European Economic Review 168 (2024) 104827

Contents lists available at ScienceDirect

European Economic Review

journal homepage: www.elsevier.com/locate/eer

Fiscal policy in the Bundestag: Textual analysis and macroeconomic effects ☆

Albina Latifi, Viktoriia Naboka-Krell, Peter Tillmann *, Peter Winker

Department of Economics, Justus Liebig University Giessen, Germany

ARTICLE INFO

Dataset link: Replication package (Original data)

JEL classification:
C89
E60
E62

Keywords:
Text mining
Word embeddings
VAR models
Identification
Government spending

ABSTRACT

Fiscal policy is made in parliaments. We go to the roots of changes of fiscal policy in Germany and use a novel data set on all parliamentary speeches in the Bundestag from 1960 to 2021. We propose an embedding-based approach, which allows the representation of words and documents in a shared vector space, in order to measure fiscal policy-related sentiment in parliamentary debates at a scale from contractionary to expansionary. We also distinguish between sentiment related to exogenous and endogenous fiscal policy. We put fiscal sentiment into a series of recursively-identified vector autoregressive models to show that a change in fiscal sentiment causes a shift in government spending and has significant effects on the macroeconomy. The results support the notion that the debate in parliament contains information for the identification of government spending shocks.

1. Introduction

Fiscal policy is made in parliaments. Legislating an increase or cut of federal public spending or taxes is a key prerogative of parliaments. This implies that changes in government spending are usually preceded by extensive debates in parliament and beyond, often stretching several quarters or even years. Hence, shifts in the tone of the parliamentary debate should indicate changes in government spending further down the road.

The empirical literature on the identification of government spending shocks and the estimation of their effects on the macroeconomy, see Ramey (2016) for a recent survey, has not yet made use of this information. In this paper, we use data on all speeches delivered in the Bundestag, the federal parliament of Germany, in order to measure fiscal policy. We argue that the measurement and the identification of fiscal policy impulses can be improved by examining the roots of fiscal policy-making, namely the parliamentary process itself. We believe that households and firms monitor the parliamentary process and adjust their expectations and decisions well before the law is eventually passed and comes into effect.

☆ We are grateful to the editor and the associate editor of this journal as well as anonymous reviewers for excellent feedback on the first version of the paper. We also thank Leif Anders Thorsrud and seminar participants at the COMPSTAT (Bologna, 2022), MACROMODELS (Wieliczka, 2022) and COMPTEXT (Glasgow, 2023) conferences as well as the 1st NEAR Conference on Narrative Economics (Bochum, 2024) and the annual meetings of the EEA (Barcelona, 2023) and the Verein für Socialpolitik (Regensburg, 2023). Bernd Hayo kindly shared his data series. Financial support from the German Research Foundation (DFG) (TI 594/4-1 and WI 2024/8-1) for the project MaFiText is gratefully acknowledged. The project also benefited from cooperation within HiTEC Cost Action CA 21163.

* Corresponding author.
E-mail addresses: albina.latifi@wirtschaft.uni-giessen.de (A. Latifi), viktoriia.naboka@wirtschaft.uni-giessen.de (V. Naboka-Krell), peter.tillmann@wirtschaft.uni-giessen.de (P. Tillmann), peter.winker@wirtschaft.uni-giessen.de (P. Winker).

https://doi.org/10.1016/j.euroecorev.2024.104827
Received 29 March 2023; Received in revised form 7 August 2024; Accepted 8 August 2024
Available online 10 August 2024
0014-2921/

As a matter of fact, parliamentary speeches are multi-dimensional objects. Extracting quantitative information about fiscal policy is not straightforward. We exploit recent advances in natural language processing (NLP) to quantify fiscal sentiment based on a large data set of parliamentary speeches.

We proceed as follows: First, we propose an embedding-based approach, which allows the representation of words and documents in a shared vector space, in order to measure fiscal policy-related sentiment in parliamentary debates at a scale from contractionary to expansionary. For this purpose, we create a dictionary containing terms related to expansionary and contractionary fiscal policy measures. Specifically, we adopt Doc2Vec, an unsupervised method to represent natural language in a high-dimensional vector space (Mikolov et al., 2013a). The resulting text vectors (also known as embeddings) capture semantic characteristics of the texts. As the context in which fiscal policy measures are discussed may change over time, we adapt the approach proposed by Kapfhammer et al. (2020) and implement a rolling forecast architecture. This provides us with three series of fiscal sentiment: for the entire Bundestag, the governing parties and the opposition. We find large fluctuations in fiscal sentiment that fit the established historical narrative.

To the best of our knowledge, we are the first to use a large body of parliamentary texts to measure shifts in fiscal sentiment and, eventually, to estimate its effect on the macroeconomy.[1] Abercrombie and Batista-Navarro (2020) provide a comprehensive literature review of 61 studies, all of which deal with the automatic analysis of sentiments and opinions as well as the positions of speakers in parliamentary debates. In their research outlook, the authors regret that most of the studies only perform a rough positional analysis (e.g. left vs. right), instead of identifying policy preferences, which constrains possible applications. Furthermore, almost all studies in Abercrombie and Batista-Navarro (2020) are limited to the analysis of a single election period. This makes this paper one of the few existing studies that deal with the automated identification of political preferences over several election periods.

A recent strand of the macroeconomic literature proposes a "narrative approach" to the identification of exogenous changes to fiscal policy, i.e. Romer and Romer (2010), Mertens and Ravn (2012) and Cloyne (2013). These authors derive tax policy shocks from text documents such as presidential speeches or parliamentary reports.[2] Guajardo et al. (2014) use historical records such as budget speeches, central bank writings, IMF staff reports and OECD documents to identify changes in fiscal policy designed to consolidate public finances. Our paper goes beyond the selected numbers of text documents used in these studies. Instead, we use data on *all* parliamentary speeches to derive a measure of fiscal sentiment.

In a second step, we follow the notion of Romer and Romer (2010) and Cloyne (2013) and distinguish between parliamentary contributions on fiscal-policy related questions that are either exogenous or endogenous with respect to the current economic situation. We apply an Latent Dirichlet Allocation (LDA) model, a probabilistic topic model, which allows us to link speeches to latent topics (Blei et al., 2003). These topics can be grouped into topics related to exogenous fiscal policy (e.g. national defense, energy policy and social welfare), endogenous fiscal policy responses (e.g. the labor market, economic growth and public investment) or policies not immediately fiscally relevant. This provides us with series of exogenous and endogenous fiscal sentiment, respectively.

Third, we put the resulting sentiment series in a battery of standard Bayesian vector autoregressive (VAR) models. The models also include variables such as real government spending, real GDP, real private consumption or real investment. Our aim is to evaluate whether fiscal sentiment causes government expenditure and, hence, macroeconomic responses. Following the pioneering work of Blanchard and Perotti (2002), a large literature uses a recursive identification scheme in order to estimate the causal effects of government spending. This draws on the notion that within a quarter government spending is predetermined such that a feedback from GDP or other macro aggregates on the level of government spending should be excluded. Fortunately, our baseline series of fiscal sentiment lends itself to a straightforward extension of this identification scheme: we order sentiment last such that a change in sentiment cannot contemporaneously drive government spending. Sentiment, on the other hand, can immediately respond to economic developments. The decomposition of sentiment into exogenous and endogenous fiscal sentiment gives rise to an extension of this identification: exogenous fiscal sentiment is ordered first, macroeconomic reactions follow, while endogenous fiscal sentiment comes last.

We find that an unexpected increase in sentiment towards a more expansionary fiscal stance causes higher government spending, an expansion of real economic activity and an increase in private consumption. Several extensions of the model show that fiscal sentiment also increases investment and inflation and reduces unemployment and the budget balance, among other variables. These responses are in line with standard (New-)Keynesian business cycle models. Furthermore, fiscal sentiment has consequences for Germany as an open economy: more expansionary sentiment leads to a real appreciation and a deterioration of the trade balance. Hence, sentiment measured from parliamentary speeches has strong and robust macroeconomic effects. These results are mostly due to shifts in exogenous fiscal sentiment, while shocks to endogenous sentiment does not result in significant macroeconomic responses.

We also contribute to the literature on fiscal news shocks (Ramey, 2011a; Ricco et al., 2016; Ben Zeev and Pappa, 2017), which argues that fiscal shocks are to some extent anticipated before they materialize. Ramey (2011a) constructs two measures of news, one from a narrative account of reports about defense spending and one from professional forecasters, and shows that both are able to predict spending shocks from recursively identified VAR models. Ricco et al. (2016) also uses data from professional forecasters to measure the future path of government spending. Ben Zeev and Pappa (2017) identify a news shock as the shock that best explains future movements in defense spending, while at the same time being orthogonal to current defense spending. By detecting changes

[1] Allard et al. (2013) and Dybowski and Adämmer (2018) quantify fiscal sentiment in central bank documents and the communication of U.S. presidents. Instead, our paper is based on a much larger text corpus.

[2] See Hayo and Uhl (2014), Hayo and Mierzwa (2022) and Christofzik et al. (2022) for similar approaches to tax policy shocks in Germany.

in fiscal sentiment in speeches of parliamentarians that precede actual legislation, we offer an alternative way to shed light on fiscal news.

Equipped with our sentiment series, we revisit the problem of fiscal foresight, which is often put forward as an argument to invalidate a recursive identification of government spending shocks, e.g. Ramey (2011a, 2016) and Ellahie and Ricco (2017). From a standard recursively-identified VAR model in the spirit of Blanchard and Perotti (2002), Fatás and Mihov (2001), Galí et al. (2007), Born and Müller (2012), Auerbach and Gorodnichenko (2012) and Ilzetzki et al. (2013) and others we obtain a series of structural government spending shocks. We then show that fiscal sentiment predicts these structural shocks six to eight quarters in advance. Hence, supposedly unanticipated government spending shocks are in fact anticipated once information from the parliamentary debate is taken into account.

The remainder of this paper is structured as follows. Section 2 describes the data from the German Bundestag as well as the preprocessing steps. Section 3 presents our approach for constructing the text-based fiscal sentiment indicator. In Section 4, we estimate a range of VAR models to understand the effects of fiscal policy on the German economy. Section 5 revisits the problem of fiscal foresight. Finally, Section 6 concludes. An online appendix contains additional material.

2. Text data

This section introduces the textual data from which we derive a measure of fiscal sentiment. Section 2.1 describes the underlying text data, i.e. the full set of parliamentary speeches. In Section 2.2, we describe specific text data preprocessing steps which are needed for the text mining methods applied later.

2.1. Bundestag speeches as a novel data source

The Bundestag is the German federal parliament. At the time of writing, it has more than 700 members with the exact number varying over time. Although we later focus on fiscal policy, we start from the full set of all parliamentary debates. The digitized debates of the plenary sessions, not the committees, are made available to the public by the German Bundestag.[3]

Unfortunately, this text data is not directly suitable for the application of NLP methods as the parliamentary speeches up to election period 18, i.e. until October 2017, are available in an unstructured form only. This means that the XML documents only have one "TEXT" tag in addition to some meta-information on the whole session such as the election period and the date. The content of the "TEXT" tag comprises the entire stenographic report — including the agenda items, the actual meeting and the annexes. Therefore, it is necessary to further structure the content of the "TEXT" tag so that each speech can be assigned to a speaker. Furthermore, each speaker should be assigned with his or her role and party affiliation. To structure the XML files, we use the workflow documented in more detail in Latifi (2023), which proceeds as follows:

First, we parse and clean the XML documents using a set of regular expressions. Second, a Named Entity Recognition (NER) model with a customized entity, which usually consists of a speaker's first and last name followed by his or her party affiliation in brackets, a colon and a newline, is developed to identify the begin of each speech. Thereby, a small hand-labeled data set is created to train the NER model. Eventually, the cleaned stenographic protocols can be split by each identified beginning of a speech. After that, one can extract roles and party affiliations of each speaker.

