inspection shows that the topic prevalence of the relevant direction - but not the opposite one - generally increases leading up to an actual tax

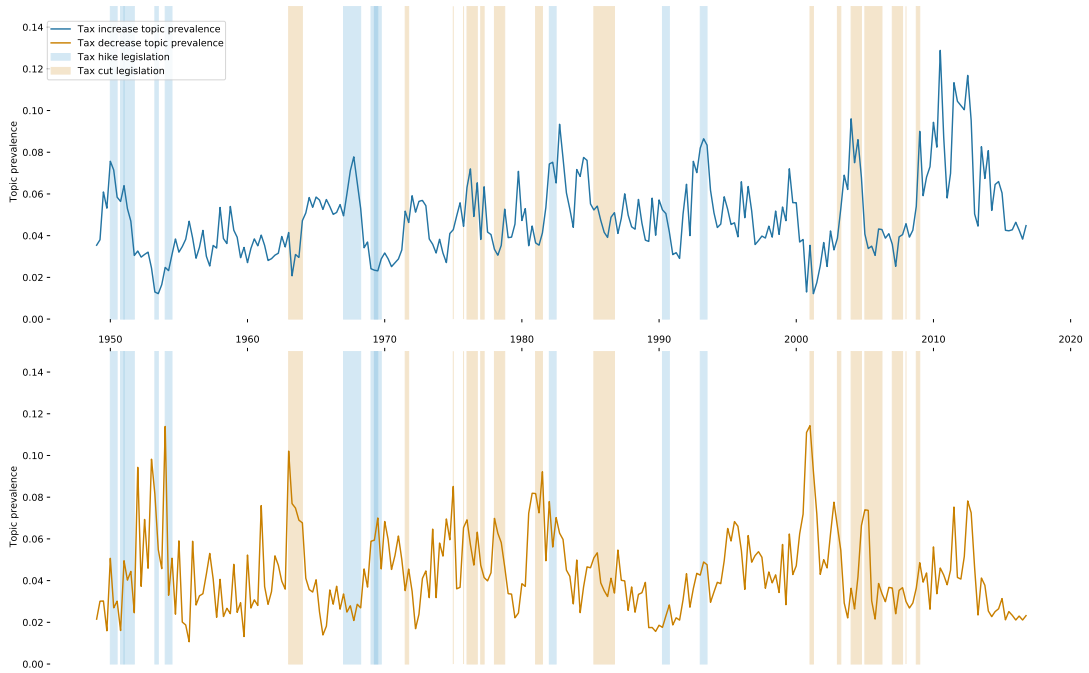



Figure 3: The prevalence of the *tax increase* (upper) and *tax decrease* (lower) topics. The shaded areas in blue (orange) indicate periods of legislative lag of tax hikes (cuts) as identified by Yang (2007).

change in that same direction, with the peak around the enactment date. While by no means a formal analysis, this does seem to indicate that our identified tax increase and decrease topics have predictive power for actual tax changes. We investigate this more formally in the next section, but first we briefly investigate the other topics.

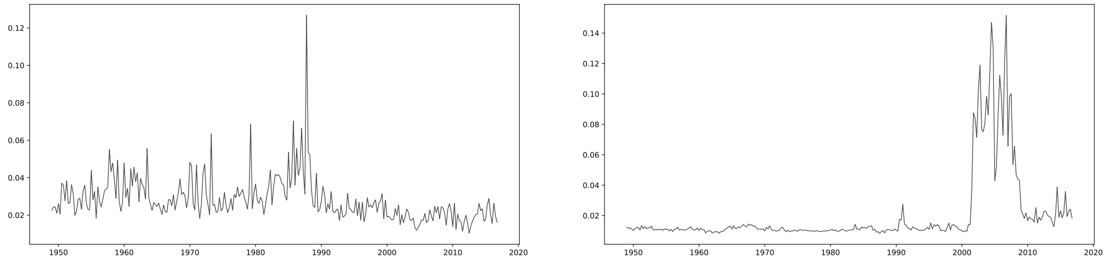
## 4.2 Other identified topics

In addition to the two tax topics, we identify 24 other topics. While those topics are not the focus of our study, they can give an indication of the overall quality of the model. Their distributions remains relatively unchanged between the first and the second step of the topic model The majority can clearly be attributed to specific policy issues such as public health, trade, and foreign policy. Figure 4 shows the distribution of a few selected topics. The distributions of all of the topics are presented in Appendix D.2.

In several cases we can quite intuitively see how the changing importance of certain issues is reflected in our prevalence measures. Perhaps the best examples are the *Cold*

(a) War on Terror (b) Cold War

Figure 4: Wordclouds of selected topics. The size of a term represents the probability attributed to it in topic's distribution.



(a) Cold War (b) War on Terror

Figure 5: Prevalences of other topics

*War* and *War on Terror* topics displayed in Figure 5. The prevalence of the *Cold War* topic has oscillated for decades, reaching a sharp peak in the late 1980s after which it declined considerably. In case of the *War on Terror* topic we see a slight peak in the early 1990s - likely the discussion surrounding the First Gulf War, a steep rise after the 9/11 attacks and a considerable decline since Barack Obama entered office, reflecting a different approach of the new president. This also shows that our LDA approach is flexible enough to handle variation over time.

## 5 Constructing a noisy tax news measure

When constructing the tax prevalence measures, our reasoning is that the obtained time series on tax decrease and increase topics should contain information on future tax changes. It is however very much possible that a tax reform put forward by the administration never comes to fruition, or its design has changed fundamentally by

the time it is signed into law. It is also possible, and contradicting our hypothesis, that the presidents strategically emphasise changes in one direction, or that they present them inaccurately.[10] Thus, it is likely that our tax prevalence measures contain both news *and* noise and a logical next step is to assess the informational content of our tax topic series. This is investigated in Section 5.1.

One should note that even 'noise' about future tax changes – signals that do not result in an actual implementation – can have macroeconomic effects if economic agents deem this signal to be credible and act upon it. Our goal is therefore not necessarily to find a series that best predicts actual tax changes, although obviously actual tax changes do need to be predicted for our measures to have meaning. In the structural analysis of Section 6 we investigate this further. However in order to do so, we need to transform them into a measure of noisy tax news for use in a structural analysis. This is the topic of Section 5.2.

## 5.1 Tax prevalence predictability

In this section we investigate how well the tax prevalence measures predict future tax changes. We initially estimate predictive regressions of various tax change series on our tax prevalence measures, assessing the strength of the correlation between tax changes and the lags of our measures by the $F$-test procedure suggested by Stock and Yogo (2005). This analysis shows that while the prevalence measure obtained from the first-stage unsupervised LDA (i.e. the general tax topic) is not a powerful predictor for any tax change measure considered, the second-step tax topics are strong predictors of future tax changes, with $F$-statistics well above the cut-off. for all considered measures of tax changes and across forecast horizons. As such, the tax topics series can be considered a strong predictor of future tax changes and could also be used as strong instruments for actual tax changes. The details of this analysis can be found in Appendix A.1.

To study (non-)predictive relationships in more detail but also to investigate to what extent our tax topics are driven by other factors such as macroeconomic conditions, we proceed with Granger causality tests. Not only are we interested in whether

[10]For example, as a result of tax cuts included in the Economic Recovery Tax Act of 1981 an immediate increase in revenue was deemed necessary to reduce the deficit. The following year the Tax Equity and Fiscal Responsibility Act was enacted. Even though the change resulted in a substantial increase in tax revenue, president Reagan's rhetoric pro-cuts remained largely unchanged during this period. Instead, the changes were presented as ways to close tax loopholes.

our prevalence measures have (joint) predictive power on various tax measures, we also investigate Granger causality the other way around, as this provides us with some understanding about the endogeneity of the tax prevalence series, which may pick up information about future government spending or other public polices affecting the federal budget.

To test for Granger causality we consider a large VAR where we include a large number of possibly relevant macroeconomic and financial variables. We also include variables that incorporate information on future policies concerning spending (Ramey, 2011; Ramey and Zubairy, 2018) and taxation (Leeper et al., 2012), as well as prevalence measures for the other topics. This allows us to determine in how far our tax prevalence series contain unique predictive power for the various tax change measures that is not found in other series. Similarly, by testing for Granger causality from all the series mentioned above (including tax change measures) to the two prevalence series we can determine what the causes of potential endogeneity are. To handle the high-dimensionality of the VAR needed for the tests, we use the post-selection test of Hecq et al. (2019) which provides valid inference about Granger causality after selection of relevant covariates via the lasso.[11]

The left column of Figure 6 illustrates which measures of tax changes are Granger-caused by our prevalence measures. It also shows whether any tax measure is predicting our tax prevalence measures. The tax prevalence topics Granger-cause all tax change measures. While predictive power is higher for aggregate measures of tax changes (federal revenue, or narratives of legislated tax changes), the tax topics also Granger-cause corporate, income, and payroll taxes. Most notably, we find that even the most exogenous RR news narrative, as well as Mertens and Ravn's (2014) unanticipated tax narrative, are Granger-caused by the tax prevalence measures. We further find evidence (p-value = 0.075) that our prevalence measures Granger-cause implicit tax rates, which are considered as proxy for tax news.[12] In contrast to these results, we do not find that any of the tax measures contains information to predict the tax prevalence measures.

We next investigate whether the prevalence measures predict macroeconomic and financial variables and vice versa. Results displayed in the right column of Figure 6

[11] All tests are based on a VAR with six lags. Nonstationary macro variables are transformed to first differences of logs.

[12] We use the risk-adjusted implicit tax rate from Leeper et al. (2012). We use rates with maturity of one year which are considered by the authors to best predict tax changes in the near future.

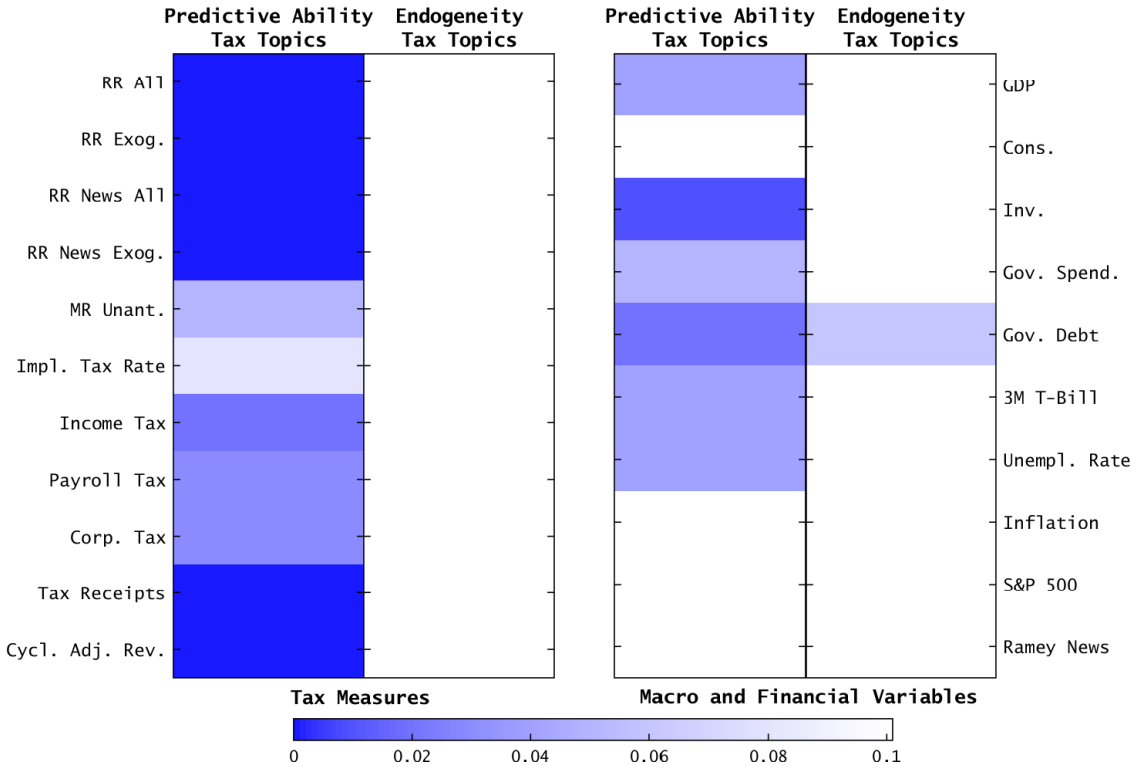



Figure 6: $p$-Values from Granger causality tests. Tax measures as well as macroeconomic and financial variables.

show that the tax prevalence topics Granger-cause the information typically included in a fiscal VAR. We find evidence for predictability in the other direction only for government debt. This suggests that the prevalence measures capture only information about tax changes that are not systematically correlated with short-run changes in economic activity, but are rather driven by other (long-run) policy goals, for example reducing the debt burden.

We also evaluate predictive relationships between the tax topics and other prevalence measures (results are shown in Figure A.1 in the Appendix). We find that some topics related to policies that affect the federal budget Granger-cause the tax topics. In particular, the *War on Terror* topic Granger-causes the tax prevalence measures. We will use this insight when selecting control variables among the topics in the structural analysis in Section 6 to eliminate possible endogeneity issues.

## 5.2 From prevalence measures to noisy tax news

In this section we construct a single measure of noisy tax news from the two tax prevalence series. For this, we need to take two things into account. First, the president may at times reference to past tax changes in his speeches. This does not add any new information about future tax changes and can therefore not be considered as news. Second, our tax prevalence measures likely contain information regarding tax reforms which were initiated or discussed by the administration but were later rejected or substantially modified throughout the legislative process. Ex-post, such statements carry no (explicit) information about future tax changes, yet they may have shaped expectations of economic agents at the time the respective message was conveyed to the public. Therefore, such "noise" should be included in our measure. Note that here we do not make any claim about how credible economic agents deem the president's speeches. This is the topic of the structural analysis in Section 6. Our goal here is simply to capture the signals, such that their effects can be analyzed in a structural analysis later.

Our tax related prevalence measures do not have a direct, economically meaningful, i.e. monetary, quantitative interpretation. If, however, we were able to transform our prevalence measures into a series of (noisy) tax news that conveys a monetary value of expected tax changes, this would simplify a structural analysis.

To tackle these issues, we proceed as follows. We first project out all information in the tax prevalence measures which is not orthogonal to all recently implemented or enacted tax changes as measured by RR. We do the same with RR's exogenous tax news series. Second, we regress RR's filtered exogenous tax news series on the filtered prevalence measures and interpret the predicted value of that regression as noisy tax news.

By predicting the RR tax news narrative we directly connect our tax topic series to actual implemented policy decisions. Moreover, we linearly combine the tax topics into a single series expressing tax news in monetary values. RR express tax changes in terms of the revenue effect (as a percentage of nominal GDP) in the quarter the change occurred. We choose to work with RR's exogenous tax news since this references the earliest point in time when economic agents know with certainty when a specific tax reform will be implemented and how the change will look like. For their tax news series RR discount the monetary value back to the quarter of enactment.

Let $\mathbf{Prev}_t = [\text{Prevalence}_{\text{inc},t}, \text{Prevalence}_{\text{dec},t}]'$ are our prevalence measures of tax

increase and tax decrease, respectively, and let $\Delta \boldsymbol{T}_{\text{all},t} = [\Delta T^{PV}_{\text{all},t}, \Delta T^{\text{Impl}}_{\text{all},t}]'$, where $\Delta T^{PV}_{\text{all},t}$ and $\Delta T^{\text{Impl}}_{\text{all},t}$ denote all tax changes dated at the time of passage and time of implementation, respectively. Then we estimate the regressions

$$\mathbf{Prev}_t = \sum_{j=1}^{6} \boldsymbol{\Gamma}_j \Delta \boldsymbol{T}_{\text{all},t-j} + \mathbf{Prev}_t^{\perp}, \tag{3}$$

and

$$\Delta T^{PV}_{\text{exo},t} = \sum_{j=1}^{6} \boldsymbol{\gamma}_j' \Delta \boldsymbol{T}_{\text{all},t-j} + \Delta \text{News}^{\perp}_{\text{exo},t} \quad \text{if } \Delta T^{PV}_{\text{exo},t} \neq 0, \tag{4}$$

where $\Delta T^{PV}_{\text{exo},t}$ is the present value of tax changes classified as exogenous by RR at time of passing of the bill.[13] The residuals $\widehat{\mathbf{Prev}}_t^{\perp}$ and $\Delta \widehat{\text{News}}^{\perp}_{\text{exo},t}$, contain now only information orthogonal to past tax policies.