With the aim of promoting computer-assisted research on parliamentary data, the German Bundestag has been publishing the XML files in a structured form since election period 19, so that the precise extraction of relevant information involves comparatively little effort and manual reworking.[4] We convert these plenary protocols into a file format that can be used for further processing by using the python package `pybundestag` (Hruzik, 2019). In a final step, we unify the data set of election periods 1–18 with the data set of election period 19 in a consistent structure. The complete data set of the election periods 1 to 19 comprises a total of 877 140 speeches.

2.2. Corpus and text preprocessing

Since most macroeconomic time series are available from 1970 onward and as we need a rolling 10-year training data set for our embedding approach, our data set begins in 1960. In a first step, speeches by the President of the Bundestag or an office holder of a similar function are excluded, as their main task is primarily to chair and moderate (including announcing voting results, calling up items on the agenda, calling up speakers) the plenary sessions. Furthermore, we exclude speeches from state ministers representing the federal states of Germany, guest speakers such as foreign dignitaries and other irregular speakers.

This leaves us with all speeches delivered by members of parliament, Chancellors, Federal Ministers and State Secretaries.[5] Furthermore, we remove very short and very long speeches from the data set. Short remarks are often made during swearing-in ceremonies or as interposed questions. We count the words of each speech and remove all speeches containing less than 100 tokens, i.e words, or more than 3573 tokens. The latter threshold corresponds to the 99.5%-percentile of the word frequency distribution

[3] All stenographic reports can be downloaded from this website as XML files packed in .zip files: https://www.bundestag.de/services/opendata.

[4] Further information on the structure of the XML documents as of election period 19 are described here: https://www.bundestag.de/resource/blob/577234/f9159cee3e045cbc37dcd6de6322fcdd/dbtplenarprotokoll_kommentiert-data.pdf.

[5] Members of the government do not have to be members of parliament, though in most cases they are.

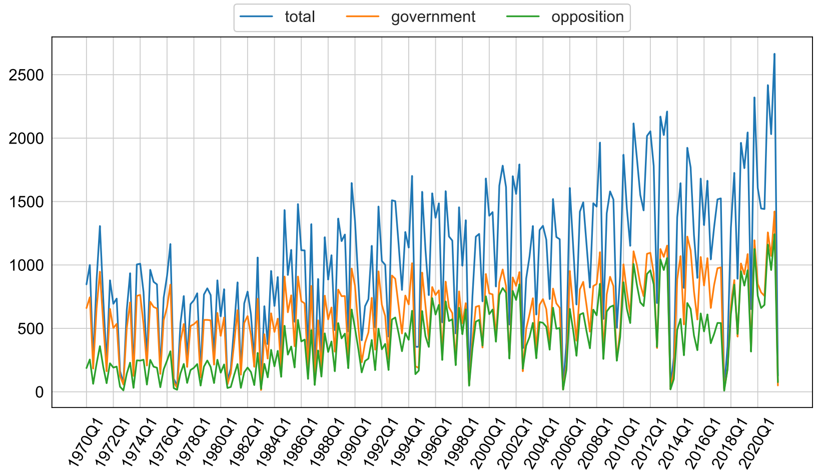

Fig. 1. Number of speeches. *Notes:* The figure shows the number of speeches over our sample period aggregated to quarterly frequency.

over the documents. After these exclusion steps, the data set contains 235 129 speeches covering the period from January 20, 1960 to September 07, 2021.

Fig. 1 shows the number of speeches on a quarterly basis. Over the entire period, the average number of speeches per quarter is 959.71, though the number of speeches increases over time. Furthermore, there are seasonal fluctuations. The third quarter contains by far the fewest speeches in the data set, with a total of 26 141 speeches, which can be attributed to the summer break.[6] In addition, Fig. 1 shows that members of the coalition parties forming the government deliver more speeches than members of the opposition parties. This is plausible given the distribution of seats. However, these differences have become smaller in the recent past.

We then prepare the corpus using common text preprocessing steps, such as those described in Grimmer and Stewart (2013) and Denny and Spirling (2018). We first lemmatize all speeches using the model *"de_core_news_lg"* from the python package `spaCy` (Honnibal et al., 2020) so that all words in the speeches are traced back to their root words. This also reduces the complexity of the corpus. In the following, we refer to a word as a token and a speech as a document. The corpus is the collection of all documents.

After lemmatizing, we convert all German umlauts and the eszett to exclude encoding errors. In addition, we remove line breaks, digits, blank sentences and special characters and convert all letters to lower case. We remove single-element tokens and tokens with more than 30 elements. The removal of further short tokens is not advisable for this domain, because meaningful tokens such as "is", "eg", "eu", "ki", "db" etc. are included in this corpus.

The next step is to create a list of stop words. Stop words are tokens that occur very frequently in the corpus, but contain little information and therefore do not contribute to the understanding of a text (e.g. personal pronouns, conjunctions, etc.). In the `nltk` package (Bird and Klein, 2009), there is a predefined list of stop words for the German language. Since a general stop words list does not include domain-specific terms such as "Bundestag", "Abgeordneter", "Deutschland", "Redezeit", "Drucksache", etc., we create a list of stop words based on the inverse document frequency (idf) value of each unique word.[7] A low idf value implies that a word occurs in a very large number of documents and is therefore not very specific. All tokens that occur in 97% of all documents in each election period are considered potential stop words. The final list of stop words is manually expanded in an iterative process, so that it comprises a total of 1405 terms. We report the final list of stop words in the appendix. These stop words were lemmatized analogously to the text and finally removed from the corpus.

Fig. 2 shows the distribution of the length of a document before and after applying the described preprocessing steps. A document contains an average of 637.10 tokens before preprocessing, but only 164.83 tokens after preprocessing. The entire corpus contains 149 800 737 tokens before and 38 755 290 tokens after preprocessing. This means that the size of the corpus is reduced by almost 75% after preprocessing.

3. A text-based fiscal sentiment indicator

This section first documents the construction of a dictionary with fiscal policy-specific terms. Next, we introduce word and document embeddings. Finally, we propose an approach to construct a fiscal sentiment indicator based on these text representations.

[6] For comparison, quarter I contains a total of 68 553 speeches, quarter II contains 73 412 speeches, and quarter IV contains 67 023 speeches.

[7] The idf value is calculated using the package `scikit-learn` (Pedregosa et al., 2011).

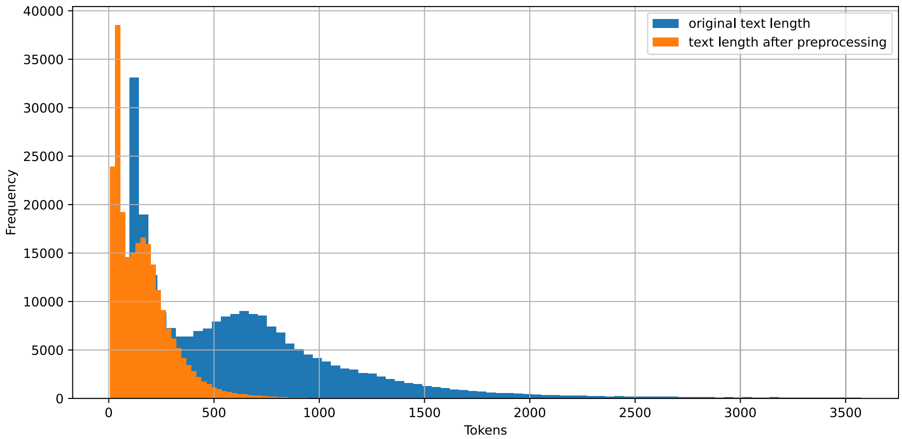

Fig. 2. Length of speeches before and after preprocessing. *Notes:* The figure shows the distribution of the length of a document (speech) before and after preprocessing.

3.1. Compiling a dictionary on fiscal policy

Although dictionary-based approaches have been widely used in economics and finance lately, e.g. to quantify the sentiment of financial reports (Loughran and McDonald, 2011), central bank communication (Picault and Renault, 2017) or to construct newspaper-based indicators of economic policy uncertainty (Baker et al., 2016), there is no dictionary available for fiscal policy. Therefore, we first need to assemble a fiscal policy-related dictionary for Germany.

We construct a list of terms relevant to fiscal policy through extensive study of the speeches in the Bundestag. Since not all randomly selected speeches address the core of fiscal policy, we also take advantage of data from the Manifesto corpus (Burst et al., 2020), in order to identify particularly informative expressions. This data contains publicly available, thematically classified quasi-sentences from election manifestos of various parties from over 50 countries. The categories capture the most relevant political issues and goals and are assigned to the respective quasi-sentences according to a strict annotation scheme. The annotation scheme is described in Werner et al. (2011). The corpus can be downloaded via the R package `manifestoR` or via the website.[8] We select 19 categories related to fiscal policy.[9]

In this way, we construct a preliminary list of 163 keywords. After that, we expand and refine the keywords, considering also the characteristics specific to the German language, such as composite words and synonyms. After this extension, the fiscal dictionary comprises a total of 322 words. We then label terms as expansionary, neutral or contractionary. It is important that we identify terms as uniquely as possible in terms of the expansionary or contractionary sentiment conveyed. In case an expression is ambiguous, we decide based on the majority vote among the authors. Eventually, our list consists of 218 terms or compound terms, of which 122 are classified as "expansionary" and 96 as "contractionary".[10]

3.2. Doc2Vec approach

Doc2Vec is an extension of Word2Vec, which is an unsupervised, neural network-based method to represent natural language in a high-dimensional vector space (Mikolov et al., 2013a). It could be considered a black box that "translates" semantic features of natural language into dimensions of a vector space. Almost no supervision is needed to train such text representations as the algorithm learns semantics from original texts. One of the most characteristic features of resulting embeddings is the interpretability of mathematical operations between them. Let us consider one specific example in the context of fiscal policy. Assume that we obtained 100-dimensional vectors for the words "staatliche" (government) and "Investitionen" (investments) each. Then, we sum up these two vectors and look for the nearest neighbors to this sum (based on cosine similarity). We obtain the following words: *oeffentlich (public)*, *investieren (invest)*, *privat (private)*, *Aufbau (development)*, *gesetzlich (legal)*, *Innovation (innovation)*, *nachhaltig (sustainable)*, *langfristig (long-term)*. A further example considers the word "Steuerentlastung" (tax relief). Its nearest neighbors are the following: *Steuererhoehung (tax increase)*, *Entlastung (relief)*, *Progression (progression)*, *Muetterrente (mothers' pension)*, *Soli (solidarity*

[8] The .csv files can be downloaded via this page: https://visuals.manifesto-project.wzb.eu/mpdb-shiny/cmp_dashboard_dataset/.

[9] This amounts to 39 .csv files of German election programs from 1998 to 2017. These files contain a total of 22 602 quasi-sentences classified with the selected categories. The following categories are selected: [303] Governmental and Administrative Efficiency, [401] Free Market Economy, [402] Incentives, [403] Market Regulation, [404] Economic Planning, [408] Economic Goals, [409] Keynesian Demand Management, [410] Economic Growth: Positive, [412] Controlled Economy, [414] Economic Orthodoxy, [416] Anti-Growth Economy: Positive, [503] Equality: Positive, [504] Welfare State Expansion, [505] Welfare State Limitation, [501] Environmental Protection: Positive, [701] Labour Groups: Positive, [702] Labour Groups: Negative, [416.1] Anti-Growth Economy: Positive, [4012] Control of Economy: Negative.

[10] The list of expansionary and contractionary terms used can be found in the appendix.

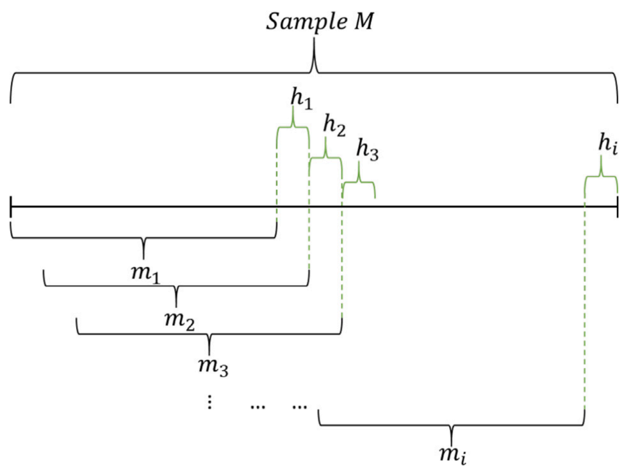

Fig. 3. Rolling-window Doc2Vec. *Notes*: This figure represents the training procedure of the dynamic Doc2Vec approach. Each training period m covers the same period of time (e.g. quarter, year, 10 years). Each forecast period h is also of fixed size. The shift length equals the defined forecast length.

tax), *Kinderfreibetrag (allowance for children)*, *Elterngeld (parental allowance)*, *Neuverschuldung (new debt)*, *Steuervereinfachung (tax simplification)* etc. These examples demonstrate the ability of embeddings to capture semantic similarity of natural language. The exemplary words and their neighbors are indeed likely to appear in the same context.

Doc2Vec allows representing sentences, paragraphs and whole documents as vectors (Le and Mikolov, 2014; Mikolov et al., 2013b). The distances between these representations could be interpreted meaningfully. In the appendix, we graphically illustrate the approach based on a selected speech of Olaf Scholz, the Federal Minister of Finance, on March 25, 2020.[11]

We propose a Doc2Vec approach to construct a fiscal policy index for the German Bundestag. Thereby, we require that values of the index at time t are based solely on information up to time t. Furthermore, the model should allow for possible changes in language usage over time. New concepts with regard to fiscal policy may occur and some of them may disappear over time. Also, the general context in which fiscal policy measures are discussed may change over time.

To address the first feature described above, we propose a rolling forecast architecture. Fig. 3 represents the training procedure for the rolling window setting. Given the sample M that corresponds to the corpus presented in Section 2.2, we divide the period into training periods m_i. For each period m_i, we train a Doc2Vec model using lemmatized and preprocessed texts. For the subsequent period h_i, document vectors are inferred based on the trained Doc2Vec model. This means that documents in the forecast period h are actually new to the model. The vectors for these documents are predicted based on the trained word dependencies and relationships. Each training and forecast period is of fixed size. However, the number of observations in each period m_i and h_i might differ depending on the number of documents.

To address the second feature, we adapt the approach proposed by Kapfhammer et al. (2020). The authors use word embeddings to measure climate change transition risk and investigate how the media speaks about risk and how the context changes over time. To address the changing context, the authors divide the time period into sub-periods and estimate separate Word2Vec models for each sub-period. Kapfhammer et al. (2020) argue that making the word embedding methodology dynamic can capture changes of the relationships between specific words over time.

Overall, the dynamic approach proceeds in four steps:

1. For a defined period of time, train the Doc2Vec model using preprocessed texts.
2. For the given training sub-period, construct an expansionary and a contractionary vector as the average vector of the identified fiscal policy-related terms. Thereby, their representation is different in each training period depending on which words occur in the learned vocabulary. For two-words terms such as "Steuern senken", "Arbeitsplätze schaffen" and others, the average of the corresponding word vectors is used to represent these terms.
3. Based on the pre-trained model, infer speeches vectors for the subsequent forecast period. Calculate the cosine similarities between the inferred document vectors and the fiscal policy vectors. Each speech receives two scores: similarity to an expansionary stance of fiscal policy and similarity to a contractionary stance of fiscal policy.
4. Construct a continuous indicator by taking the difference between the similarity to the expansionary vector and the contractionary vector. This results in a value that falls into the range from −1 (very contractionary) to 1 (very expansionary).

This procedure is fully unsupervised and language agnostic. It can be applied to other corpora and other languages in case a fiscal policy dictionary is available. Additional technical details are described in the appendix.

[11] Due to described features of word and text vectors, these are used in different text-as-data applications, e.g. Rheault and Cochrane (2020), Gennaro and Ash (2021) and Rodriguez and Spirling (2022).

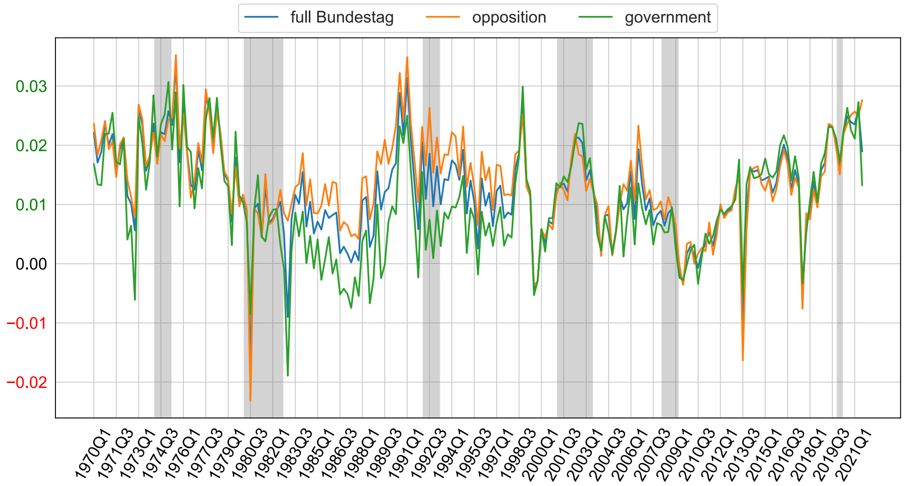

Fig. 4. Fiscal policy sentiment. *Notes:* The figure shows the fiscal sentiment of the entire Bundestag as well as the government and the opposition. The shaded areas highlight recessions identified by the German Council of Economic Experts.

3.3. Baseline sentiment

We use a training length of 40 quarters and a forecast length of one quarter. As mentioned before, this limits the sample for which we will obtain the sentiment series to 1970 to 2021. Each training sub-period contains 37 889 speeches on average, while each forecast period contains 1037 speeches on average. As described in the previous subsection, each speech receives two scores that correspond to the cosine similarities between the single speeches and the constructed expansionary and contractionary vectors. We build a continuous fiscal policy index by subtracting the similarity to the contractionary vector from the similarity to the expansionary vector. For example, a fiscal sentiment of zero could imply that a speech discusses both expansionary and contractionary fiscal policy measures and the cosine scores to both fiscal vectors cancel out.

We interpret the resulting series of fiscal sentiment as a measure of the propensity of parliamentarians to engage in expansionary fiscal policy. A higher sentiment score means that policy is more inclined to implement expansionary policies.

Fig. 4 shows the final sentiment series at quarterly frequency. These series will enter the estimated VAR models in the next section. We show one series for the entire Bundestag, i.e. for speeches of all members, one for speeches delivered by members of the governing parties and one based on speeches from opposition members. The figure also highlights recessions in Germany as identified by the German Council of Economic Experts.[12] We find a close co-movement of government and opposition sentiment. The evolution of sentiment is consistent with the established historical narrative: First, the shift from a center-left to center-right coalition in 1982, which was also motivated by concerns about fiscal sustainability, is clearly visible as a drop in fiscal sentiment of the government towards a more restrictive policy stance. Second, when in 2003 the European Commission triggered the excessive deficit procedure against Germany, which is specified in the Stability and Growth Pact, fiscal sentiment deteriorated. Third, we spot a fall in fiscal sentiment starting in 2014 when the coalition government pushed its policy of budget surpluses ("schwarze Null"). Towards the end of our sample period, i.e. after the 2017 election, there is a remarkable upward trend in sentiment.

We further analyze fiscal sentiment with regard to the changing coalition governments. Fig. 5 presents the average sentiment in the German Bundestag for each election period. In most periods, the government exhibits a more expansionary sentiment than the opposition. The only exception to this is the 2017–2021 coalition with Finance Minister Wolfgang Schäuble pushing the "schwarze Null". The distance between government and opposition varies over time and reaches a maximum in the early 1980s.

In the appendix, we compare our sentiment index with indices based on fiscal policy-specific dictionaries.

3.4. Exogenous vs. endogenous fiscal sentiment

The approach so far allows us to construct sentiment time series covering all aspects of fiscal policy. In the empirical estimation below, we need to impose restrictions on the model in order to identify shocks to fiscal sentiment. The text analysis can help us in the identification of sentiment shocks and make the restrictions imposed on the VAR model more credible.

For that purpose, we exploit the range of issues members of the Bundestag discuss in their speeches. Some issues are clearly reflecting the current economic situation, while others reflect structural decisions that have fiscal consequences but are unrelated to the current state of the economy. Consider one parliamentary speech on the state of the German military and another speech on

[12] See https://www.sachverstaendigenrat-wirtschaft.de/en/topics/business-cycles-and-growth/konjunkturzyklus-datierung.html for the recession dates.

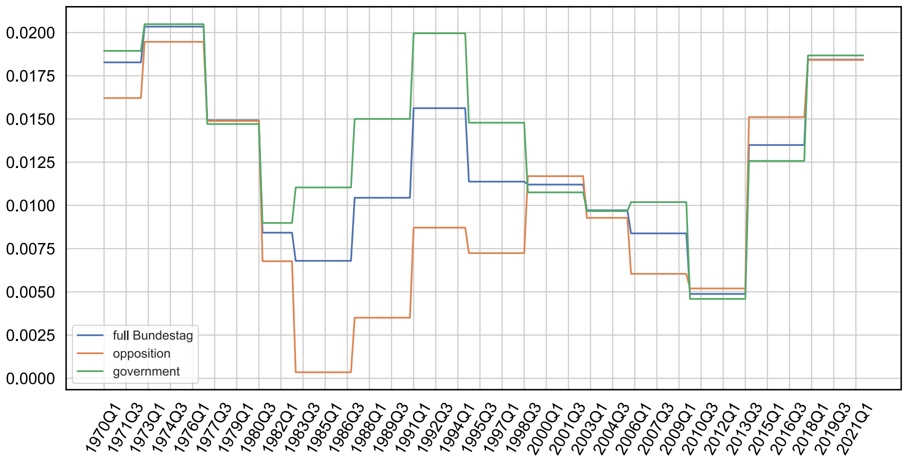

Fig. 5. Average fiscal sentiment per election period. *Notes:* The figure shows the average fiscal sentiment of the entire Bundestag the government and the opposition in each election period from 1970Q1 to 2021Q3.

the government's budget. Both speeches impact government spending and contribute to fiscal sentiment. However, the latter is a response to the state of the business cycle, while the former is unrelated to the cycle.

We follow the notion of Cloyne (2013), who picks up an approach of Romer and Romer (2010), that an "exogenous policy decision is one that was not designed to offset other macroeconomic shocks" (p. 1511). Consequently, we refer to the sentiment expressed in speeches on the German military and the budget as examples for exogenous and endogenous sentiment, respectively. Fluctuations in exogenous fiscal sentiment should reflect changes in the preferences of politicians or long-term structural issues such as infrastructure, the military, the development of the new federal states, social justice, etc., which are not an immediate response to the business cycle. Fluctuations in endogenous fiscal sentiment are a consequence of the economic cycle.