To obtain our final measure of (noisy) tax news, we regress $\Delta \widehat{\text{News}}^{\perp}_{\text{exo},t}$ on past observations of $\widehat{\mathbf{Prev}}_t^{\perp}$ as well as on lags of all enacted and implemented tax changes from the RR narratives. That is, for $h > 0$ we estimate the following regression:

$$\Delta \widehat{\text{News}}^{\perp}_{\text{exo},t+h} = \sum_{j=0}^{p_1} \boldsymbol{\beta}_{j,h}' \widehat{\mathbf{Prev}}_{t-j}^{\perp} + \sum_{j=0}^{p_2} \boldsymbol{\gamma}_{j,h}' \Delta \boldsymbol{T}_{\text{all},t-j} + u_{t,h}, \quad \text{if } \Delta T^{PV}_{\text{exo},t+h} \neq 0. \tag{5}$$

Note that in the regression equations (4) and (5) we interact the regressors with time periods $t$ (resp. $t+h$) where (exogenous) tax reforms have actually been enacted, i.e. where the regressand is non-zero. While one could estimate a zero-inflated regression model on the full series to accurately model the exact zeros on the left-hand side of (5), we take the simpler approach of focusing only on the non-zero periods. Note that the fact that no tax change is enacted over a specific time period, does not mean that the president's speeches do not shape economic agents' beliefs regarding future tax reforms during that time. Thus, a tobit-type model is rendered inappropriate if we aim at capturing noisy tax news. The president may, for example, talk about a planned tax reform which never comes to fruition, or might change dramatically in

[13] We focus on RR's exogenous tax news, i.e. news only about tax changes uncorrelated with recent economic conditions, to avoid biased estimates. However, we find that using RR's exogenous *and* endogenous tax news (that means replacing $\Delta T^{PV}_{\text{exo},t}$ in (4) by $\Delta T^{PV}_{\text{all},t}$) lead to similar results when controlling for output movements and other macroeconomic policies. Figure A.5 in Appendix A.2 compares responses of output to the two differently constructed versions of noisy tax news.

its design when finally signed into law. This is not reflected in RR's sparse tax news narrative and does not affect tax rates ex-post, but such information is, however, likely to shape agents' expectations and therefore influences economic outcomes.

As such, we want our measure to approximate not only news about *observed tax changes* but news (and noise) that may affect *latent expectations about future tax changes*. One might consider to use a method like that of Heckman (1979) to account for selection bias arising from dependence between the probability of an enacted tax reform and the right-hand side variables in our regression. However, by doing so we would be modelling the credibility of tax announcements *ex-post*, that is, the likelihood of them leading to implemented changes. This may not be representative of how economic agents shape their expectations *ex-ante*. Instead, our approach is built on the idea that RR's identified tax changes at the time of enactment are a good *proxy* for (the realization of) news driving latent expectations. As it does not seem plausible to assume that during quarters in which no tax bill is passed (and RR's narrative is zero) no tax news arrives and expectations are not changing, we consider RR's tax news series as an unsuitable *proxy* during those periods and only consider those periods where tax changes occurred.

Then, for $h > 0$, we interpret

$$\text{News}_t^h = \sum_{j=0}^{p_1} \hat{\boldsymbol{\beta}}'_{j,h} \widehat{\mathbf{Prev}}_{t-j}^{\perp}$$

as (noisy) news about tax policies expected to be enacted $h$ periods in the future. By predicting the present value of exogenous tax changes $h$ periods ahead, we aim at capturing both the perceived probability of a tax change and the present monetary value of this possible policy action. We thus capture tax news containing information which seems relevant ex-ante even though the implied policy plans never come to fruition ex-post.

The many zero observations effectively limit the number of parameters we can estimate. We set $p_1 = 3$. That is, we consider tax related statements throughout the last year to be news relevant. To make sure to capture information regarding future tax changes only, we control for tax changes enacted during the last 1.5 years, i.e we set $p_2 = 5$.

Figure 7 shows the estimated noisy tax news series for $h = 1$ together with RR's tax news narrative including both the endogenous and the exogenous component. We

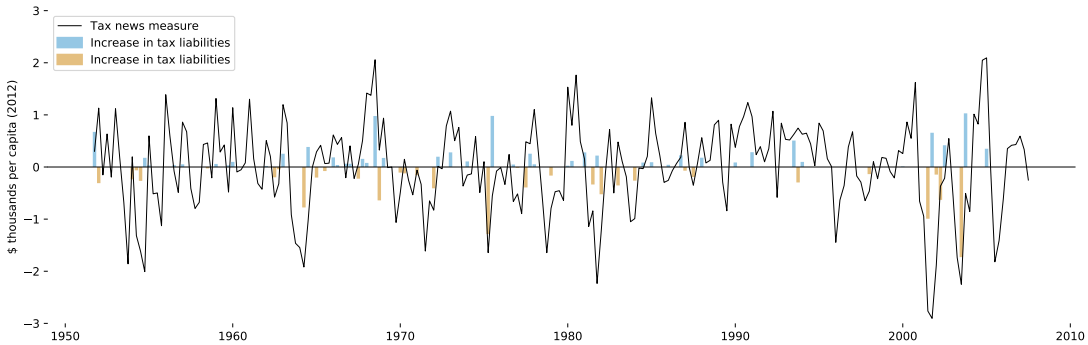



Figure 7: Our noisy tax news measure for $h = 1$ and the implemented changes in tax liabilities (Romer and Romer, 2010).

find, not surprisingly, that the noisy tax news series is a strong predictor of future tax changes regardless of how these are measured. Table A.4 shows $F$-statistics based on the regression in (A.1) with $z_{t-j}$ replaced by the estimated tax news series. We show the maximum $F$-statistic over $j = 0, 1, 2, ...6$ in (A.1) for different values of $h$ in (5). In particular, we find that our news series predicts implicit tax rates, challenging the validity of the latter series as suitable proxy for tax news.[14]

It is important to stress again that we do not identify separate noise or news shocks with our approach. The series we have labeled *noisy tax news* does not allow for a subsequent isolated investigation of the effect of either news about future tax shocks nor about the contribution of purely belief driven economic fluctuations. The information we hope to capture relates to *both* news about future fundamentals as well as changes influencing agents' beliefs only. One might argue that not being able to identify either component limits a (deep) structural interpretation. This criticism holds, however, for a variety of empirical papers. The previous investigation on the informational content of the tax prevalence measures indicates that the news signal about future tax changes, although not noise free, is strong enough allowing for a meaningful interpretation. Generally, noise and news are tightly related, further distinguishing both components and identifying them separately – if this is at all possible – is beyond the scope of this paper. We refer to Chahrour and Jurado (2018) for a detailed analysis of this relationship.

[14] An additional (yet different) concern when using implicit tax rates as a proxy for tax news is its comparably weak predictive power for future tax changes, see Table A.3.

# 6 The effects of noisy tax news

In this section we estimate structural models with our noisy tax news series as main variable of interest. First, in Section 6.1 we estimate the response of output to noisy tax news. Second, in Section 6.2, we study the implications of our series for the interpretation of structural analyses done with other narrative approaches to investigate the effects of (supposedly) exogenous tax shocks.

## 6.1 Output response to noisy tax news

We estimate output responses to a change in $\text{News}_t^h$ using local projections based on the following regressions

$$y_{t+H} = \alpha_H + \beta_H \text{News}_t^h + \boldsymbol{\gamma}_H' \boldsymbol{z}_t + \epsilon_{t,H}, \quad \text{for} \quad H = 0, 1, 2, ... \tag{6}$$

$y_{t+H}$ is real GDP in logs of levels. $\alpha_H$ includes deterministic components. If not indicated otherwise we include an intercept as well as a linear and a quadratic trend. $\boldsymbol{z}_t$ are control variables. For our benchmark specification, $\boldsymbol{z}_t$ includes lagged values of $\text{News}_t^h$, log GDP, log government spending, log government debt, and the 3-month Treasury Bill. Moreover, we include lags of two prevalence measures related to politics that are likely to affect the federal budget and which Granger-cause the tax topics (*War on Terror* and *Natural Resources, Energy & Technology*). Finally, we include lagged values of RR's exogenous tax changes (dated at time of implementation). Every variable in $\boldsymbol{z}_t$ is included with 12 lags. We set $h = 1$ because the resulting tax news estimate has the highest predictive power on tax changes, see Table A.4.[15] While including 12 lags of the dependent variable is a rather conservative choice and possibly more than necessary to capture the temporal dependency in $y_t$, it is in line with the suggestions in Montiel-Olea and Plagborg-Møller (2020) for the construction of robust inference. We construct HAR confidence intervals based on the method suggested in Lazarus et al. (2018). For all impulse responses we report 90% confidence intervals.

Figure 8 presents the results obtained from estimating (6) for various sets of control variables. For any specification, noisy news about a tax cut triggers a delayed but significant increase in output. Across all specifications the responses are quantitatively similar. While one should be careful with interpreting a change in $\text{News}_t^1$

[15] We find that altering the horizon has only little quantitative impact on output responses.

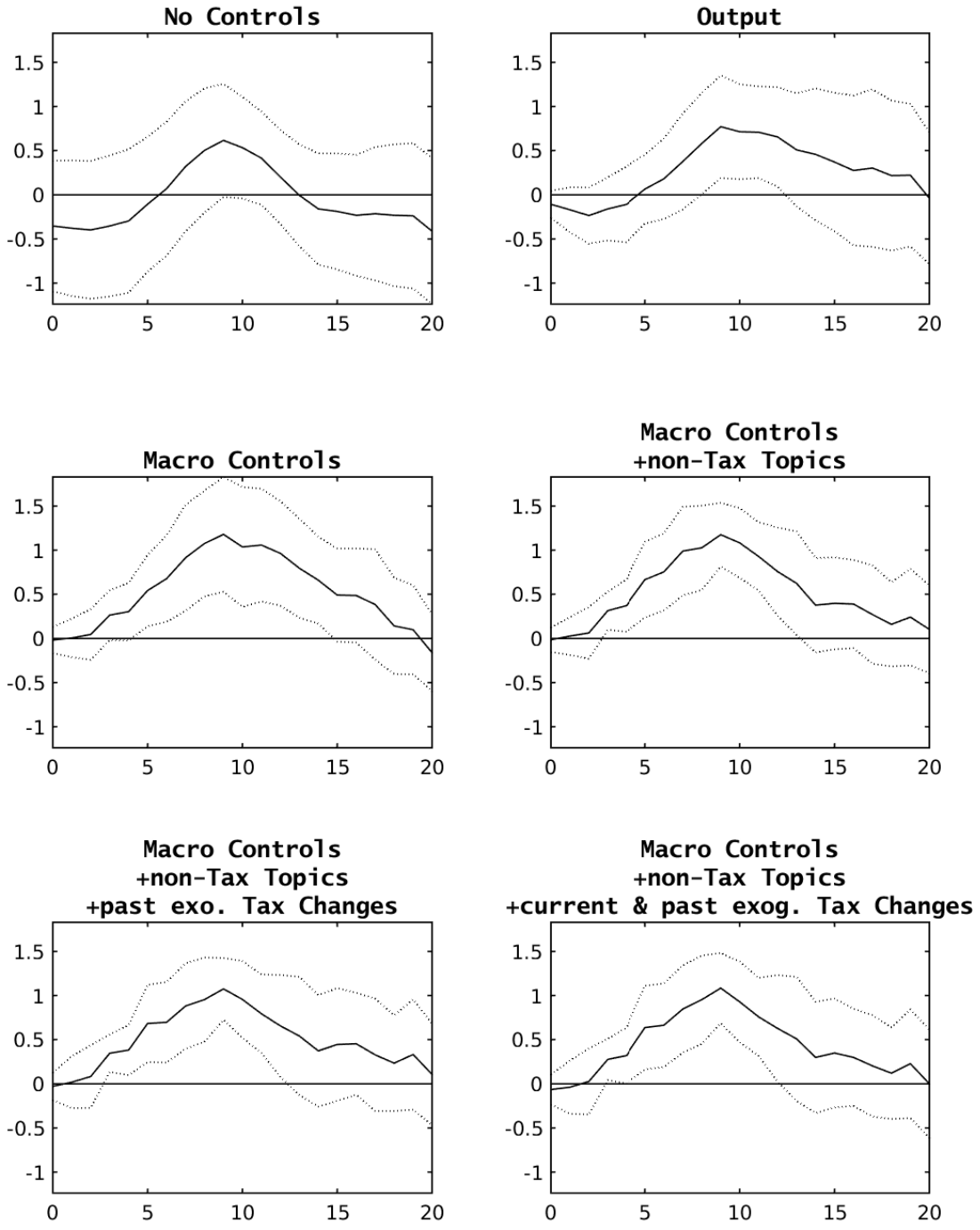



Figure 8: Output responses to a change in $\text{News}_t^1$ of the size of minus one percent of GDP. Varying set of control variables.

as pure anticipated tax shock due to the noise component, the triggered output response resembles in shape and maximum impact the one due to an anticipated tax cut estimated in Mertens and Ravn (2012). A crucial difference is, however, that we do not find evidence for contractionary effects due to future tax cuts during a possible anticipation period; the reaction of output is initially small and insignificant. Only when no additional control variables are included, we find an initial, though not significant, decline in aggregate economic activity.[16]

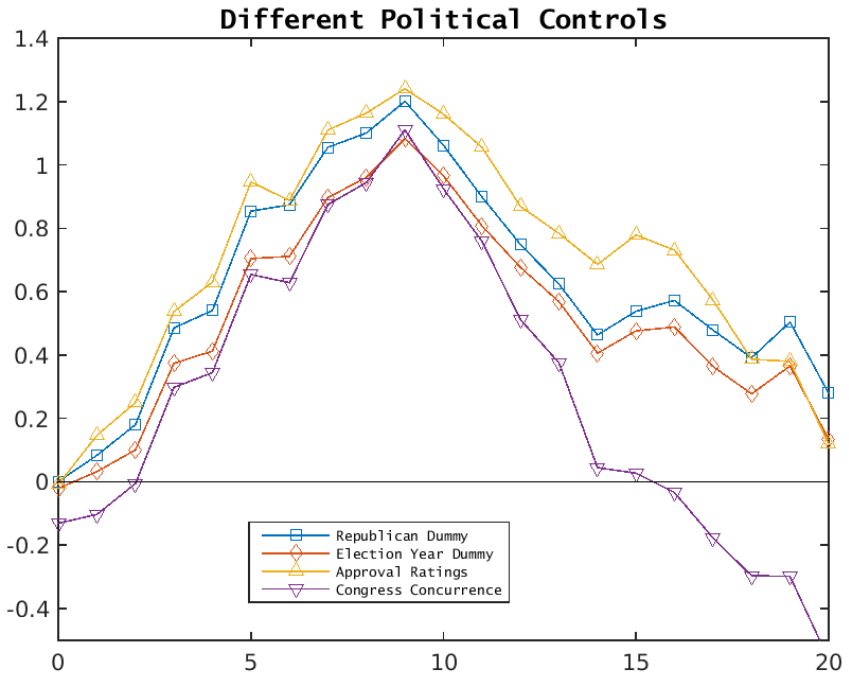



Figure 9: Output responses to a change in $\text{News}_t^1$ of the size of minus one percent of GDP. We estimate the LP regressions using the benchmark model specifications and the benchmark set of controls. Additionally we rotate in different political control variables.

It is plausible that political consideration drive the president's communication. One could imagine that a Republican president would conduct economic policies differently and people would expect tax cuts from him rather than from a Democratic president (or would deem initial announcements more credible). Similarly, an upcoming presidential election may also distort our results in case the president is

[16]This result is extremely robust across a variety of model specifications, see Figure A.2 in the Appendix.

campaigning for re-election. Moreover, a president with low approval ratings may be inclined to propose policy reforms that may increase his popularity. And finally, the president's relation with congress may influence the tone in his public communications but also the kind of policy reforms he brings forward. None of this seem to matter. Output responses to a change in $\text{News}_t^h$ when taking into account different additional political controls are displayed in Figure 9. To control for the president's popularity in public, we consider (quarterly averages of) Gallup's job approval ratings. The House and Senate concurrence is the percentage of members of congress who agree with the president's position on a roll call vote.[17] We include current observations and 12 lags when considering these additional controls in the local projection regressions.