Using topic modeling, it is possible to uncover topics within the large corpus of Bundestag speeches. These topics can be categorized as either exogenous or endogenous branches of policies. We can then categorize speeches as reflecting exogenous or endogenous policy and calculate the sentiment for each speech.[13]

One of the most popular topic modeling technique is the Latent Dirichlet Allocation (LDA) introduced by Blei et al. (2003). The LDA model is considered the prototype of probabilistic generative topic models. The model assumes that all documents in a corpus (which altogether possess exactly K predetermined topics) are generated from a random mix of different latent topics. A topic, in turn, represents a specific mix of words from the vocabulary. Each document d can contain words linked to any of these K topics. However, the documents (i.e. parliamentary speeches) differ with regard to the weights of the topics. Thus, a document can be described as a probability distribution θ_d over topics, and a topic can be seen as a probability distribution β_k over the given vocabulary. Since the topics are not known in advance, the goal of the LDA model is to learn the topic mix θ_d in each document and the word mix β_k in each topic, from the data. For more technical details see Blei and Lafferty (2009).

We estimate an LDA model with the assumption of $K = 100$ topics.[14] We manually classify the 100 topics as either exogenous or endogenous topics, based on the notion of Cloyne (2013), or as topics unrelated to fiscal policy. This classification is guided by our thorough reading of the learned topic-word distributions β_k of all topics. These topic-word distributions are usually visualized using word clouds. Fig. 6 shows, for example, two word clouds that represent the topics "Bundeswehr" on the left-hand and "Budget, Debt, Investment" on the right-hand side. We can classify the "Bundeswehr" topic as exogenous fiscal policy and the "Budget, Debt, Investment" topic as endogenous fiscal policy.

A total of nine topics are classified as endogenous fiscal policy, while 25 topics are classified as exogenous fiscal policy.[15] Based on this classification, we can aggregate the topic-document-probability θ_d for each speech and obtain three scores summing up to one that provide us with the affiliation of each speech with exogenous or endogenous policy or issues unrelated to fiscal policy. The first two scores are used to derive endogenous and exogenous sentiment series. To this end, we multiply the fiscal sentiment time series with the corresponding score and obtain exogenous and endogenous fiscal sentiment series by aggregation over all speeches.

Fig. 7 shows the evolution of these fiscal sentiment time series for the entire Bundestag. For most of the sample period, both series closely move together. The peak in exogenous sentiment coincides with the German re-unification in 1990. Endogenous sentiment is particularly high during the oil crises of the 1970s. In 2009, both series start to diverge. One reason for this could be the increased number of exogenous shocks with large fiscal consequences, i.e. the European sovereign debt crisis, the increase in migration due to the war in Syria and the Covid-19 pandemic.

---

[13] Our distinction between exogenous and endogenous fiscal policy rests on a classification of topics. In contrast, the pioneering work of Romer and Romer (2010) is based on a careful reading of actual tax laws. In a useful validation exercise, one could check whether a topic model is able to replicate the Romer and Romer (2010) shock series. We leave this for future research.

[14] We apply the `gensim`-package of Řehůřek and Sojka (2010).

[15] The thematic interpretation of the topics and the classification into exogenous and endogenous policy are listed in the appendix.

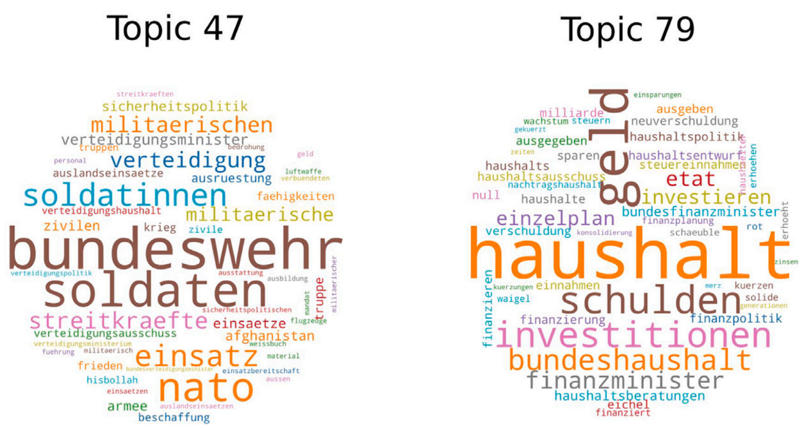

Fig. 6. Word clouds of topics "Bundeswehr" (topic 47) and "Budget, Debt, Investment" (topic 79). *Notes:* The figure shows the wordmix of a topic visualized as a word cloud of the 50 most important words. The larger the font size, the more important this word is for this topic. Topic 47 ("Bundeswehr") is classified as exogenous fiscal policy. Topic 79 ("Budget, Debt, Investment") is classified as endogenous fiscal policy.

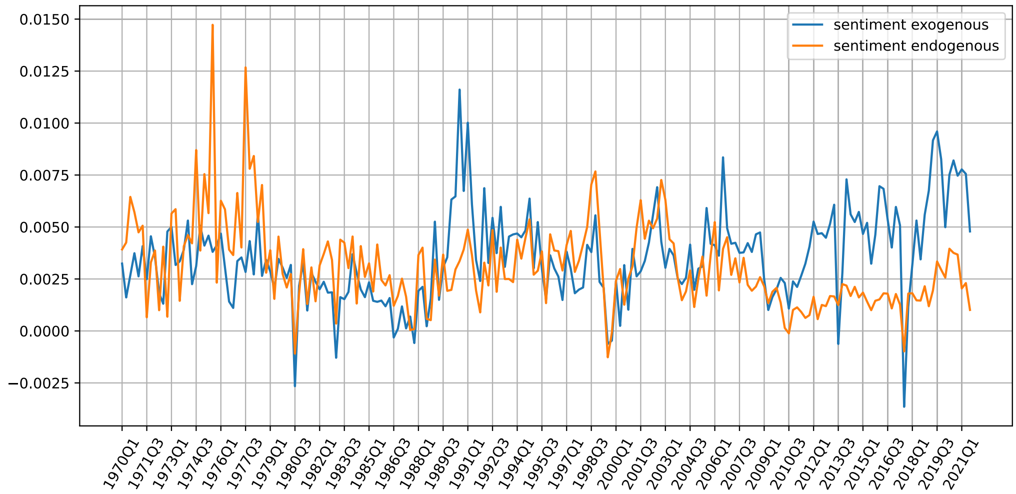

Fig. 7. Exogenous and endogenous fiscal sentiment. *Notes:* The figure shows exogenous and endogenous fiscal sentiment reflected in speeches of all members of the Bundestag.

Fig. 8 shows the correlation of our main sentiment indices with business cycle variables as well as the main series from the influential survey conducted by the ifo Institute and the World Uncertainty Index for Germany at leads and lags.[16] The aggregate sentiment series is positively correlated with growth and inflation and negatively correlated with unemployment. As expected, the endogenous sentiment is more closely correlated with growth, inflation and unemployment than the exogenous sentiment. This difference is particularly striking for the correlation with the ifo series: exogenous sentiment is procyclical, while endogenous sentiment is countercyclical. An increase in uncertainty coincides with higher exogenous sentiment and lower endogenous sentiment.

To further illustrate the evolution of exogenous and endogenous fiscal sentiment, Fig. 9 highlights selected episodes. We show sentiment following four large adverse economic shocks that could be considered exogenous: the first oil crisis in 1973, the second in 1979, the collapse of Lehman Brothers at the peak of the global financial crisis and the Covid-19 pandemic. Since we normalize sentiment to one at the beginning of each episode, this figure is not informative about the level of sentiment. Rather, it showcases the responses to these events. In each of the four episodes, we see a stronger increase in sentiment about endogenous fiscal policy topics compared to exogenous fiscal policy topics. This is intuitive as the large adverse shocks should elicit an endogenous fiscal stabilization as a response.

[16] We merged the ifo series on business climate and business expectations, respectively, from the pre-1991 period with the series from the post-1991 period. The raw data is available at https://www.ifo.de/en/ifo-time-series. The uncertainty index is available at https://worlduncertaintyindex.com/.

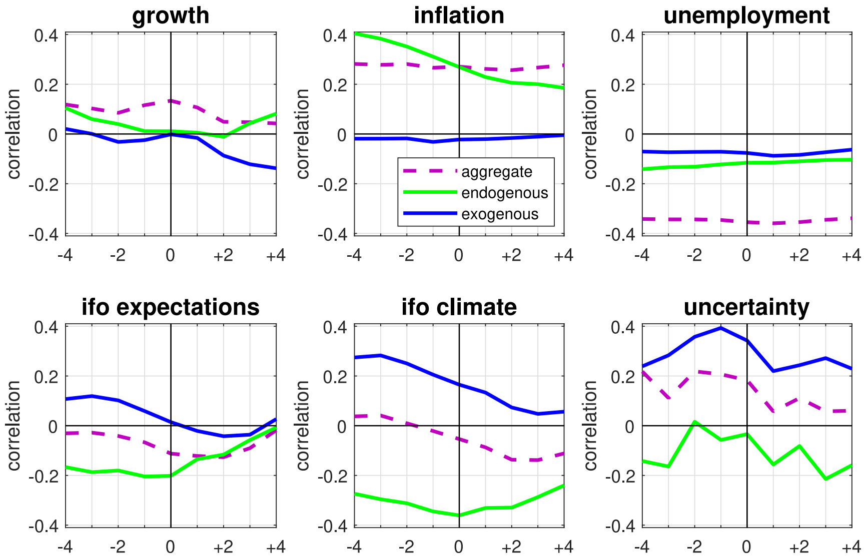

Fig. 8. Cross-correlation of sentiment with macroeconomic variables. *Notes:* The figure shows the correlation of the sentiment indices in quarter t with key macroeconomic variables as well as the indicators from the survey conducted by the ifo Institute and the World Uncertainty Index in $t \pm k$ with $k = 0, \ldots, 4$.

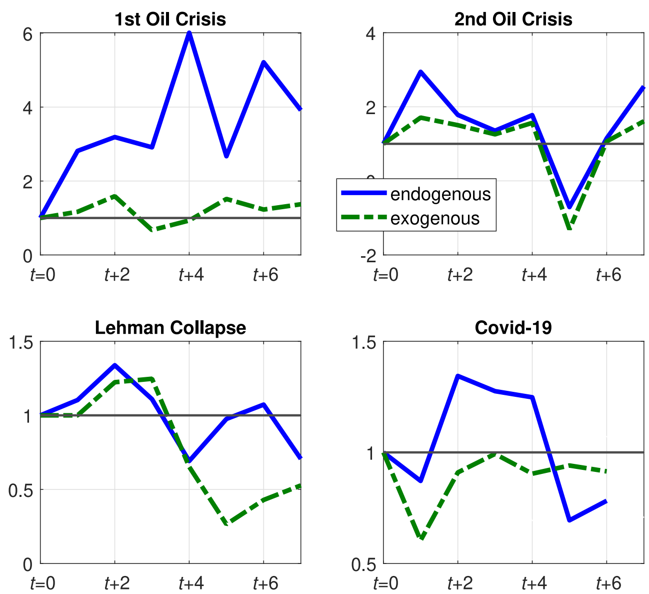

Fig. 9. Fiscal sentiment in selected episodes. *Notes:* The figure shows the standardized fiscal sentiment for the government and the opposition during selected episodes. We normalize both sentiment series to one in 1973Q3 (1st oil crisis), 1979Q2 (2nd oil crisis), 2008Q2 (collapse of Lehman Brothers) and 2019Q4 (Covid-19).