To get an idea through which channel aggregate activity is affected, we investigate how other macroeconomic aggregates react to a change in $\text{News}_t^1$. We do this by estimating the regression in (6) for different $y_{t+H}$ using the benchmark specification for $\alpha_H$ and $\boldsymbol{z}_t$. Figure 10 shows estimated responses for real private consumption, real (non-residential) investment, real government spending, real debt, unemployment, and inflation. The two main aggregates of GDP, consumption and investment, mimic the shape of the output response for the benchmark model in Figure A.2. The output response seems to be particularly triggered by the strong reaction of investment. Government spending does not respond significantly for the first three to four years and then slightly increases. The federal debt level increases in the short-run but overall is not affected significantly. Unemployment initially rises but declines afterwards, co-moving with output. Noisy news about a tax cut is persistently disinflationary albeit only significant for the first six quarters.

## 6.2 Implications for estimating tax shocks

Often narrative accounts of tax changes are used to identify the effects of tax shocks. Various studies find that indicators of economic activity, such as output, investment, or employment react significantly in the first two or three years after tax-cuts occur. A precondition for the successful identification of tax shocks is the exogeneity of the tax narrative to current and past economic conditions. As seen in Section 5, our tax prevalence measures have (strong) predictive power on several measures of federal revenue changes or implemented tax reforms, as well as on RR's and Mertens and

[17] Both series are obtained from UC Santa-Barbara's American Presidency Project https://www.presidency.ucsb.edu/.

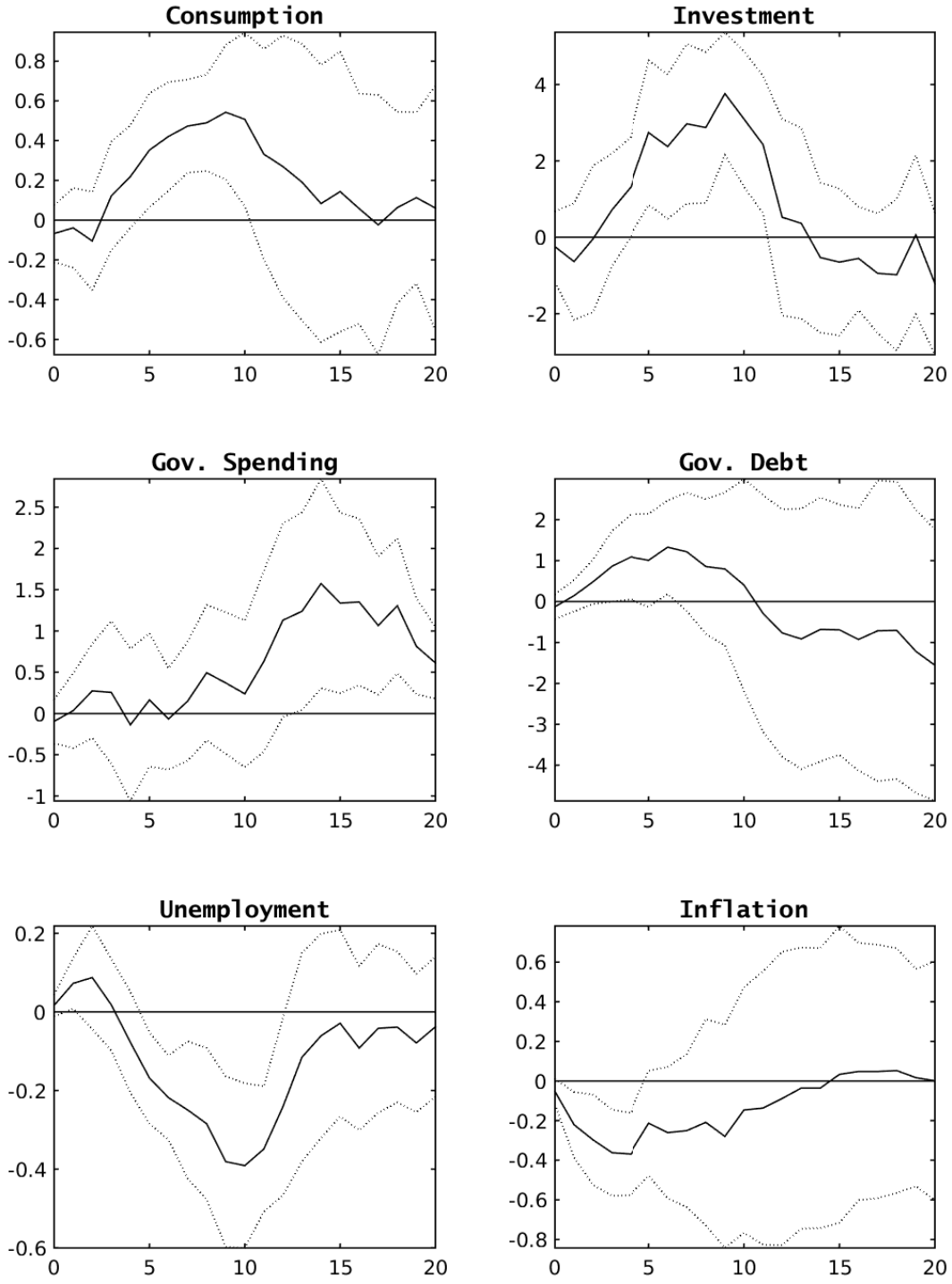



Figure 10: Responses of Components of GDP to a change in $\text{News}_t^1$ of the size of minus one percent of GDP. Benchmark model specification.

Ravn's (2014), presumed exogenous, narratives. Therefore, it stands to reason that part of the information in the narratives is erroneously used to identify exogenous tax changes and estimates could be biased.

To investigate this possible bias, we estimate output responses to changes in the original RR exogenous tax change variable and Mertens and Ravn (2014) unanticipated tax shock proxy respectively, however, conditioning on lagged observations of the tax prevalence topics.[18] That is, we estimate (6) replacing the regressor $\text{News}_t^h$ by either RR's or Mertens and Ravn's (2014) tax proxy.[19] $\boldsymbol{z}_t$ includes lagged values of the tax topic prevalence measures, lagged values of the tax shock proxy and lagged values of log GDP, log government spending, and interest rates. We include an intercept, a linear and a quadratic trend. 12 lags of all regressors are included.

Figure 11 displays output responses to a tax-cut shock identified using RR and Mertens and Ravn's (2014) exogenous tax change series. Each row compares the same estimation framework: once including our tax topic prevalence measures and once not. Regardless of the narrative/external instrument used for identification we see that including the tax topic series reduces the reaction of output. This finding suggests that part of the boost attributed to a tax shock is due to anticipation effects. As documented in the Appendix, impulse responses from VARs lead to the same conclusions.

One may expect that the tax topic series becomes less informative about tax changes in the far future (see Table A.2). If the tax topic measures are most informative about more recent tax changes, excluding more recent observations may alleviate the effect of anticipation. Figure 12 displays impulse responses of models where most recent $n$-quarters of the tax topic series are sequentially excluded from the set of regressors. These results confirm our hypothesis: if less and less recent information is included, anticipation does play a smaller and smaller role.

# 7 Conclusion

In this paper we analyze the public communications of U.S. presidents to identify the information relevant to changes in tax policy. Our semi-supervised topic modelling

[18] Mertens and Ravn (2013, 2014) investigate this bias by adding implicit tax rates to the set control variables, finding no difference in the effect of tax shocks.

[19] We ignore the fact that both proxies are only weakly correlated with (cyclically adjusted) revenue changes and do not consider weak-instrument robust inference here.

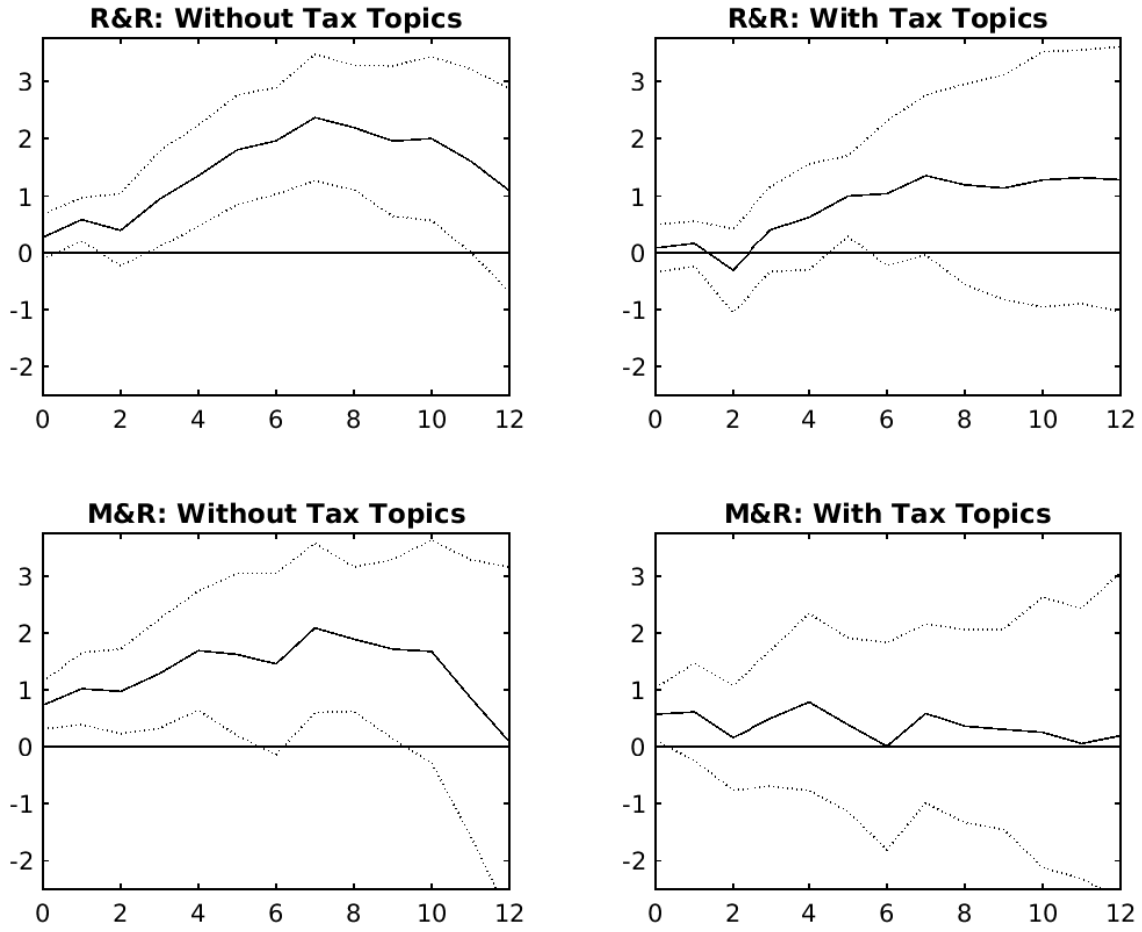



Figure 11: Output responses to a tax shocks identified using RR exogenous tax change series (first row) and Mertens and Ravn (2014) unanticipated tax shock proxy (second row).

approach allows us to automatically determine what issues are mentioned in the texts, and in particular - when the president discusses tax changes. Based on those results we create a measure of prevalence for the *tax increase* and *tax decrease* topics, which reflects the relative prominence across time of those changes on the president's agenda. We show that they are strong predictors for a variety of measures of tax changes, including those usually considered as unanticipated. This predictive power is not connected to other macroeconomic conditions, but is rather the effect of capturing the legislative process behind those changes.

By excluding the part of our prevalence measures which depends on past changes

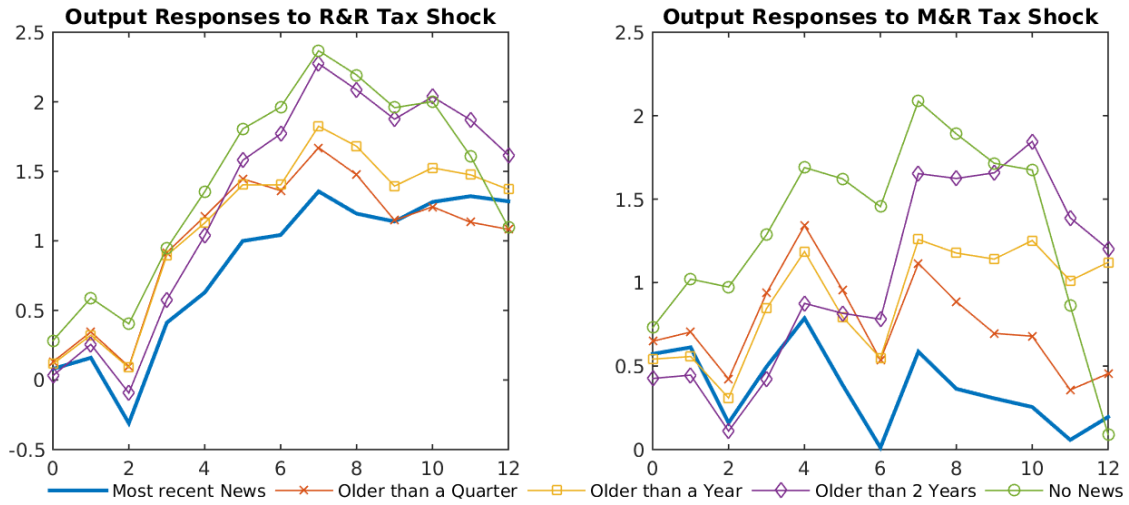



Figure 12: Output responses to a RR and Mertens and Ravn's (2014) tax shock including estimated with more and more distant lags of the tax topics.

and relating the remainder to implemented future tax changes we construct a measure of *noisy tax news*. By doing so, we quantify the signals concerning future changes in monetary terms, while retaining the proper timing of the arrival of new information. Using our noisy tax news measure we estimate the effect that information about future tax changes has on macroeconomic aggregates. While our results are mostly in accordance with the related literature, we do not find evidence for the contractionary effect of tax cut news. Finally, we empirically confirm that by omitting tax news standard approaches are prone to overestimating the impact of tax policy changes.

There are several methodological implications of our findings. First, we show that presidential speeches contain signals about future tax changes, and as such might be crucial for solving the problem of fiscal foresight. Moreover, it is possible to meaningfully quantify those signals using an automated text-analytic approach. In particular, we propose a two-step estimation procedure for the LDA model, in which informative lexical priors are constructed to differentiate between similar topics. Our approach, while relatively simple, proves crucial for determining the direction of the discussed changes.

Our results provide motivation for a variety of extensions. While our approach is able to identify news on a quarterly basis, the accuracy is lower when considering particular documents. Additional Natural Language Processing techniques could be

implemented to improve that, for example by considering the syntax of the statements. This might not only eliminate some noise in our measure, but also allow for a higher-frequency analysis.

Furthermore, our two-step approach could prove helpful in identifying news regarding different types of taxes (cf. Mertens and Ravn (2013)), however the identification of relevant lexicons would be more challenging than in this study. Finally, our approach could in principle be used to identify news about government spending. Because text-analytic methods scale well, in both cases the analysis can be augmented by considering other channels of communication such as congressional records, news outlets or even Twitter.

From an econometric perspective an important extension would be to integrate estimation of the text model and the structural econometric model into a single procedure. While considering two separate steps has the advantage that one can build on established methods for both the text mining and the structural modeling, a single unified framework could potentially better exploit the causality that runs in both directions. By directly extracting only the exogenous components of the president's speeches, one could measure the (noisy) news content of speeches more accurately and thereby mimic economic agents' reception of potential news more closely. Such an approach, however, would require the development of new models and estimation methods combining text and economic data in a single approach, which would be rather challenging. Given the current abundance of textual data sources, such an approach would be greatly beneficial to economic analysis, and therefore seems a promising future research agenda.

# References


Auerbach, A. J. and J. Slemrod (1997). The economic effects of the Tax Reform Act of 1986. *Journal of Economic Literature 35*(2), 589–632.

Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics 131*(4), 1593–1636.

Beaudry, P. and F. Portier (2014). News-driven business cycles: Insights and challenges. *Journal of Economic Literature 52*(4), 993–1074.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research 3*(Jan), 993–1022.

Chahrour, R. and K. Jurado (2018). News or noise? The missing link. *American Economic Review 108*(7), 1702–36.

Dybowski, T. and P. Adämmer (2018). The economic effects of US presidential tax communication: Evidence from a correlated topic model. *European Journal of Political Economy 55*, 511–525.

Forni, M., L. Gambetti, M. Lippi, and L. Sala (2017). Noisy news in business cycles. *American Economic Journal: Macroeconomics 9*(4), 122–52.

Gentzkow, M., B. Kelly, and M. Taddy (2019). Text as data. *Journal of Economic Literature 57*(3), 535–74.

Hansen, S. and M. McMahon (2016). Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics 99*, S114–S133.

Hansen, S., M. McMahon, and A. Prat (2018). Transparency and deliberation within the FOMC: A computational linguistics approach. *The Quarterly Journal of Economics 133*(2), 801–870.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica 47*(1), 153–161.