4. Estimating the macroeconomic effects

We now study the macroeconomic effects of exogenous changes in fiscal sentiment as reflected in the textual data. For that purpose, we augment a relatively standard VAR model by our new sentiment series.

4.1. VAR model

We estimate a reduced-form VAR model with p lags

$$y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + C x_t + \varepsilon_t, \tag{1}$$

where y_t is the $n \times 1$ vector of endogenous variables, $A_1, \dots, A_p$ are $n \times n$ coefficient matrices and x_t is a vector of exogenous regressors such as constant terms, dummies and a time trend. The vector of error terms, ε_t, follows a multivariate normal distribution, $\varepsilon_t \sim N(0, \Sigma)$, where Σ is the variance–covariance matrix with $E(\varepsilon_t \varepsilon_t') = \Sigma$. The residuals are mutually uncorrelated at all leads and lags.

The VAR model is estimated using Bayesian methods, thus treating parameters as random variables drawn from an underlying probability distribution. We assume a Normal-Wishart prior, though our results remain unchanged for alternative priors specifications.[17]

As discussed before, we adopt a recursive identification scheme. Let us write the model in its structural form

$$D_0 y_t = D_1 y_{t-1} + \cdots + D_p y_{t-p} + F x_t + \eta_t, \tag{2}$$

where $\eta \sim N(0, \Gamma)$ is the vector of structural shocks and the D matrices are defined appropriately. With $D = D_0^{-1}$, the reduced-form error terms and the structural shocks are linked by $\varepsilon_t = D\eta_t$. We assume that D is lower triangular, thus imposing restrictions on the contemporaneous interdependencies between the endogenous variables.

4.2. Data

In our baseline model, we include four endogenous variables: the log of real government expenditure, Gov_t, the log of real GDP, GDP_t, the log of real private consumption, $Cons_t$, and one of the three alternative indicators of fiscal sentiment derived in the previous sections, $Senti_t^j$, with $j \in (Bundestag, Government, Opposition)$. Hence, the vector of endogenous variables is

$$y_t' = \begin{bmatrix} Gov_t & GDP_t & Cons_t & Senti_t^j \end{bmatrix}. \tag{3}$$

In four alternative specifications, we augment the baseline VAR by additional variables. First, we include the log of real private investment and the employment rate as two additional variables reflecting the domestic business cycle. Second, we add government revenues and the real interest rate. Third, we include the federal budget balance and business expectations from the ifo survey. Fourth, we add two variables to the model that reflect the open-economy transmission of fiscal policy, i.e. the log of the real effective exchange rate and the trade balance relative to GDP. All log series are multiplied by 100. The estimation frequency is quarterly and the data spans 1970Q1–2021Q3.

We also include a time trend and an impulse dummy that is one in 2020Q2 and zero otherwise. This dummy capture the extreme drop in real economic activity due to the Covid-19 pandemic and the ensuing lockdown. As Germany went into lockdown in the second half of March 2020, we choose to set the dummy to one in the second quarter of 2020. We estimate the VAR model for $p = 8$ lags.

The core time series are taken from the OECD data file: real GDP, real private consumption and real government consumption. In an extended model, we also use real gross fixed capital formation, i.e. investment, and real government revenues (interpolated to quarterly frequency). Both are also taken from the OECD. All series are seasonally adjusted. The data for the federal budget balance is taken from the Bundesbank and interpolated to quarterly frequency.[18]

4.3. Identification

We draw on the extensive literature on the identification of exogenous fiscal policy shocks pioneered by Blanchard and Perotti (2002) and applied, among others, by Fatás and Mihov (2001), Galí et al. (2007), Born and Müller (2012), Auerbach and Gorodnichenko (2012) and Ilzetzki et al. (2013) and impose a recursive ordering onto the variables.[19] The ordering of the variables as in (3) implies that in a given quarter government expenditure is predetermined. Changes in GDP or consumption, respectively, do not contemporaneously affect government spending. Our specific application lends itself to a straightforward extension of this line of literature. The starting point of our analysis is that fiscal policy is made in parliaments and that parliamentary decisions take time. This is exactly why spending is predetermined in a given quarter. Our text data reflects this parliamentary debate. In fact, as argued by Mertens and Ravn (2010) and Ramey (2011a, 2016), among others, changes in government spending could be anticipated several quarters in advance. Ordering government spending first thus implies that VARs do not identify unanticipated government spending shocks. We will revisit this issue in the next section.

Including information on the fiscal sentiment expressed in parliament alleviates this concern. We order sentiment last. Hence, a change in fiscal sentiment as expressed in Bundestag speeches should not contemporaneously drive either government expenditure, nor real GDP or real consumption. At the same time, fiscal sentiment is contemporaneously responding to the business cycle.[20] If sentiment is informative about fiscal policy, we should expect that an exogenous increase in sentiment, i.e. a shift towards a more expansionary policy stance, raises government expenditure and economic activity. This effect should be more pronounced for the sentiment of speakers of the parties forming the government compared to opposition speakers.

[17] In order to estimate the model, we rely on the BEAR toolbox for MATLAB, see https://www.ecb.europa.eu/pub/research/working-papers/html/bear-toolbox.en.html.

[18] The additional data series are taken from the FRED database of the Federal Reserve Bank of St. Louis. The series ID are: unemployment rate (LMUNRRTTDEQ156S), consumer price index (DEUCPIALLMINMEI), long-term bond yield (IRLTCT01DEQ156N), short-term interbank rate (IR3TIB01DEQ156N), real effective exchange (CCRETT01DEQ661N), nominal exports (DEUGDPNQDSMEI), nominal imports (DEUIMPORTQDSMEI) and nominal GDP (DEUEXPORTQDSMEI).

[19] Tenhofen et al. (2010) apply this identification for VAR model with German data.

[20] The results remain unchanged if we order fiscal sentiment second, i.e. after government spending but before GDP and consumption.

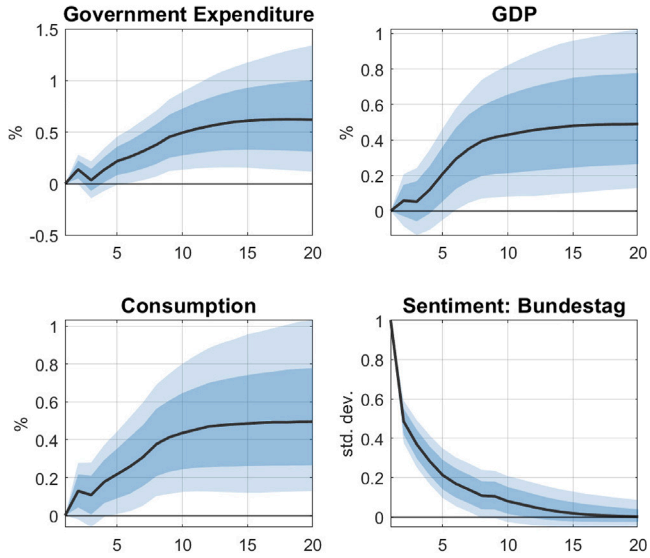

Fig. 10. Response to fiscal sentiment (entire Bundestag). *Notes:* The figure shows the responses of the endogenous variables to an increase in fiscal sentiment as reflected in speeches of all members of the Bundestag. All responses are derived from a recursively identified Bayesian VAR model with 8 lags and Normal-Wishart priors. The shaded areas cover 68% and 90% of all draws.

4.4. Results

Fig. 10 shows the responses of our endogenous variables to an increase in the fiscal sentiment of the entire Bundestag one standard deviation in size. All figures also include probability bands that cover 68% and 90% of all draws. As a consequence of the shock, government expenditure increase strongly by about 0.5%. This response is highly significant and very persistent.[21] Hence, a shift towards a more expansionary policy stance as reflected in the speeches held in the Bundestag does indeed cause a subsequent increase in government spending. We also see that the increase in spending needs time to unfold: the increase becomes significant six quarters after the impulse to sentiment. The additional spending has real economic effects: real GDP as well as real private consumption increase by about 0.4%.

In Fig. 11, we depict the responses to fiscal sentiment as reflected in the speeches of the members of parliament who belong to the governing parties. While the increase in government spending is slightly smaller than in Fig. 10, the overall macroeconomic effects are somewhat larger. Again, expansionary fiscal policy has a strong impact on income and consumption. If we include only sentiment in those speeches that are delivered by the opposition parties, see Fig. 12, the increase in government spending is similar. Nevertheless, the implied fiscal multiplier, i.e. the response of GDP relative to the response of spending, is higher for a shock to government sentiment compared to opposition sentiment as we will show below.

We now extend the baseline model by additional variables. In the first alternative model, we include real investment and unemployment. Both variables are also ordered behind government expenditure but before our sentiment indicator. Fig. 13 reports the corresponding impulse response functions. The shift in sentiment causes a strong increase in government expenditure and a significant increase in private consumption and investment. As expected, the response of investment is larger than the response of consumption. The unemployment rate falls after the shock, which is in line with the economic expansion.

In the second alternative model, we augment the baseline variables by government revenues and the real interest rate measured by the difference between the long-term German bond yield and the year-on-year inflation rate. Fig. 14 shows that more expansionary sentiment causes an increase in government revenues. As our sentiment variables reflects both sentiment on spending and taxation, this result suggests that the expansion of economic activity, which should raise revenues, is stronger than the drop in revenues, which could results from lower taxes. As in Ramey (2016), a fiscal expansion leads to a drop in the real interest rate. In the appendix, we show that this is because the increase in inflation overcompensates the response of the nominal interest rate. Hence, the expansionary fiscal impulse is leading to inflationary pressure, which is in line with standard New-Keynesian models.

In a third extension, we include the central government's budget balance and the ifo index of business expectations. Due to a lack of quarterly data, we had to interpolate the annual data series of the budget balance in percent of GDP. When sentiment in speeches turns more expansionary, the budget balance falls significantly and business expectations improve, see Fig. 15.

We study a fourth alternative model specification that reflects the open-economy transmission of fiscal policy. In addition to the four variables of the baseline model, we include the real effective exchange rate and the trade balance relative to GDP. Again, both

[21] Fisher and Peters (2010) also find quite persistent effects of (defense) spending.

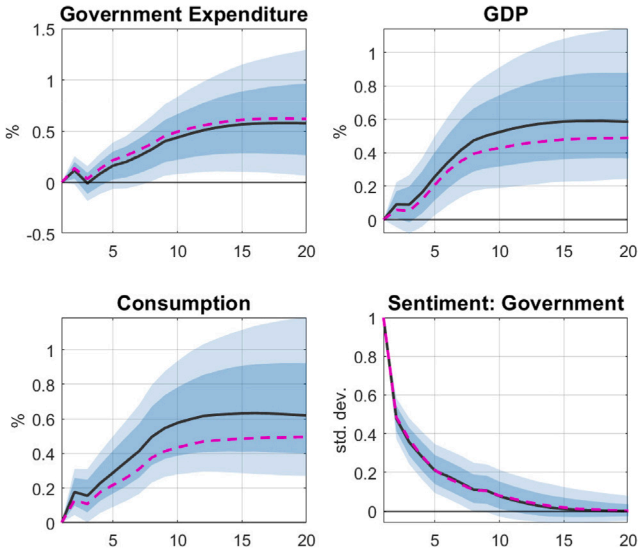

Fig. 11. Response to fiscal sentiment (government). *Notes:* The figure shows the responses of the endogenous variables to an increase in fiscal sentiment as reflected in speeches of members of the parties forming the government. The dashed line is the response from the model with speeches from the entire Bundestag. All responses are derived from a recursively identified Bayesian VAR model with 8 lags and Normal-Wishart priors. The shaded areas cover 68% and 90% of all draws.

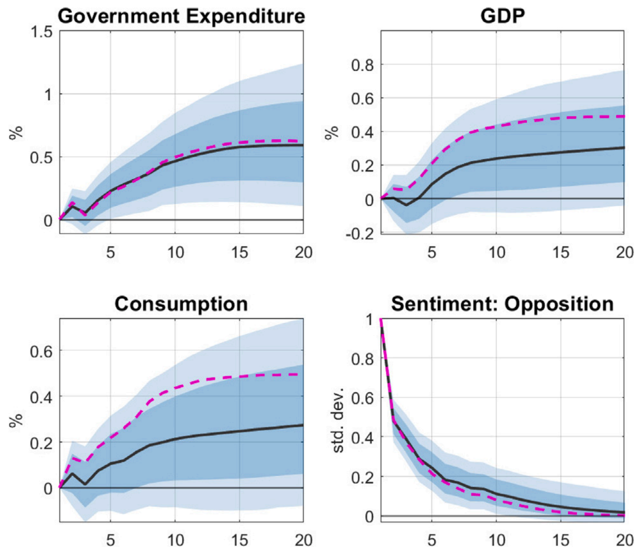

Fig. 12. Response to fiscal sentiment (opposition). *Notes:* The figure shows the responses of the endogenous variables to an increase in fiscal sentiment as reflected in speeches of all members in the opposition. The dashed line is the response from the model with speeches from the entire Bundestag. All responses are derived from a recursively identified Bayesian VAR model with 8 lags and Normal-Wishart priors. The shaded areas cover 68% and 90% of all draws.

variables are ordered after government expenditure but before fiscal sentiment. Textbook models of an open economy suggest that a fiscal expansion, here reflected by a shifts towards a more expansionary sentiment, causes a real appreciation of the domestic currency and a deterioration of the trade balance. Fig. 16 shows that the impulse responses are perfectly in line with standard models. Germany experiences a real appreciation of about 0.7% after 10 quarters as well as a drop in the trade balance by 0.2 percentage points.[22]

[22] Thus, our findings are in line with theory and do not exhibit a puzzling depreciation after an expansionary policy, see Forni and Gambetti (2016) and Ferrara et al. (2021) for this debate.

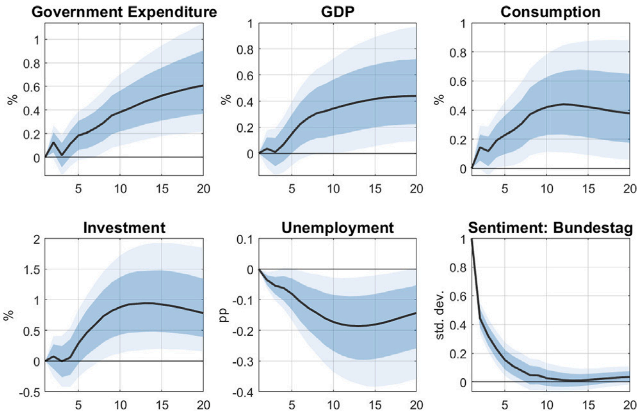

Fig. 13. Response to fiscal sentiment (Bundestag): extended VAR. *Notes:* The figure shows the responses of the endogenous variables to an increase in fiscal sentiment as reflected in speeches of all members of the Bundestag. All responses are derived from a recursively identified Bayesian VAR model with 8 lags and Normal-Wishart priors. The shaded areas cover 68% and 90% of all draws.

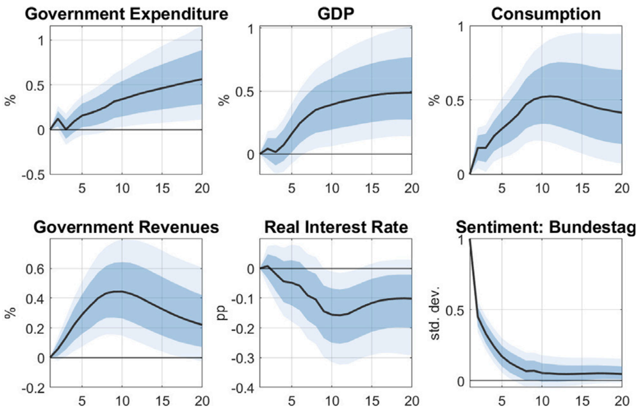

Fig. 14. Response to fiscal sentiment (Bundestag): extended VAR. *Notes:* The figure shows the responses of the endogenous variables to an increase in fiscal sentiment as reflected in speeches of all members of the Bundestag. All responses are derived from a recursively identified Bayesian VAR model with 8 lags and Normal-Wishart priors. The shaded areas cover 68% and 90% of all draws.

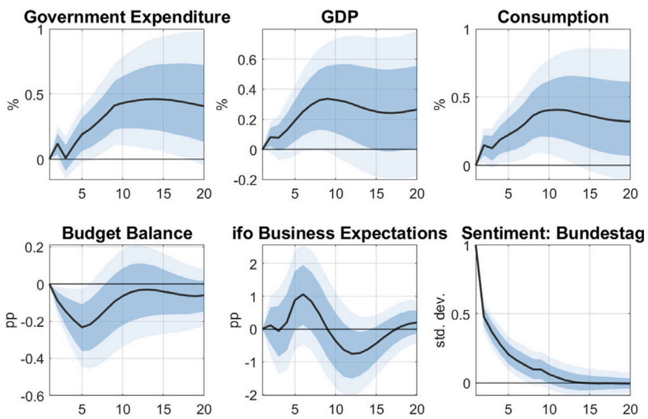

Fig. 15. Response to fiscal sentiment (Bundestag): extended VAR. *Notes:* The figure shows the responses of the endogenous variables to an increase in fiscal sentiment as reflected in speeches of all members of the Bundestag. All responses are derived from a recursively identified Bayesian VAR model with 8 lags and Normal-Wishart priors. The shaded areas cover 68% and 90% of all draws.

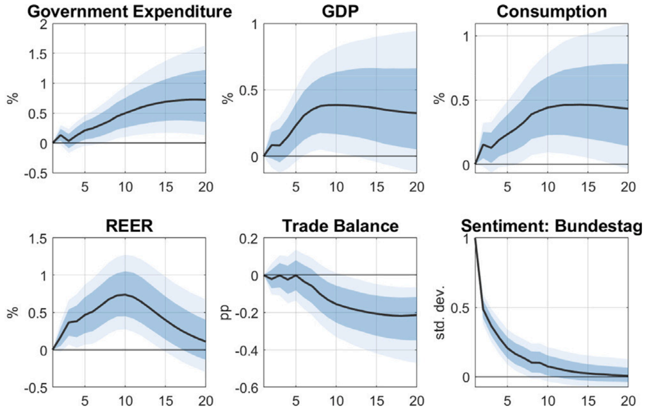

Fig. 16. Response to fiscal sentiment: open-economy VAR. *Notes:* The figure shows the responses of the endogenous variables to an increase in fiscal sentiment as reflected in speeches of all members of the Bundestag. All responses are derived from a recursively identified Bayesian VAR model with 8 lags and Normal-Wishart priors. The shaded areas cover 68% and 90% of all draws.

Taken together, these findings suggest that a change in the fiscal sentiment expressed in the Bundestag does indeed have real economic effects. Our results lend support to the Keynesian paradigm, i.e. suggesting that expansionary fiscal policy does indeed increase income and consumption.[23]

4.5. Robustness

In our baseline model, we include the log-levels (times 100) of macroeconomic aggregates such as GDP, consumption and government expenditure. An alternative would be to detrend the three macroeconomic variables. We follow Gordon and Krenn (2010), Ramey (2016), Ramey and Zubairy (2018) and Ilori et al. (2012) and detrend each variable using the trend in real GDP, i.e. we include variable x_t, which is either real GDP, real consumption or real government expenditure, as $100 \times (\ln x_t - \ln y_t^{trend})$, where y_t^{trend} is the estimated trend in real GDP. We derive y_t^{trend} either from fitting a quadratic trend to log GDP as in Gordon and Krenn (2010), Ramey (2016) and Ramey and Zubairy (2018) or from applying the Hamilton (2018) filter to log real GDP as in Ilori et al. (2012).

The results based on the quadratic trend in GDP are shown in Fig. 17. An unexpected increase in fiscal sentiment raises government expenditure, GDP as well as private consumption. All responses are distinct from zero and look similar to the results based on the log-level variables presented in the previous section. In Fig. 18, we show the impulse responses based on the GDP trend derived from the Hamilton (2018) detrending procedure. The change in fiscal sentiment still pushes up private consumption.

How strong is the response of GDP relative to the increase in government spending? We address this question by calculating the cumulative response of GDP divided by the cumulative response of government expenditure. This calculation is often used to quantify the fiscal multiplier, e.g. Ramey and Zubairy (2018). However, it should be emphasized that this is an *implied* multiplier only as we do not look at the consequences of an increase in government spending but rather an increase in fiscal sentiment. Fig. 19 reports these multipliers for each horizon of the impulse response. For the baseline model, the multiplier remains slightly below one.[24] It is larger for the Gordon and Krenn (2010) detrending and lower if we apply the Hamilton (2018) filter. Interestingly, the implied multiplier is much larger following an increase in government sentiment compared to opposition sentiment. One potential reason for that could be that the composition of spending is different when the increase in spending is triggered by opposition sentiment compared to government sentiment.

4.6. Exogenous and endogenous fiscal sentiment

In the estimated VAR model of the previous section, we ordered sentiment last, thus treating each aspect of fiscal policy alike by assuming that it drives real economic activity with a lag of at least one quarter. We now estimate the model based on the distinction between sentiment about exogenous and endogenous fiscal policy, respectively.

[23] In the appendix, we show that the presence of fiscal sentiment in the VAR model also tends to weaken the transmission of shocks to government spending to GDP.

[24] The size of the multiplier is in line with the literature: Ramey (2011b) argues that the spending multiplier is between 0.8 and 1.5 for the U.S. economy. For Germany, Tenhofen et al. (2010) estimates multipliers larger than ours, while Berg (2016) finds smaller multipliers. Our results are consistent with the literature on fiscal news shocks. Ramey (2011a) and Ben Zeev and Pappa (2017) also find that spending multipliers for anticipated fiscal spending are rather large.

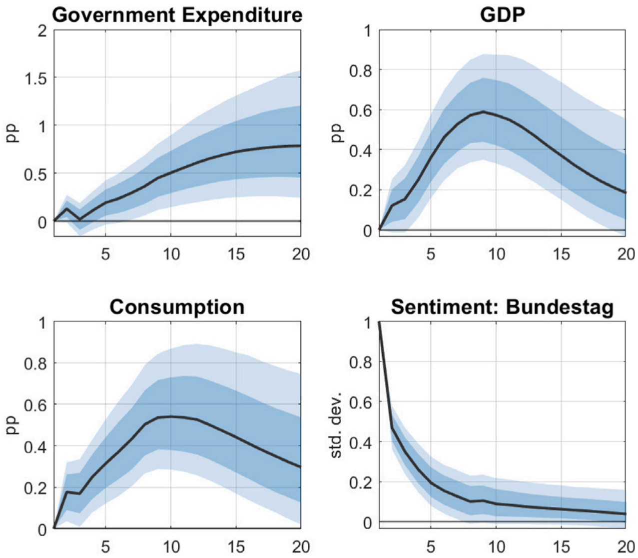

Fig. 17. Response to fiscal sentiment: detrending. *Notes:* The figure shows the responses of the endogenous variables to an increase in fiscal sentiment. The three macroeconomic variables are detrended by the quadratic trend in GDP. All responses are derived from a recursively identified Bayesian VAR model with alternative lags orders p and Normal-Wishart priors. The purple responses are derived from the model estimated by OLS. The shaded areas cover 68% and 90% of all draws.