Hecq, A., L. Margaritella, and S. Smeekes (2019). Granger causality testing in high-dimensional VARs: a post-double-selection procedure. arXiv e-print 1902.10991.

Jaimovich, N. and S. Rebelo (2009). Can news about the future drive the business cycle? *American Economic Review 99*(4), 1097–1118.

Jelodar, H., Y. Wang, C. Yuan, X. Feng, X. Jiang, Y. Li, and L. Zhao (2019). Latent Dirichlet Allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications 78*(11), 15169–15211.

Larsen, V. H. and L. A. Thorsrud (2019). The value of news for economic developments. *Journal of Econometrics 210*(1), 203–218.

Lazarus, E., D. J. Lewis, J. H. Stock, and M. W. Watson (2018). HAR inference: Recommendations for practice. *Journal of Business and Economic Statistics 36*(4), 541–559.

Leeper, E. M., A. W. Richter, and T. B. Walker (2012). Quantitative effects of fiscal foresight. *American Economic Journal: Economic Policy 4*(2), 115–44.

Leeper, E. M., T. B. Walker, and S.-C. S. Yang (2013). Fiscal foresight and information flows. *Econometrica 81*(3), 1115–1145.

Mertens, K. and M. O. Ravn (2012). Empirical evidence on the aggregate effects of anticipated and unanticipated US tax policy shocks. *American Economic Journal: Economic Policy 4*(2), 145–81.

Mertens, K. and M. O. Ravn (2013). The dynamic effects of personal and corporate income tax changes in the United States. *American Economic Review 103*(4), 1212–47.

Mertens, K. and M. O. Ravn (2014). A reconciliation of SVAR and narrative estimates of tax multipliers. *Journal of Monetary Economics 68*, S1–S19.

Minka, T. (2000). Estimating a dirichlet distribution. Technical report, MIT.

Montiel-Olea, J. L. and M. Plagborg-Møller (2020). Local projection inference is simpler and more robust than you think. arXiv e-print 2007.13888.

Nakamura, E. and J. Steinsson (2018). High-frequency identification of monetary non-neutrality: The information effect. *Quarterly Journal of Economics 133*(3), 1283–1330.

Plagborg-Møller, M. and C. K. Wolf (2021). Local projections and vars estimate the same impulse responses. *Econometrica 89*, 955–980.

Ramey, V. A. (2009). Defense news shocks, 1939–2008: Estimates based on news sources. Working paper, University of California, San Diego.

Ramey, V. A. (2011). Identifying government spending shocks: It's all in the timing. *Quarterly Journal of Economics 126*(1), 1–50.

Ramey, V. A. (2016). Macroeconomic shocks and their propagation. NBER Working Papers 21978, National Bureau of Economic Research.

Ramey, V. A. and S. Zubairy (2018). Government spending multipliers in good times and in bad: evidence from us historical data. *Journal of Political Economy 126*(2), 850–901.

Romer, C. D. and D. H. Romer (2009a). A narrative analysis of postwar tax changes. Working paper, University of California, Berkeley.

Romer, C. D. and D. H. Romer (2009b). Replication data for: The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. Nashville, TN: American Economic Association [publisher], 2010. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E112357V1.

Romer, C. D. and D. H. Romer (2010). The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. *American Economic Review 100*(3), 763–801.

Sato, I. and H. Nakagawa (2010). Topic models with power-law using pitman-yor process. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 673–682.

Shapiro, A. H., M. Sudhof, and D. J. Wilson (2020). Measuring news sentiment. *Journal of Econometrics*, forthcoming.

Stock, J. and M. Yogo (2005). Asymptotic distributions of instrumental variables statistics with many instruments. In *Identification and Inference for Econometric Models*, pp. 109–120. New York: Cambridge University Press.

Wallach, H. M. (2008). *Structured topic models for language*. Ph. D. thesis, University of Cambridge Cambridge, UK.

Wallach, H. M., D. M. Mimno, and A. McCallum (2009). Rethinking LDA: Why priors matter. In *Advances in Neural Information Processing Systems*, pp. 1973–1981.

Yang, S.-C. S. (2007). A chronology of postwar US federal income tax policy. CAEPR Working Papers 2007-021, Center for Applied Economics and Policy Research, Indiana University.

# Appendix A Additional results for the prevalence and tax news measures

## A.1 Results about predictability

In Tables A.1-A.4 we investigate the predictive ability of the tax prevalence and noisy tax news measures for a variety of tax change measures by means of $F$-tests (Stock and Yogo, 2005). We regress various measures of tax changes on the tax topics and a set of control variables. More specifically, we estimate the following predictive regressions

$$\Delta T_t = \gamma_0 + \boldsymbol{\gamma}_1' \boldsymbol{z}_{t-j} + \sum_{i=1}^{4} \boldsymbol{\beta}_i' \boldsymbol{x}_{t-i} + u_t \quad \text{for } j = 0, 1, 2, ..., 6, \tag{A.1}$$

where $\Delta T_t$ is a tax change series, $\boldsymbol{z}_t$ contains tax topic prevalence measures, and $\boldsymbol{x}_t$ contains control variables. We estimate (A.1) for several endogenous tax change series. Controls always include first difference of log GDP and log government spending, as well as interest rates. Moreover four lags of the dependent variable are included.[20] The strength of the correlation between the tax topics contained in $\boldsymbol{z}_{t-j}$ and $\Delta T_t$ $j$ quarters ahead can be assessed by comparing the (marginal) $F$-statistic to suggestions in Stock and Yogo (2005). As measures of tax changes we consider changes in (cyclically adjusted) tax receipts and two of RR's narrative measures of tax changes.[21]

Table A.1 reveals that the prevalence measure obtained from the first-stage unsupervised LDA (i.e. the general tax topic) is not a powerful predictor for any tax change measure considered. This finding serves as a note of caution against a too confident use of unsupervised text analytics. Table A.2 shows the (marginal) $F$-statistic on the (joint) exclusion of the tax cut and tax hike topics from various predictive regressions. Regardless of how we measure future tax changes, the $F$-statistic is, for several forecast horizons, well above the cut-off. This suggests that the tax topics series can be considered a strong predictor of future tax changes. Therefore, the tax prevalence measures could also be used as strong instruments for actual tax changes.

One may argue that, due to their construction, our prevalence series may pick

[20] Our analysis is build on time series sampled at quarterly frequency from 1948Q1-2007Q4. Details about the economic data used in this paper can be found in Appendix C.2.

[21] RR link tax changes directly to the legislative process and, thus, their narrative contains more precise information about the timing of tax changes. We consider RR's series quantifying expected changes in current tax liabilities at time of implementation (i.e. for the first fiscal year the law was scheduled to be in effect) and the associated discounted present value at time of enactment. The latter measure is interpreted by RR as a signal containing tax news (in the strict sense of providing (almost) perfect foresight, i.e. there is no uncertainty of whether the tax change will happen or not). Further, RR construct sub-components of those measures capturing "exogenous" tax changes only (i.e. changes not driven by recent economic conditions). All narratives taken from RR are expressed as ratio to nominal GDP.

| Forecast Horizon | 0 | 1 | 2 | 3 | 4 | 5 | 6 | $H_{max}$ | $F_{max}$ |
|---|---|---|---|---|---|---|---|---|---|
| **Tax change measures**: | | | | | | | | | |
| Change in aggregate tax receipts | 0.03 | 0 | 0.11 | 0.3 | 0.63 | 0.04 | 0.31 | 4 | 0.63 |
| Change in cycl. adj. revenue | 0.06 | 2.44 | 3.4 | 2.1 | 3.18 | 0 | 2.9 | 2 | 3.4 |
| R&R (2010) – Expected change in tax liabilities at time of implementation | 0.73 | 5.38 | 1.59 | 1.04 | 0.83 | 0.58 | 6.37 | 6 | 6.37 |
| R&R (2010) – Expected change in tax liabilities at time of enactment | 0.45 | 0.2 | 0.34 | 1.97 | 1.71 | 0.75 | 1.83 | 3 | 1.97 |
| **Spending change measures**: | | | | | | | | | |
| Change in government consumption expenditures | 1.73 | 0 | 3.6 | 1.3 | 3.38 | 0.01 | 0.13 | 2 | 3.6 |
| Ramey's (2011) spending news | 0.11 | 0.66 | 0.18 | 1.17 | 0.65 | 0.12 | 0.03 | 3 | 1.17 |

Table A.1: $F$-statistics on the exclusion of the prevalence measure obtained from the first-stage unsupervised LDA.

up news signals about other policy plans that would affect the federal budget in the future, notably government spending. However, as indicated in the last rows of Table A.2, our measures do not possess much power in predicting federal spending (news).[22]

For comparison, we additionally summarize $F$-tests indicating the predictive strength of the implicit tax rates from Leeper et al. (2012) (see Table A.3). Finally, Table A.4 shows $F$-statistics based on the regression in (A.1) with $z_{t-j}$ replaced by the estimated tax news series. Displayed is the maximum $F$-statistic over $j = 0, 1, 2, ...6$ in (A.1) for different values of $h$ in (5).

Finally, In Figure A.1 we report the results of the Granger causality tests of Hecq et al. (2019) as described in Section 5.1 for the prevalence measures of the other topics.

[22]Overall, the above findings are robust across different specifications of the regression in (A.1). Unreported results, considering other control variables such as the party of the sitting president, election year, etc., do not alter the results.

| Forecast Horizon | 0 | 1 | 2 | 3 | 4 | 5 | 6 | $H_{max}$ | $F_{max}$ |
|---|---|---|---|---|---|---|---|---|---|
| **Tax change measures**: | | | | | | | | | |
| Change in aggregate tax receipts | 1.76 | 10.25 | 11.65 | 19.87 | 6.48 | 5.77 | 0.7 | 3 | 19.87 |
| Change in cycl. adj. revenue | 0.09 | 6.2 | 10.87 | 11.85 | 6.56 | 7.34 | 2.95 | 3 | 11.85 |
| R&R (2010) – Expected change in tax liabilities at time of implementation | 1.76 | 4.44 | 9.56 | 16.76 | 19.23 | 12.05 | 13.09 | 4 | 19.23 |
| R&R (2010) – Expected change in tax liabilities at time of enactment | 2 | 0.26 | 8.33 | 9.1 | 14.28 | 14.55 | 8.83 | 5 | 14.55 |
| **Spending change measures**: | | | | | | | | | |
| Change in government consumption expenditures | 2.27 | 0.93 | 2.2 | 0.55 | 0.99 | 1.24 | 0.82 | 0 | 2.27 |
| Ramey's (2011) spending news | 0.38 | 0.8 | 0.73 | 0.54 | 0.54 | 0.04 | 0.01 | 1 | 0.8 |

Table A.2: $F$-statistics on the joint exclusion of the tax cut and tax hike topics.

| Forecast Horizon | 0 | 1 | 2 | 3 | 4 | 5 | 6 | $H_{max}$ | $F_{max}$ |
|---|---|---|---|---|---|---|---|---|---|
| **Tax change measures**: | | | | | | | | | |
| Change in aggregate tax receipts | 2.15 | 3.62 | 0.06 | 0.27 | 0.57 | 0.34 | 0.17 | 1 | 3.62 |
| Change in cycl. adj. revenue | 0.42 | 2.5 | 0.27 | 0.75 | 0.28 | 0.44 | 0.01 | 1 | 2.5 |
| R&R (2010) – Expected change in tax liabilities at time of implementation | 1.51 | 0.36 | 0.81 | 0.05 | 1.57 | 2.69 | 3.8 | 6 | 3.8 |
| R&R (2010) – Expected change in tax liabilities at time of enactment | 2.02 | 0.48 | 1.23 | 0.01 | 2.18 | 2.36 | 4.09 | 6 | 4.09 |
| **Spending change measures**: | | | | | | | | | |
| Change in government consumption expenditures | 0.05 | 0.06 | 0 | 0.59 | 0.66 | 1.68 | 0.32 | 5 | 1.68 |
| Ramey's (2011) spending news | 2.39 | 0.55 | 1.74 | 1.43 | 0.08 | 0.34 | 1.42 | 0 | 2.39 |

Table A.3: $F$-statistics on the exclusion of implicit tax rates from Leeper et al. (2012).

| Forecast Horizon | 1 | 2 | 3 | 4 | 5 | 6 | $h_{max}$ | $F_{max}$ |
|---|---|---|---|---|---|---|---|---|
| **Tax change measures**: | | | | | | | | |
| Change in aggregate tax receipts | 25.73 | 19.26 | 25.24 | 8.29 | 2.24 | 11.26 | 1 | 25.73 |
| Change in cycl. adj. revenue | 35.1 | 25.58 | 14.2 | 10.39 | 1.09 | 8.49 | 1 | 35.1 |
| R&R (2010) – Expected change in tax liabilities at time of implementation | 43.55 | 24 | 24.06 | 14.65 | 13.53 | 23.61 | 1 | 43.55 |
| R&R (2010) – Expected change in tax liabilities at time of enactment | 53.45 | 28.3 | 29.54 | 16.65 | 14.85 | 10.07 | 1 | 53.45 |
| Leeper et al. (2012) implicit tax rates | 7.91 | 5.3 | 24.07 | 12.25 | 6.16 | 1.57 | 3 | 24.07 |
| **Spending change measures**: | | | | | | | | |
| Change in government consumption expenditures | 2.33 | 3.89 | 0.73 | 1.64 | 4.1 | 2.38 | 5 | 4.1 |
| Ramey's (2011) spending news | 1.59 | 2.08 | 0.65 | 2.69 | 4.72 | 2.69 | 5 | 4.72 |

Table A.4: $F$-statistics on the exclusion of the (noisy) tax news series. The statistics are based on the regression in (A.1) with $z_{t-j}$ replaced by the estimated tax news series. Displayed is the maximum $F$-statistic over $j = 0, 1, 2, ...6$ in (A.1) for different values of $h$ in (5).

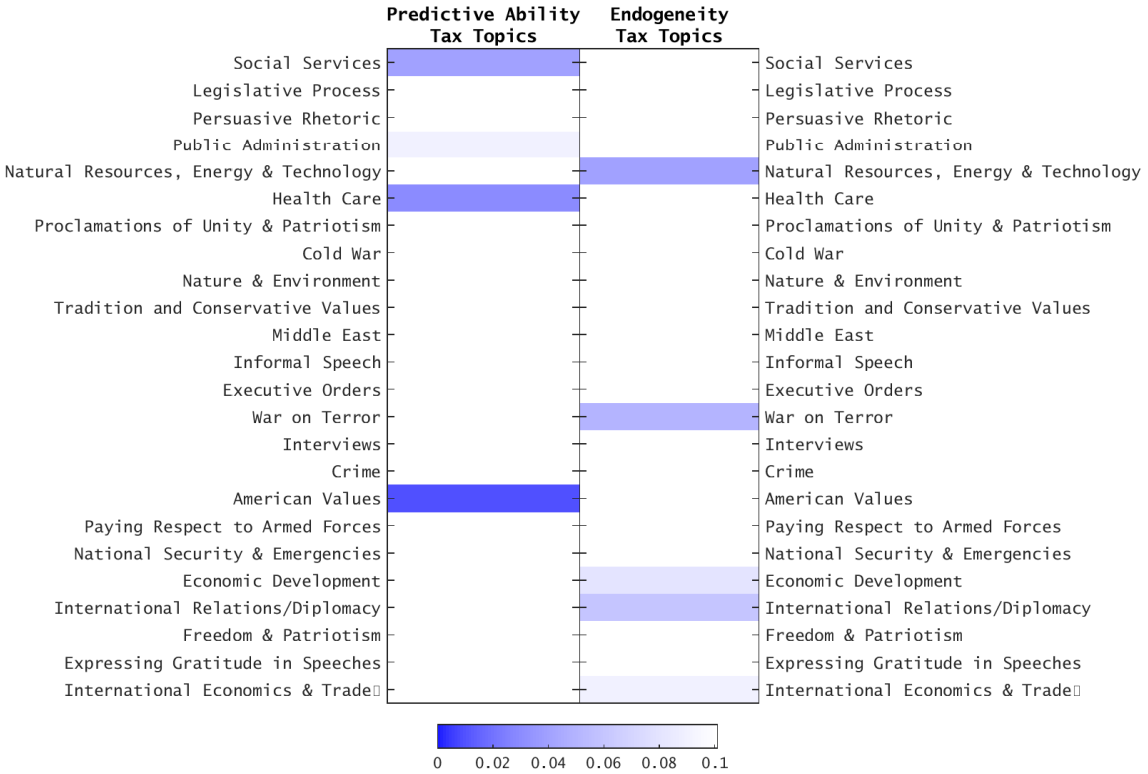



Figure A.1: $p$-Values from Granger causality tests. Prevalence measures of other identified topics.