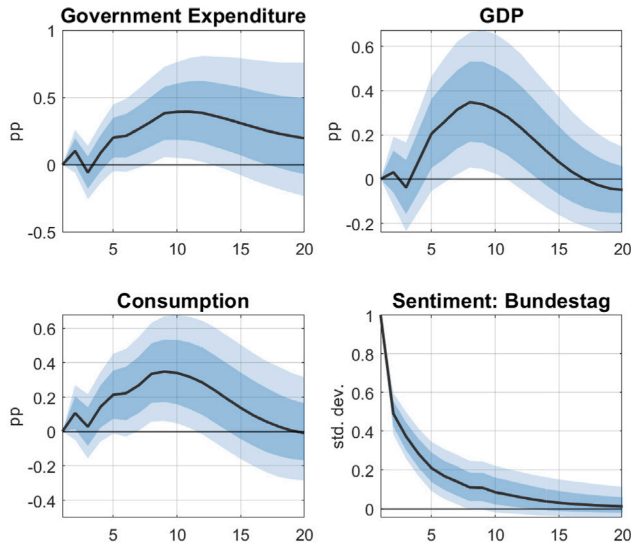

Fig. 18. Response to fiscal sentiment: detrending. *Notes:* The figure shows the responses of the endogenous variables to an increase in fiscal sentiment. The three macroeconomic variables are detrended by the Hamilton (2018) trend in GDP. The black responses are derived from a recursively identified Bayesian VAR model with alternative lags orders p and Normal-Wishart priors. The purple responses are derived from the model estimated by OLS. The shaded areas cover 68% and 90% of all draws.

Exogenous sentiment, $Senti_t^{j,exo}$, is ordered first as it is not contemporaneously driven by the remaining variables.[25] Endogenous sentiment, $Senti_t^{j,endo}$, is ordered last as it is contemporaneously responsive to the other variables in the model. The vector of variables becomes

$$y_t' = \left[Senti_t^{j,exo} \quad Gov_t \quad GDP_t \quad Cons_t \quad Senti_t^{j,endo} \right]. \tag{4}$$

[25] It should be stressed that despite the labeling the sentiment series is still endogenous in the sense that it potentially responds to the other variables with a delay of one quarter.

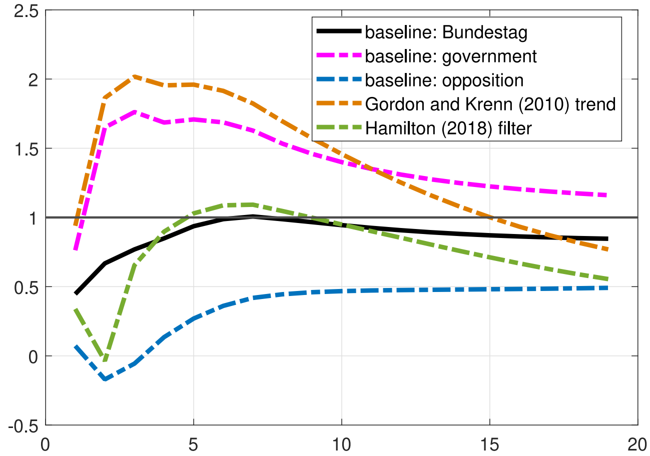

Fig. 19. Cumulative fiscal multipliers. *Notes:* The figure shows the cumulative responses of GDP to an increase in fiscal sentiment relative to the cumulative response of government expenditure for alternative model specifications.

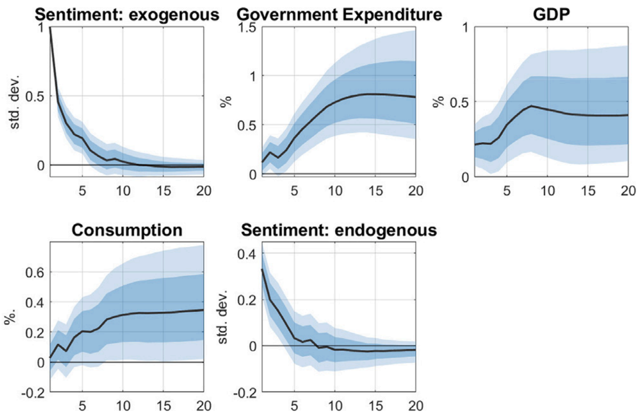

Fig. 20. Response to exogenous fiscal sentiment. *Notes:* The figure shows the responses of the endogenous variables to an increase in exogenous fiscal sentiment as reflected in speeches of all members of the Bundestag. All responses are derived from a recursively identified Bayesian VAR model with 8 lags and Normal-Wishart priors. The shaded areas cover 68% and 90% of all draws.

Fig. 20 reports the estimated responses to an increase in exogenous sentiment. The shock to exogenous sentiment causes a significant increase in each of the four other variables. The responses are quantitatively similar to the responses in the baseline model with a cumulative multiplier below one. An increase in the endogenous fiscal sentiment, Fig. 21, has no significant effect on the other variables. Think about shocks to endogenous sentiment as deviations of fiscal stabilization from the implicit policy rule. These deviations are not reflected in higher expenditure and have no effect on GDP and private consumption. The results suggest that most of our baseline findings are in fact driven by shifts in exogenous fiscal sentiment.

5. Fiscal foresight revisited

The literature on government spending shocks argues that fiscal foresight invalidates the recursive Blanchard–Perotti identification, see Mertens and Ravn (2010), Ramey (2011a, 2016) and Ellahie and Ricco (2017). Our data set allows us to examine the degree to which fiscal sentiment expressed in parliamentary speeches allows the public to forecast government spending shocks. In other words, we check whether the estimated government spending shock in a Blanchard and Perotti (2002) style model with government spending ordered first in a recursive VAR is predicted by our fiscal sentiment index. We estimate a VAR model with the three core variables used before: government spending, GDP and consumption. We use all three alternative treatments of the variables, i.e. log-levels, quadratic detrending and Hamilton-detrending and do not include sentiment at this stage. Importantly, we adopt the recursive Blanchard–Perotti identification scheme that orders government spending first. Hence, government spending is predetermined with respect to output and consumption. This provides us with three alternative series of structural government spending shocks — one for each treatment of the endogenous variables.

In the next step, we assess whether sentiment contains information that allows us to predict future government spending shocks. The following model in the spirit of Jordà (2005) regresses the government spending shock for model k at time $t+h$ on the

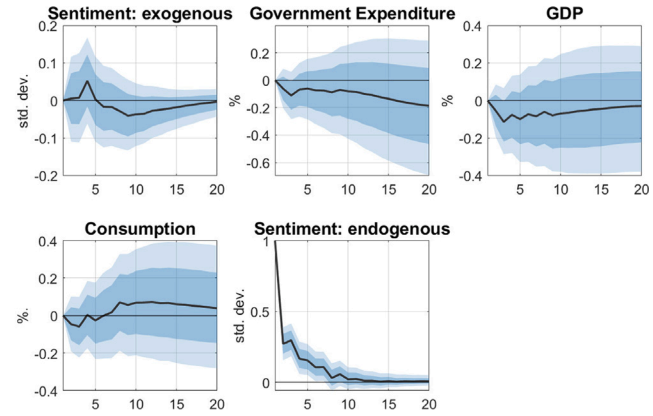

Fig. 21. Response to endogenous fiscal sentiment. *Notes:* The figure shows the responses of the endogenous variables to an increase in endogenous fiscal sentiment as reflected in speeches of all members of the Bundestag. All responses are derived from a recursively identified Bayesian VAR model with 8 lags and Normal-Wishart priors. The shaded areas cover 68% and 90% of all draws.

parliamentary sentiment at time t

$$shock^k_{t+h} = \alpha_h + \beta_h Senti^j_t + \gamma_h X_{t-1} + \varepsilon_{t+h} \tag{5}$$

with $j \in (Bundestag, Government, Opposition)$. A significant β_h would indicate that current sentiment predicts future government spending shocks. We estimate this model for each shock k as well as for the standardized sentiment of the government, the opposition and the entire Bundestag. The vector X_t contains contemporaneous and lagged realizations of GDP, consumption and government expenditure as control variables. As the dependent variable is the result of a structural identification, it should be orthogonal to these control variables. Nevertheless, we include these variables as controls.

Panel (a) of Fig. 22 plots the estimated β_h as a function of h for the sentiment of the entire Bundestag. The results are consistent across the alternative treatments of the variables: A shift towards a more expansionary fiscal sentiment in t predicts an increase in government spending six to eight quarters later. Hence, the government spending shocks are predictable from parliamentary speeches. The results are weakly significant at the 90% level. When we narrow the set of speeches to the members of the governing parties, see panel (b) of the figure, we obtain similar results. Information from speeches of politicians from the opposition parties, see panel (c), does not predict future government spending shocks.

Overall, this section supports the notion that government spending shocks identified from a recursive ordering can indeed be anticipated. This also underlines the relevance of the parliamentary process as a source of information for upcoming changes to fiscal policy.

6. Conclusions

This paper went to the roots of fiscal policy-making — the debate in parliament. We use the full set of parliamentary speeches delivered in the German Bundestag as a source of information about fiscal preferences. An embedding-based approach using the latest advances in text mining provides us with a sentiment index on a scale from expansionary to restrictive that summarizes the debate about fiscal policy. This sentiment series has real economic effects: recursively identified VAR models suggest that an increase in fiscal sentiment towards a more expansionary policy stance increases government spending, output and consumption. Hence, a change in fiscal sentiment has macroeconomic effects consistent with standard New-Keynesian business cycle models.

We draw two main conclusions: First, we believe textual data to be very informative about economic policy-making. The rich information incorporated in parliamentary speeches is particularly promising for researchers interested in fiscal policy. In this paper, we focused on the consequences of fiscal sentiment for government expenditure and the macroeconomy. In follow-up work, we will study the consequences of disagreement about fiscal policy between the government and the opposition. Using this data set to assess the consequences of sentiment on the revenue side of public finances could also be an interesting way forward.

Second, the identification of government spending shocks often rests on the assumption that the part of government spending not forecastable from lags of the endogenous variable is a suitable exogenous shock. Information from parliamentary debates about fiscal policy might help enhance this identification. As parliamentary speeches partly forecast future expenditure, only the part of government expenditure that is orthogonal to lags of business cycle variables as well as lags of textual information from the parliamentary debates should qualify as a government expenditure shock.