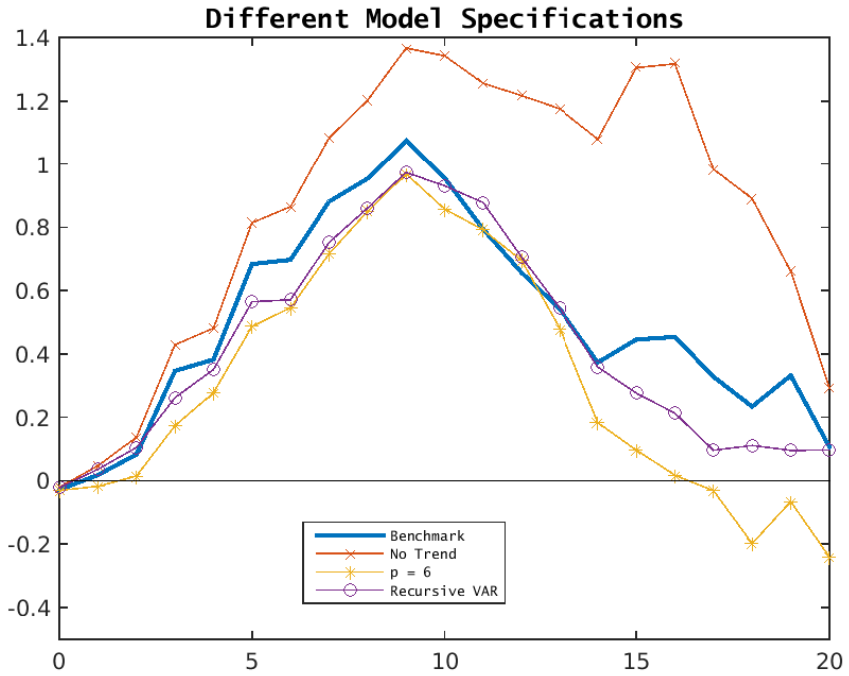



Figure A.2: Output responses to a change in $\text{News}_t^1$ of the size of minus one percent of GDP. Different model specifications, setting $h = 1$ and the benchmark set of controls. The VAR is estimated with 12 lags, intercept, linear and quadratic trend.

## A.2 Effects of tax news for various model specifications

Figure A.2 shows output responses for our benchmark set of controls yet for different model specifications. Including six instead of 12 lags has little quantitative impact on the response of output. Excluding any trend leads - unsurprisingly - to a more persistent response. We also compare our results to a SVAR estimate of output response to a shock in $\text{News}_t^1$. The SVAR is identified by a recursive structure, ordering our news variable first. Instead of estimating a proxy-VAR, we follow the suggestion in Plagborg-Møller and Wolf (2021) to use a Cholesky identification scheme instead. The authors show that both approaches are conceptually the same, but unlike the proxy-VAR a 'Cholesky-VAR' remains valid even if the shock of interest is non-invertible.

Next, we extend our benchmark framework in (6) in several directions. First, we study whether the sign of the change in $\text{News}_t^1$ causes asymmetric reactions of output. For this we split $\text{News}_t^1$ in two series: one containing only positive values (or zeros) the other one only negative ones (or zeros). The first row in Figure A.3 shows impulse responses to both of these series estimated using our benchmark specification. We do not find supporting evidence for possible nonlinear effects. Second, we interact our news signal with the recession index from the NBER to shed some light on non-linear

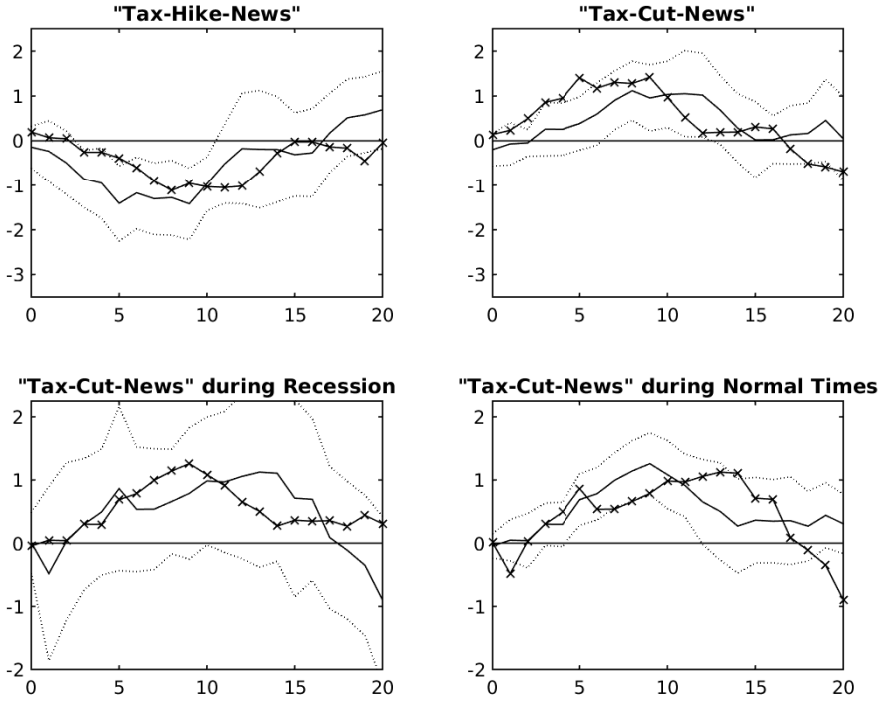



Figure A.3: Output responses to tax news

effects of tax changes over the business cycle. The results shown in the second row of Figure A.3 suggest again that the effects of tax news is not influenced by the state of the economy. The insignificant output response during times of economic slack, is likely driven by the small number of observations in the recession regime.[23]

Figure A.4 shows output responses to a tax shock identified using RR's exogenous tax change series and Mertens and Ravn's (2014) unanticipated tax shock proxy. The VAR is estimated with the same set of regressors and using the same trend specification as the local projection regression in Section 6.2. 12 lags of all endogenous variables are included. SVAR responses are identified recursively, ordering the tax change narrative (RR or Mertens and Ravn's (2014)) first.

Finally, Figure A.5 compares responses of output to two slightly differently constructed versions of (noisy) tax news. The left quadrant displays responses when $News_t^1$ is estimated as per (3)-(5). For estimating the responses in the right quadrant we have replaced the regressand $\Delta T_{\text{exo},t}^{PV}$ with $\Delta T_{\text{all},t}^{PV}$ in (4) and have added six lags of log output and log government spending as additional controls. Responses are then estimated in both cases using the benchmark specification of (6).

[23]Indeed, we find that output responses are significant in both economic up- and downswings when choosing an alternative threshold of the 20th percentile of real (annualised) year-on-year GDP growth.

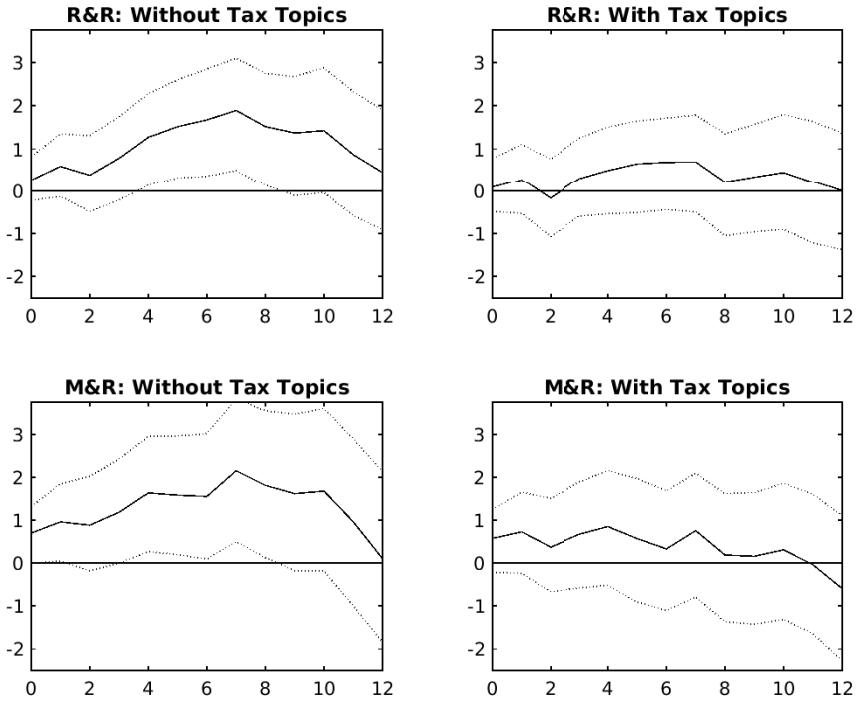



Figure A.4: Output responses to a tax shocks identified using RR exogenous tax change series (first row) and Mertens and Ravn (2014) unanticipated tax shock proxy (second row).

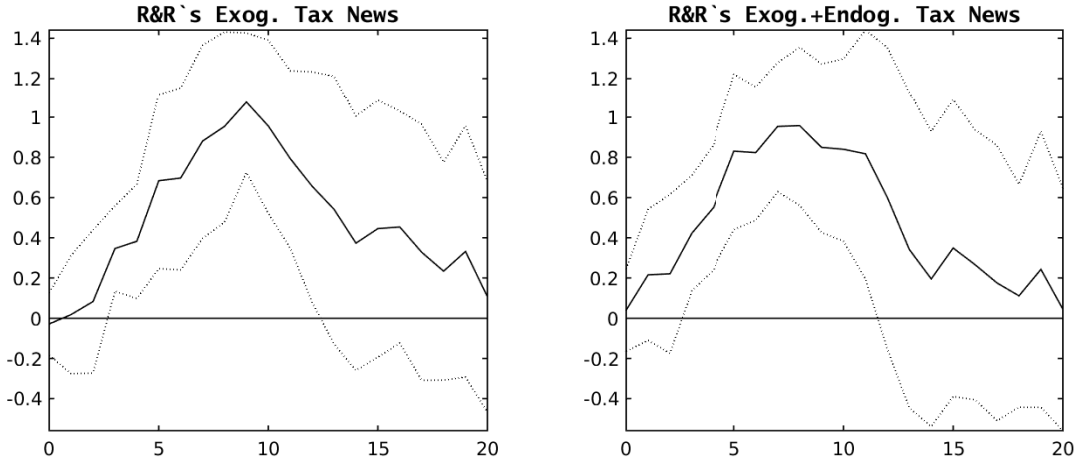



Figure A.5: Responses of output to two slightly different versions of (noisy) tax news $News_t^1$. Left panel: using $\Delta T_{\text{exo},t}^{PV}$ as regressand in (4). Right panel: using $\Delta T_{\text{all},t}^{PV}$ as regressand in and adding six lags of log output and log government spending as further controls in (4).

# Appendix B Details for the two-step LDA model

## B.1 Dirichlet distribution

A $K$-dimensional Dirichlet distribution is a continuous probability distribution, where a draw from the distribution is a vector of $K$ non-negative numbers summing up to 1. For $\boldsymbol{\theta} \sim \text{Dir}(\boldsymbol{\alpha})$ the parameter vector $\boldsymbol{\alpha} = a\boldsymbol{\mu}$ is decomposed into two parts - a scalar concentration parameter $a = \sum_{k=1}^{K} \alpha_k$ and the mean vector $\boldsymbol{\mu} = \mathbb{E}[\boldsymbol{\theta}] = \frac{1}{a}\boldsymbol{\alpha}$. The higher the $a$ the more concentrated the draws will be around the mean, following $\text{Var}[\theta_k] = \frac{\mu_k(1-\mu_k)}{a+1}$. This is visualised in Figure B.1.

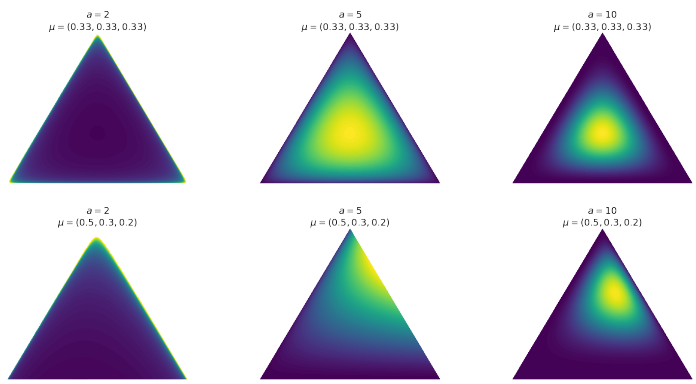



Figure B.1: Probability density graphs for $K = 3$-dimensional Dirichlet distributions with mean $\boldsymbol{\mu}$ and concentration parameter $a$. The lighter the colour, the higher the density.

A specific advantage of the Dirichlet distribution for estimating LDA models is the conjugacy property of this distribution in Bayesian inference. As an example, consider a 3-topic model with $\theta_d \sim \text{Dir}(2, 1, 0.5)$. If based on the data we believe that the document $d$ contains 0 tokens assigned to topic 1, 10 tokens assigned to topic 2 and 5 tokens assigned to topic 3, the posterior distribution for its mixing proportions is given by $\text{Dir}(0+2, 10+1, 5+0.5)$. Intuitively, the parameter vector for the prior can then be interpreted as a count vector of past observations, where we can influence the relative importance of terms through $\boldsymbol{\mu}$, and the weight of prior information through $a$.

The use of Dirichlet priors leads to the LDA's tendency to concentrate big part of the probability mass of the mixing proportions in a relatively small number of topics, and big part of the probability mass of the topics' distributions in a relatively small number of terms. This behavior can be best seen when considering the distribution over the topic assignment $z_{d,n}$ for some token $w_{d,n}$, conditional on the topic

assignments for the rest of the corpus $\mathcal{W}^{\neg d,n}, \mathcal{Z}^{\neg d,n}$:

$$\mathbb{P}(z_{d,n} = k | w_{d,n} = v_i, \mathcal{W}^{\neg d,n}, \mathcal{Z}^{\neg d,n}) \propto \underbrace{(N_d^{(k)} + \alpha_k)}_{\text{document}} \underbrace{\frac{N^{(k,i)} + \eta_{k,i}}{\sum_{d=1}^{D} N_d^{(k)} + \sum_{j=1}^{V} \eta_{k,j}}}_{\text{topic}},$$

where $N_d^{(k)} = \sum_{n=1}^{N_d} \mathbb{1}(z_{d,n} = k)$ is the number of tokens assigned to topic $k$ in document $d$ and $N^{(k,i)} = \sum_{d=1}^{D} \sum_{n=1}^{N_d} \mathbb{1}(w_{d,n} = v_i)\mathbb{1}(z_{d,n} = k)$ is the number of tokens with the term $v_i$ assigned to topic $k$. This probability consists of two parts:

- In the document part, the more tokens within a document are identified as coming from topic $k$, the higher is $N_d^{(k)}$ and the higher the likelihood that the next token is also identified as coming from topic $k$.
- In the topic part, the more times the given term $v_i$ is identified as coming from topic $k$ (across the whole corpus), the higher $N^{(k,i)}$ and the higher the likelihood that the next instance of that term is also identified as coming from topic $k$.

## B.2 Model specification

Our approach requires that the following parameters are specified. In the first, unsupervised step the number of topics $K$ is 25; the shape and scale parameters $(c, s)$ for the Gamma priors of the elements of $\boldsymbol{\eta}$ and $\boldsymbol{\alpha}$ are chosen to be $100^{-1}$ and 100 respectively. In the second step we split the *tax* topic into *tax increase* and *tax decrease* topic, resulting in $K = 26$; the parameters $(c, s)$ for the Gamma priors of the elements of $\boldsymbol{\alpha}$ remain at $100^{-1}$ and 100 respectively, while the vectors $\boldsymbol{\eta}_k, \quad k = 1, \ldots, K$ are constructed according to (2).

In the first step the model consists of the following elements:

$$\eta_{k,i} = \eta, \quad \eta \sim \text{Gamma}(c, s) \qquad p(\eta) = \frac{1}{\Gamma(s)c^s} \eta^{s-1} e^{-\frac{\eta}{c}} \tag{B.1}$$

$$\boldsymbol{\phi}_k | \boldsymbol{\eta}_k \sim \text{Dir}(\boldsymbol{\eta}_k) \qquad p(\boldsymbol{\phi}_k | \boldsymbol{\eta}_k) = \frac{\Gamma\left(\sum_{i=1}^{V} \eta_{k,i}\right)}{\prod_{i=1}^{V} \Gamma\left(\eta_{k,i}\right)} \prod_{i=1}^{V} {\phi_{k,i}}^{\eta_{k,i}-1} \tag{B.2}$$

$$\alpha_k \sim \text{Gamma}(c, s) \qquad p(\alpha_k) = \frac{1}{\Gamma(s)c^s} \alpha_k^{s-1} e^{-\frac{\alpha_k}{c}} \tag{B.3}$$

$$\boldsymbol{\theta}_d | \boldsymbol{\alpha} \sim \text{Dir}(\boldsymbol{\alpha}) \qquad p(\boldsymbol{\theta}_d | \boldsymbol{\alpha}) = \frac{\Gamma\left(\sum_{k=1}^{K} \alpha_k\right)}{\prod_{k=1}^{K} \Gamma\left(\alpha_k\right)} \prod_{k=1}^{K} {\theta_{d,k}}^{\alpha_k - 1} \tag{B.4}$$

$$z_{d,n} | \boldsymbol{\theta}_d \sim \text{Cat}(\boldsymbol{\theta}_d) \qquad \mathbb{P}(z_{d,n} = k | \boldsymbol{\theta}_d) = \theta_{d,k}$$

$$w_{d,n} | z_{d,n} = k, \boldsymbol{\Phi} \sim \text{Cat}(\boldsymbol{\phi}_k) \qquad \mathbb{P}(w_{d,n} = v_i | z_{d,n} = k) = \phi_{k,i}$$

In the second step, instead of estimating $\boldsymbol{\eta}_k, \quad k = 1, \ldots, K$ according to (B.1) we construct them according to (2) and treat them as known.

## B.3 Estimation of the topic model

Here we describe the Bayesian estimation procedure to obtain the posterior distribution of the document mixing proportions $\boldsymbol{\Theta}$ and topics' term probabilities $\boldsymbol{\Phi}$, conditional on the observed corpus of documents $\mathcal{W}$ and the Dirichlet distribution parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\eta}_k$'s.

The complete data likelihood of the LDA model is

$$p\big(\mathcal{W}, \mathcal{Z} | \boldsymbol{\Theta}, \boldsymbol{\Phi}\big) = \prod_{d=1}^{D} \prod_{n=1}^{N_d} \prod_{i=1}^{V} \prod_{k=1}^{K} \left(\theta_{d,k} \phi_{k,i}\right)^{\mathbb{1}(z_{d,n}=k)\mathbb{1}(w_{d,n}=v_i)}. \tag{B.5}$$

The posterior of the model is obtained by combining the likelihood in (B.5) and the priors in (B.2) and (B.4):

$$\begin{aligned}
&p\big(\boldsymbol{\Theta}, \boldsymbol{\Phi}, \boldsymbol{\alpha}, \boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_K, \mathcal{Z} | \mathcal{W}\big) \\
&\quad \propto p\big(\mathcal{W}, \mathcal{Z} | \boldsymbol{\Theta}, \boldsymbol{\Phi}\big) p\big(\boldsymbol{\Theta} | \boldsymbol{\alpha}\big) p\big(\boldsymbol{\Phi} | \boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_K\big) p\big(\boldsymbol{\alpha}\big) p\big(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_K\big) \\
&\quad \propto \left( \frac{\Gamma\left(\sum_{k=1}^{K} \alpha_k\right)}{\prod_{k=1}^{K} \Gamma\left(\alpha_k\right)} \right)^{D} \prod_{k=1}^{K} \left( \frac{\Gamma\left(\sum_{i=1}^{V} \eta_{k,i}\right)}{\prod_{i=1}^{V} \Gamma(\eta_{k,i})} \right) p\big(\boldsymbol{\alpha}\big) p\big(\boldsymbol{\eta}_1, \ldots \boldsymbol{\eta}_K\big) \\
&\qquad \times \prod_{d=1}^{D} \prod_{k=1}^{K} \theta_{d,k}^{\alpha_k - 1 + \sum_{i=1}^{V} \sum_{n=1}^{N_d} \mathbb{1}(z_{d,n}=k)\mathbb{1}(w_{d,n}=v_i)} \\
&\qquad \times \prod_{k=1}^{K} \prod_{i=1}^{V} \phi_{k,i}^{\eta_{k,i} - 1 + \sum_{d=1}^{D} \sum_{n=1}^{N_d} \mathbb{1}(z_{d,n}=k)\mathbb{1}(w_{d,n}=v_i)}.
\end{aligned}$$

Posterior inference of the LDA mdel is computationally demanding due to the size of the corpus. More efficient methods, such as the variational Bayes algorithm of Blei et al. (2003), are applicable. We opt for the Gibbs sampler as it has better convergence properties in our application.

Let $\Omega = \{\boldsymbol{\Theta}, \boldsymbol{\Phi}, \boldsymbol{\alpha}, \boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_K\}$ denote the set of all model parameters and $\Omega_{-\kappa} = \Omega \setminus \{\kappa\}$ for variable $\kappa$ and variable set $\Omega$. The Gibbs sampling steps for the model parameters and latent states are as follows. First,

$$p\big(\boldsymbol{\theta}_d | \Omega_{-\boldsymbol{\theta}_d}, \mathcal{Z}, \mathcal{W}\big) \propto \prod_{k=1}^{K} \theta_{d,k}^{\alpha_k - 1 + \sum_{i=1}^{V} \sum_{n=1}^{N_d} \mathbb{1}(z_{d,n}=k)\mathbb{1}(w_{d,n}=v_i)}, \text{ for } d = 1, \ldots, D, \tag{B.6}$$

is proportional to a Dirichlet distribution, as in the conjugate prior in (B.4). Next,

$$p\big(\boldsymbol{\phi}_k|\Omega_{-\boldsymbol{\phi}_k},\mathcal{Z},\mathcal{W}\big) \propto \prod_{i=1}^{V} \phi_{k,i}{}^{\eta_{k,i}-1+\sum_{d=1}^{D}\sum_{n=1}^{N_d}\mathbb{1}(z_{d,n}=k)\mathbb{1}(w_{d,n}=v_i)}, \text{ for } k=1,\ldots,K, \quad \text{(B.7)}$$

is proportional to a Dirichlet distribution, as in the conjugate prior in (B.2). Similarly,

$$\begin{aligned} &p\big(z_{d,n}=k|\Omega,\mathcal{Z}_{-z_{d,n}},\mathcal{W}\big) \\ &\quad\propto \theta_{d,k}^{\alpha_k-1+\sum_{i=1}^{V}\mathbb{1}(w_{d,n}=v_i)} \prod_{i=1}^{V} \phi_{k,i}{}^{\eta_{k,i}-1+\mathbb{1}(w_{d,n}=v_i)}, \text{ for } k=1,\ldots,K \end{aligned} \quad \text{(B.8)}$$

which is a categorical distribution for $d=1,\ldots,D$ and $n=1,\ldots,N_d$ with probabilities proportional to the right hand side of (B.8).

For the remaining model parameters, $(\boldsymbol{\alpha},\boldsymbol{\eta}_1,\ldots,\boldsymbol{\eta}_K)$, we use the fixed-point iteration (MAP estimator) of Minka (2000); Wallach (2008) at each Gibbs iteration. We optimize $\boldsymbol{\alpha}=a\boldsymbol{\mu}$ as

$$[a\mu_k]^* = a\mu_k \frac{\sum_{d=1}^{D}\left[\Psi\left(N_d^{(k)}+a\mu_k\right)-\Psi\left(a\mu_k\right)\right]+c}{\sum_{d=1}^{D}\left[\Psi\left(N_d+a\right)-\Psi\left(a\right)\right]-\frac{1}{s}}, \quad \text{(B.9)}$$

and, only in the first step, we optimize $\eta$ using

$$\eta^* = \frac{\eta\sum_{k=1}^{K}\sum_{i=1}^{V}\left[\Psi\left(N^{(k,i)}+\eta\right)-\Psi\left(\eta\right)\right]+c}{V\sum_{k=1}^{K}\left[\Psi\left(\sum_{i=1}^{V}N^{(k,i)}+V\eta\right)-\Psi\left(V\eta\right)\right]-\frac{1}{s}}, \quad \text{(B.10)}$$

where $N_d^{(k)}=\sum_{n=1}^{N_d}\mathbb{1}(z_{d,n}=k)$ is the number of tokens assigned to topic $k$ in document $d$, $N^{(k,i)}=\sum_{d=1}^{D}\sum_{n=1}^{N_d}\mathbb{1}(w_{d,n}=v_i)\mathbb{1}(z_{d,n}=k)$ is the number of tokens with the term $v_i$ assigned to topic $k$, and $(c,s)$ are the shape and scale parameters, respectively, for the Gamma priors exemplified in (B.3) and (B.1). The steps in B.9 and B.10 are repeated until convergence[24]. The Gibbs sampler for the proposed 2-step LDA model is outlined in Algorithm B.1 below, for a corpus $\mathcal{W}$ obtained from pre-processed data, and $K$ topics.

[24]Until $\max_k \frac{[a\mu_k]^*-a\mu_k}{a\mu_k} < 0.01$ and $\max_k \frac{\eta^*-\eta}{\eta} < 0.01$

**Algorithm B.1** Gibbs sampling algorithm

LDA step 1:

1: Set $m = 0$.
2: Initialize $\alpha_k^{(m)} = 1$ for $k = 1, \dots K$ and $\eta = 1$.
3: Initialize $\boldsymbol{\Theta}^{(m)}$ using the Dirichlet distribution in (B.4); $z_{d,n}^{(m)}$ for $d = 1, \dots, D$ and $n = 1, \dots N_d$ using the categorical distribution in (1); $\boldsymbol{\phi}_k^{(m)}$ for all $k$ using the Dirichlet distribution in (B.2).
4: Set $\Omega^{(m)} = \{\boldsymbol{\Theta}^{(m)}, \boldsymbol{\Phi}^{(m)}, \boldsymbol{\alpha}^{(m)}, \boldsymbol{\eta}_1^{(m)}, \dots, \boldsymbol{\eta}_K^{(m)}\}$
5: **while** $m \leq 15000$ **do**
6: Set $m = m + 1$
7: Draw $\boldsymbol{\Theta}^{(m)} \mid \Omega_{-\boldsymbol{\Theta}}^{(m-1)}, \mathcal{Z}^{(m-1)}, \mathcal{W}^{(m-1)}$ using the Dirichlet distribution in (B.6)
8: Update $\Omega^{(m)} = \{\boldsymbol{\Theta}^{(m)}, \boldsymbol{\Phi}^{(m-1)}, \boldsymbol{\alpha}^{(m-1)}, \boldsymbol{\eta}_1^{(m-1)}, \dots, \boldsymbol{\eta}_K^{(m-1)}\}$
9: **for** $k = 1, \dots, K$ **do**
10: Draw $\boldsymbol{\phi}_k^{(m)} \mid \Omega_{\boldsymbol{\phi}_k}^{(m)}, \mathcal{Z}^{(m-1)}, \mathcal{W}$ for all $k$ using the Dirichlet distribution in (B.7)
11: Update $\boldsymbol{\Phi}^{(m)} = \{\boldsymbol{\phi}_1^{(m)}, \dots, \boldsymbol{\phi}_k^{(m)}, \boldsymbol{\phi}_{k+1}^{(m-1)}, \dots, \boldsymbol{\phi}_K^{(m-1)}\}$
12: Update $\Omega^{(m)} = \{\boldsymbol{\Theta}^{(m)}, \boldsymbol{\Phi}^{(m)}, \boldsymbol{\alpha}^{(m-1)}, \boldsymbol{\eta}_1^{(m-1)}, \dots, \boldsymbol{\eta}_K^{(m-1)}\}$
13: **end for**
14: Set $\mathcal{Z}^{(m)} = \{z_{1,1}^{(m-1)}, \dots, z_{1,N_1}^{(m-1)}, \dots, z_{D,1}^{(m-1)}, \dots, z_{D,N_D}^{(m-1)}\}$
15: **for** $d = 1, \dots, D$ **do**
16: **for** $n = 1, \dots, N_d$ **do**
17: Draw $z_{d,n}^{(m)} \mid \Omega^{(m)}, \mathcal{Z}_{-z_{d,n}}^{(m)}, \mathcal{W}$ using the categorical distribution in (B.8)
18: Update $\mathcal{Z}^{(m)} = \{z_{1,1}^{(m)}, \dots, z_{d,n}^{(m)}, z_{d,n+1}^{(m-1)}, \dots, z_{d+1,1}^{(m-1)} \dots, z_{D,N_D}^{(m-1)}\}$
19: **end for**
20: **end for**
21: Update $\boldsymbol{\alpha}^{(m)} = a\boldsymbol{\mu}$ using the fixed-point iteration in (B.9).
22: Set $\boldsymbol{\eta}_k^{(m)} = \boldsymbol{\eta}^{(m)} = a_\eta \boldsymbol{\iota}_V$ where $\boldsymbol{\iota}_V$ is a $V$-dimensional vector of $1/V$. Update $\boldsymbol{\eta}^{(m)}$ using the fixed-point iteration in (B.9).
23: **end while**

LDA step 2

1: Set $K = K + 1$.
2: Set $m = 0$.
3: Initialize $\alpha_k^{(m)} \sim \text{Gamma}^{(m)}(100^{-1}, 100)$ for $k = 1, \dots K$
4: Set $\boldsymbol{\eta}_k, \quad k = 1, \dots, K$ from step 1, using equation (2).
5: Initialize $\boldsymbol{\Theta}^{(m)}$ using the Dirichlet distribution in (B.4); $z_{d,n}^{(m)}$ for $d = 1, \dots, D$ and $n = 1, \dots N_d$ using the categorical distribution in (1); $\boldsymbol{\phi}_k^{(m)}$ for all $k$ using the Dirichlet distribution in (B.2).
6: Set $\Omega^{(m)} = \{\boldsymbol{\Theta}^{(m)}, \boldsymbol{\Phi}^{(m)}, \boldsymbol{\alpha}^{(m)}, \boldsymbol{\eta}_1^{(m)}, \dots, \boldsymbol{\eta}_K^{(m)}\}$
7: Repeat LDA step 1 lines 5–23, except for 22.

## B.4 Lexicons

Our procedure for constructing priors for the *tax increase* and *tax decrease* topics require specifying the terms whose usage differentiates the two topics. We do that based on reading presidential speeches which are known to provide information about future tax changes. We select a total of 69 speeches used by Romer and Romer (2009a) and Yang (2007) in their analyses of tax legislations. The list of the speeches is provided in Appendix D.1. In each speech we note the terms used to announce and motivate changes in particular direction. We then aggregate those results and choose terms whose usage is indicative of discussing a tax change with the particular direction.

Importantly, we do not assume that, for example, a term in the *tax increase* lexicon cannot be used when discussing tax cuts. We use those lexicons only to modify the prior probabilities of those terms in the two topics of interest.

**Tax increase:** "additional cost", "additional revenue", "additional tax", "balance budget", "budget deficit", "cut deficit", "defense spend", "deficit", "deficit reduction", "fair balance", "fair share", "fairness", "federal revenue", "fiscal responsibility", "fiscally responsible", "government revenue", "higher tax", "increase revenue", "increase tax", "military spend", "new spend", "new tax", "propose increase", "propose tax", "raise revenue", "raise tax", "reduce debt", "reduce deficit", "revenue increase", "rise cost", "sound fiscal", "tax impose", "tax increase", "tax revenue", "tax rich"

**Tax decrease:** "boost economy", "business incentive", "create incentive", "create job", "cut", "cut propose", "cut tax", "ease burden", "economic growth", "economic incentive", "excessive", "grow economy", "high rate", "incentive", "incentive invest", "incentivize", "increase employment", "increase investment", "increase production", "increase productivity", "increase prosperity", "investment", "investment tax", "lower", "lower tax", "new job", "propose cut", "provide incentive", "rate drop", "rate reduction", "rebuild economy", "recession", "reduce burden", "reduce rate", "reduce unemployment", "reduction", "relief", "relief package", "relief program", "slow growth", "stimulate economic", "stimulate economy", "stimulate growth", "stimulate investment", "strengthen economic", "strengthen economy", "tax break", "tax burden", "tax credit", "tax cut", "tax incentive", "tax policy", "tax rate", "tax rebate", "tax reduce", "tax reduction", "tax relief"

# Appendix C Data

## C.1 Text data

We analyze the texts of all the documents contained in *The Public Papers of the Presidents*. This includes speeches, but also other public communications (e.g. interviews, letters). We obtained the raw texts from the *American Presidency Project* (www.presidency.ucsb.edu) on 2019-03-25. Our dataset consists of 59,214 texts spanning from 1949-01-20 to 2017-01-19.