Data availability

Replication package (Original data) (GitHub)

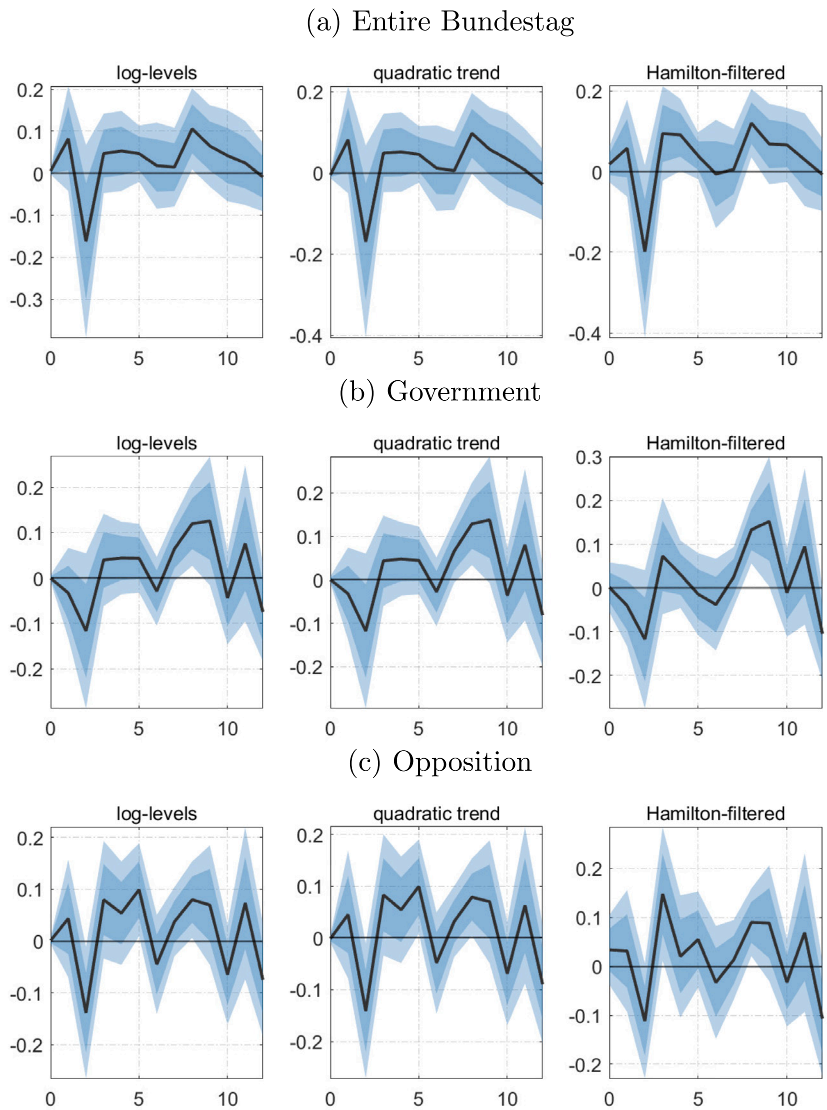

Fig. 22. Response of government expenditure shock to sentiment. *Notes:* The figure shows the responses of the recursively identified government expenditure shock to an increase in fiscal sentiment. The shaded areas reflect 68% and 90% confidence bands constructed from Newey–West standard errors.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.euroecorev.2024.104827.

References

Abercrombie, G., Batista-Navarro, R., 2020. Sentiment and position-taking analysis of parliamentary debates: a systematic literature review. J. Comput. Soc. Sci. 3 (1), 245–270, URL: https://doi.org/10.1007/s42001-019-00060-w.

Allard, J., Catenaro, M., Vidal, J.-P., Wolswijk, G., 2013. Central bank communication on fiscal policy. Eur. J. Political Econ. 30, 1–14.

Auerbach, A., Gorodnichenko, Y., 2012. Measuring the output responses to fiscal policy. Am. Econ. J.: Econ. Policy 4, 1–27.

Baker, S.R., Bloom, N., Davis, S.J., 2016. Measuring economic policy uncertainty. Q. J. Econ. 131 (4), 1593–1636.

Ben Zeev, N., Pappa, E., 2017. Chronicle of a war foretold: the macroeconomic effects of anticipated defence spending shocks. Econ. J. 127, 1568–1597.

Berg, T.O., 2016. Time varying fiscal multipliers in Germany. Rev. Econ. 66, 13–46.

Bird, E.L., Klein, E., 2009. Natural Language Processing with Python. O'Reilly Media Inc.

Blanchard, O., Perotti, R., 2002. An empirical characterization of the dynamic effects of changes in government spending and taxes on output. Q. J. Econ. 117, 1329–1368.

Blei, D.M., Lafferty, J.D., 2009. Topic models. In: Srivastava Ashok, N., Sahami, M. (Eds.), Text Mining. Chapman and Hall/CRC, pp. 101–124.

Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent dirichlet allocation. J. Mach. Learn. Res. 3, 993–1022.

Born, B., Müller, G., 2012. Government spending shocks in quarterly and annual time series. J. Money Credit Bank. 44, 507–517.
Burst, T., Krause, W., Lehmann, P., Lewandowski, J., Matthieß, T., Merz, N., Regel, S., Zehnter, L., 2020. Manifesto Corpus. Version: 2020-1. WZB Berlin Social Science Center, Berlin.
Christofzik, D.I., Fuest, A., Jessen, R., 2022. Macroeconomic effects of the anticipation and implementation of tax changes Germany: Evidence from a narrative account. Economica 89, 62–81.
Cloyne, J., 2013. Discretionary tax changes and the macroeconomy: new narrative evidence from the United Kingdom. Amer. Econ. Rev. 103, 1507–1528.
Denny, M.J., Spirling, A., 2018. Text preprocessing for unsupervised learning: why it matters, when it misleads, and what to do about it. Polit. Anal. 26 (2), 168–189.
Dybowski, T.P., Adämmer, P., 2018. The economic effects of U.S. presidential tax communication: evidence from a correlated topic model. Eur. J. Political Econ. 55, 511–525.
Ellahie, A., Ricco, G., 2017. Government purchases reloaded: Informational insufficiency and heterogeneity in fiscal VARs. J. Monetary Econ. 90, 13–27.
Fatás, A., Mihov, I., 2001. The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence. Working Paper, INSEAD.
Ferrara, L., Metelli, L., Natoli, F., Siena, D., 2021. Questioning the puzzle: fiscal policy, real exchange rate and inflation. J. Int. Econ. 133, 103524.
Fisher, J.D.M., Peters, R., 2010. Using stock returns to identify government spending shocks. Econ. J. 120, 414–436.
Forni, M., Gambetti, L., 2016. Government spending shocks in open economy VARs. J. Int. Econ. 99, 68–84.
Galí, J., López-Salido, D., Vallés, J., 2007. Understanding the effects of government spending on consumption. J. Eur. Econom. Assoc. 5, 227–270.
Gennaro, G., Ash, E., 2021. Emotion and reason in political language. Econ. J. 132 (643), 1037–1059.
Gordon, R.J., Krenn, R., 2010. The End of the Great Depression: VAR Insight on the Roles of Monetary and Fiscal Policy. Working Paper No 16380, National Bureau of Economic Research.
Grimmer, J., Stewart, B.M., 2013. Text as data: the promise and pitfalls of automatic content analysis methods for political texts. Political Analysis 21 (3), 267–297.
Guajardo, J., Leigh, D., Pescatori, A., 2014. Expansionary austerity? International evidence. J. Eur. Econom. Assoc. 12 (4), 949–968.
Hamilton, J., 2018. Why you should never use the Hodrick-Prescott filter. Rev. Econ. Stat. 100, 831–843.
Hayo, B., Mierzwa, S., 2022. Legislative tax announcements and GDP: evidence from the United States, Germany, and the United Kingdom. Econom. Lett. 216, 110548.
Hayo, B., Uhl, M., 2014. The macroeconomic effects of legislated tax changes in Germany. Oxf. Econ. Pap. 66, 397–418.
Honnibal, M., Montani, I., Van Landeghem, S., Boyd, A., 2020. SpaCy: industrial-strength natural language processing in Python. URL: https://spacy.io/.
Hruzik, J., 2019. Pybundestag. URL: https://github.com/Jhruzik/pybundestag.
Ilori, A.E., Paez-Farrell, J., Thoenissen, C., 2012. Fiscal policy shocks and international spillovers. Eur. Econ. Rev. 141, 103969.
Ilzetzki, E., Mendoza, E.G., Vegh, C.A., 2013. How big (small?) are fiscal multipliers? J. Monetary Econ. 60, 239–254.
Jordà, O., 2005. Estimation and inference of impulse responses by local projections. Amer. Econ. Rev. 95, 161–182.
Kapfhammer, F., Larsen, V.H., Thorsrud, L.A., 2020. Climate Risk and Commodity Currencies. Working Papers No 10/2020, Centre for Applied Macro- and Petroleum economics (CAMP), BI Norwegian Business School, URL: https://ideas.repec.org/p/bny/wpaper/0093.html.
Latifi, A., 2023. MaFiText-bundestag speeches: Processing stenographic protocols of the German Bundestag. Unpublished. in preparation.
Le, Q., Mikolov, T., 2014. Distributed representations of sentences and documents. In: Proceedings of the 31st International Conference on International Conference on Machine Learning - Volume 32. ICML '14, JMLR.org, pp. II–1188–II–1196.
Loughran, T., McDonald, B., 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. J. Finance 66, 35–65.
Mertens, K., Ravn, M., 2010. Measuring the impact of fiscal policy in the face of anticipation: a structural VAR approach. Econ. J. 120, 393–413.
Mertens, K., Ravn, M., 2012. Empirical evidence on the aggregate effects of anticipated and unanticipated US tax policy shocks. Am. Econ. J.: Econ. Policy 4, 145–181.
Mikolov, T., Chen, K., Corrado, G., Dean, J., 2013a. Efficient estimation of word representations in vector space. In: Bengio, Y., LeCun, Y. (Eds.), 1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings. URL: http://arxiv.org/abs/1301.3781.
Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., Dean, J., 2013b. Distributed representations of words and phrases and their compositionality. In: Burges, C., Bottou, L., Welling, M., Ghahramani, Z., Weinberger, K. (Eds.), Advances in Neural Information Processing Systems. Vol. 26, Curran Associates, Inc., pp. 3111–3119, URL: https://proceedings.neurips.cc/paper_files/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.
Picault, M., Renault, T., 2017. Words are not all created equal: a new measure of ECB communication. J. Int. Money Finance 79, 136–156, URL: https://www.sciencedirect.com/science/article/pii/S0261560617301808.
Ramey, V., 2011a. Identifying government spending shocks: it's all in the timing. Q. J. Econ. 126, 1–50.
Ramey, V.A., 2011b. Can government purchases stimulate the economy? J. Econ. Lit. 49, 673–685.
Ramey, V.A., 2016. Macroeconomic shocks and their propagation. Handb. Macroecon. 2A, 71–162.
Ramey, V.A., Zubairy, S., 2018. Government spending multipliers in good times and in bad: evidence from US historical data. J. Polit. Econ. 126, 850–901.
Řehůřek, R., Sojka, P., 2010. Software framework for topic modelling with large corpora. In: Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks. ELRA, Valletta, Malta, pp. 45–50.
Rheault, L., Cochrane, C., 2020. Word embeddings for the analysis of ideological placement in parliamentary corpora. Polit. Anal. 28 (1), 112–133.
Ricco, G., Callegari, G., Cimadomo, J., 2016. Signals from the government: policy disagreement and the transmission of fiscal shocks. J. Monetary Econ. 82, 107–118.
Rodriguez, P.L., Spirling, A., 2022. Word embeddings: what works, what doesn't, and how to tell the difference for applied research. J. Polit. 84 (1), 101–115.
Romer, C.D., Romer, D.H., 2010. The macroeconomic effects of tax changes: estimates based on a new measure of fiscal shocks. Amer. Econ. Rev. 100, 763–801.
Tenhofen, J., Wolff, G.B., Heppke-Falck, K.H., 2010. The macroeconomic effects of exogenous fiscal policy shocks in Germany: a disaggregated SVAR analysis. Jahrb. Nationalökon. Stat. 230, 328–355.
Werner, A., Lacewell, O., Volkens, A., 2011. Manifesto Coding Instructions, fourth fully revised ed. URL: https://manifesto-project.wzb.eu/down/papers/handbook_2011_version_4.pdf, May 2011.