The pre-processing starts with breaking the texts into individual paragraphs, resulting in a total of 1,119,200 documents. The text is then broken along non-alphanumeric characters into tokens (individual words). We then remove stopwords - function words such as prepositions and pronouns, since they are not informative in the *bags-of-words* approach; and rare words which appear in less than 20 documents - those are often spelling mistakes, since they are too rare to be informative about the term co-occurrence pattern. The remaining tokens are then lemmatized - all the nouns are turned into singular form and the verbs are turned into present tense first person form. Lastly, we identify meaningful collocations of two words - bigrams. If a combination of two consecutive words (e.g. "tax" and "cut") appears more often than predicted by their individual frequency (as measured by a $\chi^2$ score) it is considered a bigram (i.e. "tax_cut"). A bigram is treated as a separate term in the vocabulary and each instance of a bigram in the texts is treated as a single token. The final vocabulary consists of 50,851 terms, including 29,815 bigrams.

## C.2 Other variables

We use four series of changes in tax liabilities developed by Romer and Romer (2010). The data-set is part of the paper's supplementary materials (Romer and Romer, 2009b). The construction of their measures is described in the companion background paper (Romer and Romer, 2009a). To create the measures they first identify all the major changes in tax liabilities. For each change they analyze the narrative record, including presidential speeches, to determine the motivation behind it. This can be either exogenous to the macroeconomic conditions (deficit-driven or long-run growth) or endogenous (spending-driven or countercyclical). The size of the changes is the estimated change in tax liabilities at the time of implementation based on contemporary sources. The changes are timed either at the implementation - when they go into effect, or at enactment - when the legislation is signed by the president, in which case the present value of the changes is given.

We additionally use the speeches identified by Romer and Romer (2009a) as giving motivation behind particular changes to verify the ability of our model to properly classify tax-related content. The list of those speeches is provided in Appendix D.1.

| Name | Source | Description |
|---|---|---|
| *Tax Measures:* | | |
| | | Federal government current tax receipts: |
| Tax receipts | BEA: W006RC | - all * |
| Income tax | BEA: A074RC | - personal taxes * |
| Corporate tax | BEA: B075RC | - taxes on corporate income * |
| Payroll tax | BEA: W780RC | Contributions for government social insurance * |
| Cyclically adjusted revenue | Romer and Romer (2010) | Change in real cyclically adjusted revenues as a percent of real GDP |
| Implicit tax rate | Leeper et al. (2012) | Risk-adjusted implicit tax rate based on U.S. municipal bonds |
| | | Legislated changes in tax liabilities: |
| RR All | Romer and Romer (2010) | - all, timed at implementation * |
| RR Exogenous | Romer and Romer (2010) | - exogenous motivation, timed at implementation * |
| RR News All | Romer and Romer (2010) | - all, present value at enactment * |
| RR News Exogenous | Romer and Romer (2010) | - exogenous motivation, present value at enactment * |
| MR Unanticipated | Mertens and Ravn (2014) | - exogenous motivation, only changes implemented less than a quarter after enactment, timed at implementation * |
| *Macroeconomic and financial variables:* | | |
| GDP | BEA: A191RC | Gross Domestic Product * |
| Consumption | BEA: DPCERC | Personal Consumption Expenditures * |
| Investment | BEA: A006RC | Gross Private Domestic Investment * |
| Government spending | BEA: W013RC | Federal Government: Current Expenditures * |
| Government debt | Federal Reserve Bank of Dallas | Change in Par Value of U.S. Government Debt * |
| 3-Month treasury bill | Federal Reserve Board, H.15 | Secondary Market Rate 3-Month Treasury Bill |
| Unemployment rate | BLS: LNS14000000 | Unemployed as percentage of the labor force |
| Inflation | BLS, CPI-U | Consumer Price Index for All Urban Consumers: All Items in U.S. City |
| SP500 | The data-set of Robert Shiller | Average Daily-Close * |
| Spending news | Ramey (2009) | * |
| *Other:* | | |
| Population | BLS: LNU00000000 | Number of persons, 16 years of age or older, noninstitutional |
| GDP deflator | BEA: A191RD | Gross domestic product implicit price deflator |
| Job approval rating | American Presidency Project | Gallup Poll approval of the way the president is handling his job |
| House and Senate concurrence | American Presidency Project | Percent of times a majority of members of Congress vote with the president's position on roll call votes |

Table C.1: List of variables used in Sections 5 and 6. * indicates variables that are divided by the population and the GDP deflator. BEA: U.S. Bureau of Economic Analysis; BLS: U.S. Bureau of Labour Statistics

# Appendix D Topic model results

## D.1 Selected tax-related speeches

Romer and Romer (2009a) and Yang (2007) use presidential speeches to analyze to motivation behind particular tax legislations. We combine their selection of speeches and use them for two purposes. Firstly, the speeches are used to develop the lexicons of terms related to tax increase and tax decrease, as described in Appendix B.4. Secondly, we use them to verify the ability of our model to properly classify a text based on which direction of tax changes it is discussing.

Below we present a list of speeches announcing legislations whose overall effect was a tax increase (Table D.1) and tax decrease (Table D.2). Importantly, in our estimation we treat individual paragraphs as documents, each having separate mixing proportions. Especially in case of long speeches, such as the *State of the Union Addresses*, many of those paragraphs discuss issues other than taxation. Additionally, certain legislations consisted of measures that increased some taxes, but cut others. In the tables those are indicated as "mixed". For speeches related to those legislations we expect some paragraphs to have high *tax increase* proportions, and other to have high *tax decrease* proportions. For each speech we therefore show the minimum, the average and the maximum mixing proportion across all of its paragraphs, for both *tax increase* and *tax decrease* topics, as estimated in the second step. Overall the results fit our expectations, supporting the claim that our model is able to distinguish between the direction of the discussed changes.

| Date | Announced legislation | Title of the speech | | *Tax increase* proportion | | | *Tax decrease* proportion | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Mixed | Number of paragraphs | Min. | Av. | Max. | Min. | Av. | Max. |
| *In a special message to Congress on tax policy* | | | | | | | | | |
| 1950/01/23 | Revenue Act of 1950 | Yes | 58 | 0.010 | 0.530 | 0.855 | 0.004 | 0.015 | 0.197 |
| *Midyear Economic Report of the President* | | | | | | | | | |
| 1950/07/26 | Revenue Act of 1950 | Yes | 92 | 0.004 | 0.383 | 0.848 | 0.003 | 0.059 | 0.387 |
| *Letter to Committee Chairmen on Taxation of Excess Profits* | | | | | | | | | |
| 1950/11/14 | Revenue Act of 1950 | Yes | 9 | 0.016 | 0.373 | 0.788 | 0.004 | 0.012 | 0.027 |
| *Annual Budget Message to Congress* | | | | | | | | | |
| 1951/01/15 | Revenue Act of 1951 | Yes | 399 | 0.002 | 0.262 | 0.907 | 0.002 | 0.017 | 0.352 |
| *Special Message to the Congress Recommending a "Pay as We Go" Tax Program* | | | | | | | | | |
| 1951/02/02 | Revenue Act of 1951 | Yes | 48 | 0.019 | 0.500 | 0.845 | 0.004 | 0.020 | 0.223 |
| *Radio address* | | | | | | | | | |
| 1953/05/19 | Extending the Excess Profits Tax Act of 1950 | No | 76 | 0.005 | 0.217 | 0.759 | 0.004 | 0.017 | 0.060 |
| *Annual Message to the Congress on the State of the Union* | | | | | | | | | |
| 1956/01/05 | Federal-Aid Highway Act of 1956 | No | 128 | 0.004 | 0.126 | 0.822 | 0.002 | 0.021 | 0.274 |
| *Annual Budget Message to the Congress* | | | | | | | | | |
| 1965/01/25 | Social Security Amendments of 1965 | No | 251 | 0.006 | 0.203 | 0.785 | 0.003 | 0.028 | 0.474 |
| *Annual Budget Message to the Congress* | | | | | | | | | |
| 1966/01/24 | Tax Adjustment Act of 1966 | No | 304 | 0.004 | 0.177 | 0.802 | 0.003 | 0.024 | 0.357 |
| *Special Message to the Congress on Fiscal Policy* | | | | | | | | | |
| 1966/09/08 | Public Law 89-800 | No | 97 | 0.007 | 0.277 | 0.809 | 0.003 | 0.064 | 0.608 |
| *State of Union address* | | | | | | | | | |
| 1967/01/10 | Revenue and Expenditure Control Act of 1968 | Yes | 178 | 0.004 | 0.075 | 0.851 | 0.004 | 0.028 | 0.498 |
| *Annual Budget Message to the Congress* | | | | | | | | | |
| 1967/01/24 | Revenue and Expenditure Control Act of 1968 | Yes | 285 | 0.005 | 0.176 | 0.871 | 0.003 | 0.024 | 0.558 |
| *Special Message to the Congress: The State of the Budget and the Economy* | | | | | | | | | |
| 1967/08/03 | Revenue and Expenditure Control Act of 1968 | Yes | 130 | 0.006 | 0.303 | 0.776 | 0.005 | 0.033 | 0.498 |
| *Annual Budget Message to the Congress* | | | | | | | | | |
| 1968/01/29 | Revenue and Expenditure Control Act of 1968 | Yes | 419 | 0.004 | 0.165 | 0.824 | 0.003 | 0.029 | 0.581 |
| *Special Message to Congress on Fiscal Policy* | | | | | | | | | |
| 1969/03/26 | Extending the Ten Percent Surtax | No | 18 | 0.124 | 0.435 | 0.805 | 0.005 | 0.035 | 0.248 |
| *Windfall Profits Tax and Energy Security Trust Fund Message to the Congress* | | | | | | | | | |
| 1979/04/26 | Crude Oil Windfall Profit Tax Act of 1980 | No | 77 | 0.011 | 0.304 | 0.782 | 0.003 | 0.019 | 0.150 |
| *Statement Announcing the Establishment of the National Commission on Social Security Reform* | | | | | | | | | |
| 1981/12/16 | Social Security Amendments of 1983 | No | 7 | 0.040 | 0.302 | 0.547 | 0.005 | 0.009 | 0.015 |
| *State of the Union address* | | | | | | | | | |
| 1982/01/26 | Tax Equity and Fiscal Responsibility Act of 1982 | No | 77 | 0.004 | 0.205 | 0.826 | 0.003 | 0.068 | 0.576 |

| Date | Announced legislation | Mixed | Number of paragraphs | *Tax increase* proportion Min. | Av. | Max. | *Tax decrease* proportion Min. | Av. | Max. |
|---|---|---|---|---|---|---|---|---|---|
| | | **Title of the speech** | | | | | | | |
| *Address Before a Joint Session of the Congress on the State of the Union* | | | | | | | | | |
| 1984/01/25 | Deficit Reduction Act of 1984 | No | 80 | 0.003 | 0.116 | 0.753 | 0.003 | 0.030 | 0.516 |
| *Message to the Congress Transmitting the Fiscal Year 1985 Budget* | | | | | | | | | |
| 1984/02/01 | Deficit Reduction Act of 1984 | No | 105 | 0.003 | 0.208 | 0.807 | 0.002 | 0.037 | 0.440 |
| *Remarks to Reporters Announcing a Deficit Reduction Plan* | | | | | | | | | |
| 1984/03/15 | Deficit Reduction Act of 1984 | No | 26 | 0.012 | 0.184 | 0.772 | 0.004 | 0.016 | 0.028 |
| *Joint Session of Congress on the State of the Union* | | | | | | | | | |
| 1987/01/27 | Omnibus Budget Reconciliation Act of 1987 | No | 34 | 0.003 | 0.073 | 0.631 | 0.003 | 0.040 | 0.682 |
| *Statement on Proposed Tax Increases* | | | | | | | | | |
| 1987/10/15 | Omnibus Budget Reconciliation Act of 1987 | No | 4 | 0.350 | 0.461 | 0.618 | 0.004 | 0.006 | 0.010 |
| *Statement on the federal budget negotiations* | | | | | | | | | |
| 1990/06/26 | Omnibus Budget Reconciliation Act of 1990 | No | 3 | 0.010 | 0.268 | 0.777 | 0.008 | 0.011 | 0.017 |
| *Address to the Nation on the Federal Budget Agreement* | | | | | | | | | |
| 1990/10/02 | Omnibus Budget Reconciliation Act of 1990 | No | 15 | 0.009 | 0.261 | 0.704 | 0.004 | 0.057 | 0.334 |
| *Address to the nation on the economic program* | | | | | | | | | |
| 1993/02/15 | Omnibus Budget Reconciliation Act of 1993 | Yes | 20 | 0.007 | 0.243 | 0.740 | 0.005 | 0.109 | 0.666 |
| *Address Before a Joint Session of Congress on Administration Goals* | | | | | | | | | |
| 1993/02/17 | Omnibus Budget Reconciliation Act of 1993 | Yes | 75 | 0.003 | 0.262 | 0.789 | 0.002 | 0.054 | 0.634 |
| *Radio Address* | | | | | | | | | |
| 1993/05/15 | Omnibus Budget Reconciliation Act of 1993 | Yes | 17 | 0.006 | 0.346 | 0.819 | 0.005 | 0.098 | 0.486 |

Table D.1: List of speeches announcing legislation whose overall effect was an increase in tax revenue. Mixed legislation includes both tax hikes and cuts.

| Date | Announced legislation | Title of the speech: Mixed | No. paragraphs | *Tax increase* proportion Min. | Av. | Max. | *Tax decrease* proportion Min. | Av. | Max. |
|---|---|---|---|---|---|---|---|---|---|
| *Letter to the President of the Senate and to the Speaker of the House of Representatives Urging Continuation of Corporation Tax Rates* | | | | | | | | | |
| 1958/05/26 | Tax Rate Extension Act of 1958 | No | 3 | 0.008 | 0.013 | 0.016 | 0.011 | 0.205 | 0.472 |
| *Radio and Television Report to the American People on the State of the National Economy* | | | | | | | | | |
| 1962/08/13 | Changes in Depreciation Guidelines and Revenue Act of 1962 | Yes | 68 | 0.004 | 0.150 | 0.750 | 0.005 | 0.218 | 0.744 |
| *Annual Message to the Congress on the State of the Union* | | | | | | | | | |
| 1963/01/14 | Revenue Act of 1964 | Yes | 86 | 0.002 | 0.037 | 0.525 | 0.003 | 0.120 | 0.895 |
| *Special Message to the Congress on Tax Reduction and Reform* | | | | | | | | | |
| 1963/01/24 | Revenue Act of 1964 | Yes | 145 | 0.002 | 0.025 | 0.331 | 0.011 | 0.572 | 0.899 |
| *Remarks at the National Conference of the Business Committee for Tax Reduction* | | | | | | | | | |
| 1963/09/10 | Changes in Depreciation Guidelines and Revenue Act of 1962 | Yes | 36 | 0.004 | 0.078 | 0.590 | 0.012 | 0.452 | 0.847 |
| *Remarks Upon Announcing Plans to Recommend a Reduction in Excise Taxes* | | | | | | | | | |
| 1965/05/15 | Excise Tax Reduction Act of 1965 | No | 14 | 0.007 | 0.132 | 0.571 | 0.010 | 0.252 | 0.647 |
| *Special Message to the Congress Recommending Reduction of Excise Taxes and Increases in User Charges* | | | | | | | | | |
| 1965/05/17 | Excise Tax Reduction Act of 1965 | No | 112 | 0.003 | 0.048 | 0.579 | 0.011 | 0.311 | 0.729 |
| *Special Message to the Congress Recommending Reinstatement of the Investment Tax Credit and Accelerated Depreciation Investment Incentives* | | | | | | | | | |
| 1967/03/09 | Public Law 90-26 | Yes | 33 | 0.008 | 0.090 | 0.546 | 0.007 | 0.202 | 0.668 |
| *Special Message to the Congress on Reform of the Federal Tax System* | | | | | | | | | |
| 1969/04/21 | Tax Reform Act of 1969 | No | 38 | 0.004 | 0.044 | 0.434 | 0.022 | 0.400 | 0.786 |
| *Statement Announcing Changes in Administration of the Depreciation Provisions of the Tax Laws* | | | | | | | | | |
| 1971/01/11 | Reform of Depreciation Rules | No | 17 | 0.005 | 0.067 | 0.337 | 0.012 | 0.300 | 0.801 |
| *Annual Budget Message to the Congress* | | | | | | | | | |
| 1971/01/29 | Reform of Depreciation Rules | No | 224 | 0.003 | 0.023 | 0.524 | 0.005 | 0.140 | 0.794 |
| *Challenge of Peace* | | | | | | | | | |
| 1971/08/15 | Revenue Act of 1971 | No | 56 | 0.006 | 0.090 | 0.395 | 0.004 | 0.125 | 0.652 |
| *Address to the Nation on Energy and Economic Programs* | | | | | | | | | |
| 1975/01/13 | Tax Reduction Act of 1975 | Yes | 41 | 0.005 | 0.062 | 0.499 | 0.009 | 0.203 | 0.818 |
| *Address Before a Joint Session of the Congress Reporting on the State of the Union* | | | | | | | | | |
| 1975/01/15 | Tax Reduction Act of 1975 | Yes | 91 | 0.003 | 0.040 | 0.622 | 0.005 | 0.188 | 0.748 |
| *Annual Budget Message to the Congress* | | | | | | | | | |
| 1975/02/03 | Tax Reduction Act of 1975 | Yes | 89 | 0.002 | 0.033 | 0.449 | 0.003 | 0.319 | 0.875 |
| *Address to the Nation on Federal Tax and Spending Reduction* | | | | | | | | | |
| 1975/10/06 | Revenue Adjustment Act of 1975 | No | 28 | 0.003 | 0.081 | 0.439 | 0.013 | 0.337 | 0.766 |
| *Annual Message to Congress* | | | | | | | | | |
| 1976/01/26 | Tax Reform Act of 1976 | No | 24 | 0.002 | 0.130 | 0.737 | 0.002 | 0.168 | 0.687 |
| *Economic Recovery Program—Message to the Congress* | | | | | | | | | |
| 1977/01/31 | Tax Reduction and Simplification Act of 1977 | Yes | 34 | 0.007 | 0.056 | 0.458 | 0.022 | 0.334 | 0.711 |
| *Fiscal Year 1978 Budget Revisions Message to the Congress Transmitting the Revisions* | | | | | | | | | |
| 1977/02/22 | Tax Reduction and Simplification Act of 1977 | Yes | 13 | 0.004 | 0.027 | 0.126 | 0.013 | 0.286 | 0.696 |
| *State of the Union address* | | | | | | | | | |
| 1978/01/19 | Revenue Act of 1978 | No | 305 | 0.003 | 0.028 | 0.616 | 0.002 | 0.085 | 0.695 |

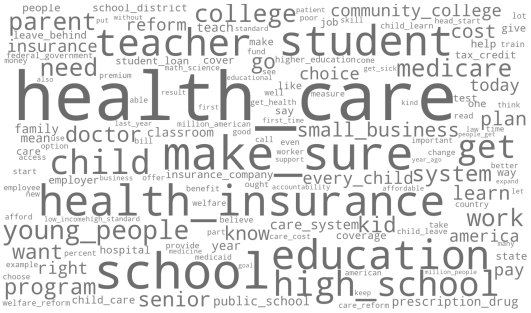
| Date | Announced legislation | Mixed | No. paragraphs | *Tax increase* proportion Min. | Av. | Max. | *Tax decrease* proportion Min. | Av. | Max. |
|---|---|---|---|---|---|---|---|---|---|
| | **Title of the speech** | | | | | | | | |
| *Tax Reduction and Reform Message to the Congress* | | | | | | | | | |
| 1978/01/20 | Revenue Act of 1978 | No | 204 | 0.003 | 0.042 | 0.533 | 0.018 | 0.432 | 0.870 |
| *Budget Message to the Congress Transmitting the Fiscal Year 1979 Budget* | | | | | | | | | |
| 1978/01/20 | Revenue Act of 1978 | No | 35 | 0.003 | 0.025 | 0.168 | 0.004 | 0.187 | 0.751 |
| *Inaugural Address* | | | | | | | | | |
| 1981/01/20 | Economic Recovery Tax Act of 1981 | Yes | 35 | 0.004 | 0.022 | 0.336 | 0.004 | 0.054 | 0.471 |
| *Address to the Nation on the Economy* | | | | | | | | | |
| 1981/02/05 | Economic Recovery Tax Act of 1981 | Yes | 50 | 0.004 | 0.080 | 0.624 | 0.008 | 0.349 | 0.775 |
| *Address before a Joint Session of the Congress on the Program for Economic Recovery* | | | | | | | | | |
| 1981/02/18 | Economic Recovery Tax Act of 1981 | Yes | 63 | 0.004 | 0.082 | 0.645 | 0.005 | 0.329 | 0.778 |
| *Address Before a Joint Session of the Congress on the Program for Economic Recovery* | | | | | | | | | |
| 1981/04/28 | Economic Recovery Tax Act of 1981 | Yes | 40 | 0.004 | 0.053 | 0.762 | 0.005 | 0.201 | 0.820 |
| *Address to the Nation on Tax Reform* | | | | | | | | | |
| 1985/05/28 | Tax Reform Act of 1986 | Yes | 53 | 0.003 | 0.057 | 0.402 | 0.005 | 0.360 | 0.818 |
| *Address Before a Joint Session of Congress on the State of the Union* | | | | | | | | | |
| 1986/02/04 | Tax Reform Act of 1986 | Yes | 34 | 0.002 | 0.028 | 0.451 | 0.003 | 0.115 | 0.750 |
| *Radio Address to the Nation on Tax Reform* | | | | | | | | | |
| 1986/05/10 | Tax Reform Act of 1986 | Yes | 7 | 0.006 | 0.051 | 0.205 | 0.015 | 0.376 | 0.738 |
| *The President's Radio Address* | | | | | | | | | |
| 1997/02/22 | Taxpayer Relief Act of 1997 and Balanced Budget Act of 1997 | Yes | 8 | 0.003 | 0.119 | 0.530 | 0.009 | 0.354 | 0.865 |
| *Remarks on Departure for Boston* | | | | | | | | | |
| 1997/06/30 | Taxpayer Relief Act of 1997 and Balanced Budget Act of 1997 | Yes | 40 | 0.005 | 0.067 | 0.487 | 0.003 | 0.236 | 0.760 |
| *A press conference* | | | | | | | | | |
| 2001/02/05 | Economic Growth and Tax Relief Reconciliation Act of 2001 | No | 20 | 0.009 | 0.043 | 0.188 | 0.008 | 0.397 | 0.740 |
| *The President's Agenda for Tax Relief* | | | | | | | | | |
| 2001/02/08 | Economic Growth and Tax Relief Reconciliation Act of 2001 | No | 1 | 0.129 | 0.129 | 0.129 | 0.207 | 0.207 | 0.207 |
| *The President's Radio Address* | | | | | | | | | |
| 2001/03/17 | Economic Growth and Tax Relief Reconciliation Act of 2001 | No | 10 | 0.010 | 0.169 | 0.615 | 0.083 | 0.303 | 0.738 |
| *Remarks to Business* | | | | | | | | | |
| 2001/10/26 | Job Creation and Worker Assistance Act of 2002 | Yes | 43 | 0.005 | 0.054 | 0.426 | 0.003 | 0.084 | 0.745 |
| *State of the Union address* | | | | | | | | | |
| 2002/01/29 | Job Creation and Worker Assistance Act of 2002 | Yes | 64 | 0.003 | 0.034 | 0.689 | 0.003 | 0.051 | 0.680 |
| *Remarks to the Economic Club of Chicago in Chicago* | | | | | | | | | |
| 2003/01/07 | Jobs and Growth Tax Relief Reconciliation Act of 2003 | Yes | 52 | 0.003 | 0.114 | 0.702 | 0.004 | 0.223 | 0.815 |
| *The President's Radio Address* | | | | | | | | | |
| 2003/01/11 | Jobs and Growth Tax Relief Reconciliation Act of 2003 | Yes | 11 | 0.008 | 0.086 | 0.396 | 0.011 | 0.226 | 0.795 |

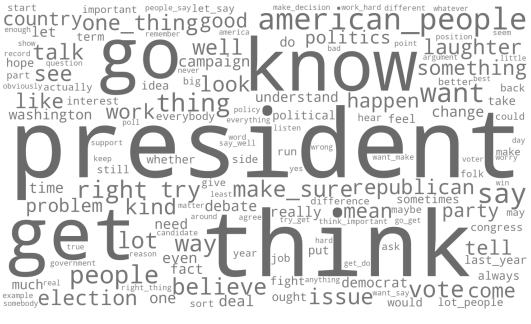
| Date | Announced legislation | Title of the speech: Mixed | Title of the speech: No. paragraphs | *Tax increase* proportion: Min. | *Tax increase* proportion: Av. | *Tax increase* proportion: Max. | *Tax decrease* proportion: Min. | *Tax decrease* proportion: Av. | *Tax decrease* proportion: Max. |
|---|---|---|---|---|---|---|---|---|---|
| *Remarks to the Tax Relief Coalition* | | | | | | | | | |
| 2003/05/06 | Jobs and Growth Tax Relief Reconciliation Act of 2003 | Yes | 47 | 0.004 | 0.090 | 0.748 | 0.004 | 0.194 | 0.839 |
| *State of the Union address* | | | | | | | | | |
| 2004/01/20 | Working Families Tax Relief Act of 2004 | No | 68 | 0.004 | 0.050 | 0.728 | 0.004 | 0.058 | 0.747 |
| *Remarks on the national economy* | | | | | | | | | |
| 2008/01/18 | Economic Stimulus Act of 2008 | No | 14 | 0.004 | 0.094 | 0.563 | 0.007 | 0.182 | 0.824 |

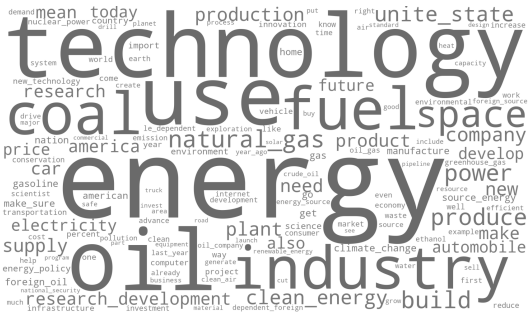
Table D.2: List of speeches announcing legislation whose overall effect was a decrease in tax revenue. Mixed legislation includes both tax hikes and cuts.

## D.2 Wordclouds - composition of the topics

Figure D.1 shows the 26 topic distributions estimated in the second step of our LDA approach in the form of wordclouds. For each topic 150 terms with the highest probability of being used are shown. The size of a term indicates its relative probability. For each topic we provide an interpretation which we believe best captures its likely usage. As expected the majority of estimated topics relate to particular issues in U.S. politics. It is important to note however that this is not the case for all of them. Topics are based on terms that tend to co-occur in the documents, and that co-occurrence can be caused not only by the what issues are discussed. For example, our data-set includes interviews with the president, and the particular language used in those seems to "picked up" by topic shown in Fig. D.1o. Another example is the topic which we call *Informal Speech* shown in Fig. D.1l. The fact that it heavily features the term "laughter" used in the transcripts to indicate when the president or the audience is laughing - suggests that it concerns the informal part of presidential speeches. Other terms however do not seem to be connected by any particular theme, and as such, it might be an artifact of our pre-processing approach. In particular, terms that show up frequently regardless of the actual content might co-occur "naturally" and be grouped into a common topic.







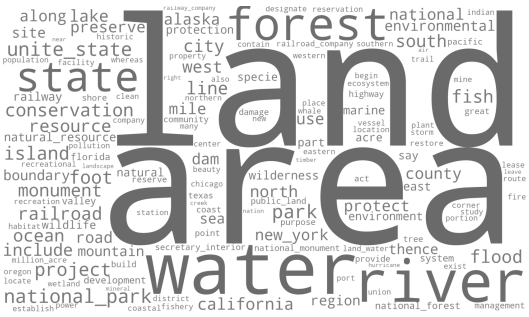
(a) Social Services (b) Legislative Process

(c) Persuasive Rhetoric (d) Public Administration

(e) Natural Resources, Energy & Technology (f) Health Care

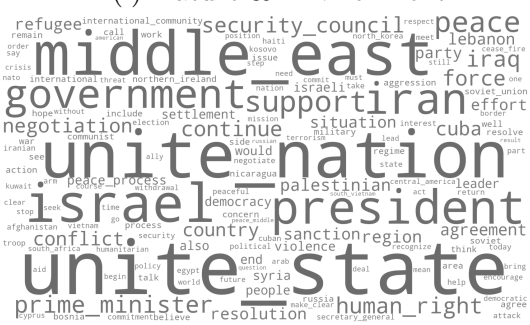
(g) Proclamations of Unity & Patriotism (h) Cold War

Figure D.1: Wordclouds of the non-tax topics

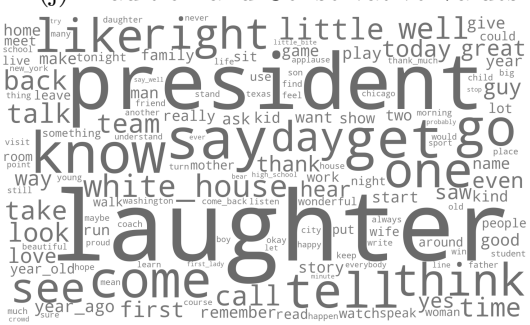



(i) Nature & Environment


(j) Tradition and Conservative Values

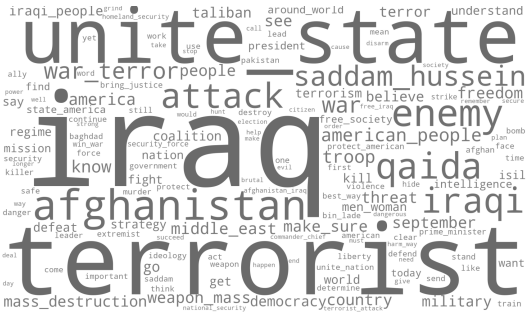



(k) Middle East




(l) Informal Speech


(m) Executive Orders

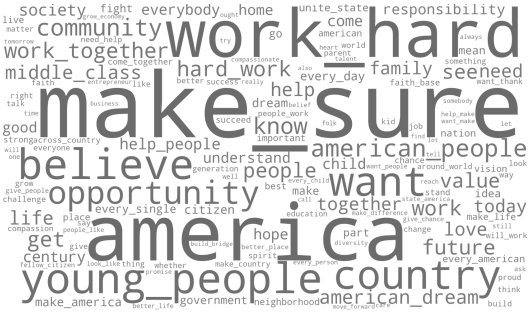



(n) War on Terror


(o) Interviews


(p) Crime

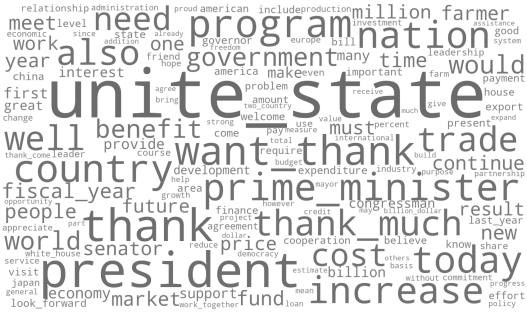
Figure D.1: Wordclouds of the non-tax topics



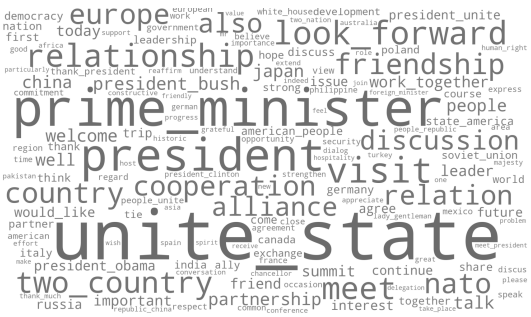

(q) American Values


(r) Paying Respect to Armed Forces


(s) National Security & Emergencies

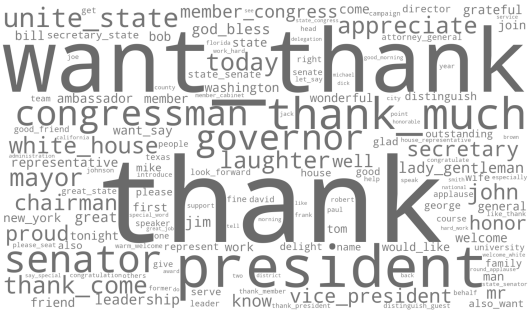



(t) Economic Development




(u) International Relations/Diplomacy


(v) Freedom & Patriotism




(w) Expressing Gratitude in Speeches


(x) International Economics & Trade

Figure D.1: Wordclouds of the non-tax topics