36 Years of Textual Data on Populism: Using a Dynamic Topic Model and a Novel National Parliament Dataset to Analyze Austrian Right-Wing Populism

By: Josh Klier

2022

Fulbright-Mach Research Scholar

Introduction:

The use of text in quantitative political methods has increased over recent years. This increase has included a rise in the use of parliamentary records and databases. Parliamentary records are a useful source of data for not just political scientists, but humanities, other social sciences and computational linguistics. The most obvious reason for this being that they are databases of what politicians have said in parliaments through time. This allows for a database that not only encapsulates the national political discourse through time but also has a common denominator in where the speeches occur for valid and sound comparison and analysis.[1] The most famous and expansive parliamentary database is that of Hansard, the U.K's parliamentary database that has its beginnings in the 1800s. However, the U.K. is not the only country with an expansive parliamentary corpus. Many European countries have corpora created from stenographical protocols.[2] Austria's longest available corpus runs from 1997-2017.[3]

The aim of this paper is twofold: namely, to publish a widespread database of all the Austrian national parliament plenary sittings from 1986-2021 and run a dynamic topic model on a filtered version of the dataset to analyze what right-wing populists have said in parliament during this time period.[4] The novel dataset introduced in this paper will allow further analysis of the past decades of Austrian politics. The length of the dataset is also of importance specifically

---

[1] Take, for example, running a textual analysis with speeches made in parliament and representatives' speeches at rallies. The rhetoric would be very different and lead to a very biased analysis.

[2] CLARIN is a collection of these parliamentary databases online. See more here: https://www.clarin.eu

[3] The longest Austrian corpus and can be found here: (INSERT LINK) This database is part of a larger project that encapsulates texts from different European parliaments. I have used the Austrian part of this database that stretches from 1996-2017, meaning that this new database adds 14 years of plenary sitting data.

[4] Austrian Plenary sittings of the national council are sittings that happen a couple of times per month and involve all the representatives of parliament. Plenary sittings largely consist of deliberations of bills, matters of topical interest or question time. In this sense, the plenary sittings are used as a space in which politicians stress the points that are important to their agenda - whether through questions, criticisms or support of a certain bill. Thus, the plenary sittings provide an easily observable space to analyze what politicians/their parties stand for and how this evolves through time.

for Austrian populism studies as the timeframe corresponds with the recent successes of the Freiheitliche Partei Österreich (FPÖ) - the largest right-wing populist party in Austria. Namely, the dataset stretches from 1986 onwards, corresponding with the beginning of the party's time of successful leadership under their then leader Jörg Haider.

This paper investigates the recent history of right-wing populism in Austria by textually analyzing all speeches made by populist party members in Austrian plenary sittings from 1986-2021. The intent behind this is to discover the latent semantic structures in a large corpus of plenary sitting speeches. As politicians have a limited amount of time to deliberate in plenary sittings, they must prioritize which topics to engage. Plenary sittings, therefore, provide a valid environment for textual analysis as the setting and the importance of the sitting remains the same through time.

By utilizing a dynamic topic model, I analyze the latent semantic structures and how they change through time – in both composition and attention paid to by national council members. Meaning, this paper seeks to create a roadmap of what right-wing populists have said in parliament over the last 36 years, and in turn give attention to: what populists stand for, seeing how these topics change over time in their semantic makeup, and analyzing the prevalence of these topics over time.

This paper adds to populism literature by using a novel dataset of speakers in Austrian plenary sittings from right-wing populist parties that spans from 1986-2021. In total the dataset has 320,000 speeches. After filtering for just FPÖ party members, the dataset included 30,010 speeches. Additionally, this paper adds to populism research as it is the first paper to use a dynamic topic model to study populist speeches in national parliaments over time.

Related work/literature review:

During the past few decades, Europe has seen a resurgence in right-wing populism. With this has come a plethora of corresponding populism research attempting to figure out why populism is enjoying such success and what this might mean for the future. One of the points that populist researchers seem to agree on through all this research is that populist parties, apart from their anti-establishment nature, stand for a variety of different issues across countries. Researchers have observed this trend and described populist parties as having 'chameleon-like character'.[5] In this sense, populism may adapt to different contexts, or likewise through time.[6] Studying populism through time and how parties adapt could give valuable insight to the theoretical assumptions of what populist parties stand for. Additionally, if one assumes that populist politicians are elected based on what they stand for and how they stand against other parties on issues in parliament, textually analyzing populist speeches in parliament could offer valuable insight into why populists are elected.

Populism, as well as exhibiting chameleon-like character, has been described as a thin-centered ideology that has 'a restricted core attached to a narrower range of political concepts'.[7] Furthermore, populism does not have 'the same level of intellectual refinement and consistency' as other political ideologies.[8] The nature of having chameleon-like character can be seen as an effect of being a political ideology that lacks core values; it is easy to flip positions on an issue if it isn't a core value of your ideology.

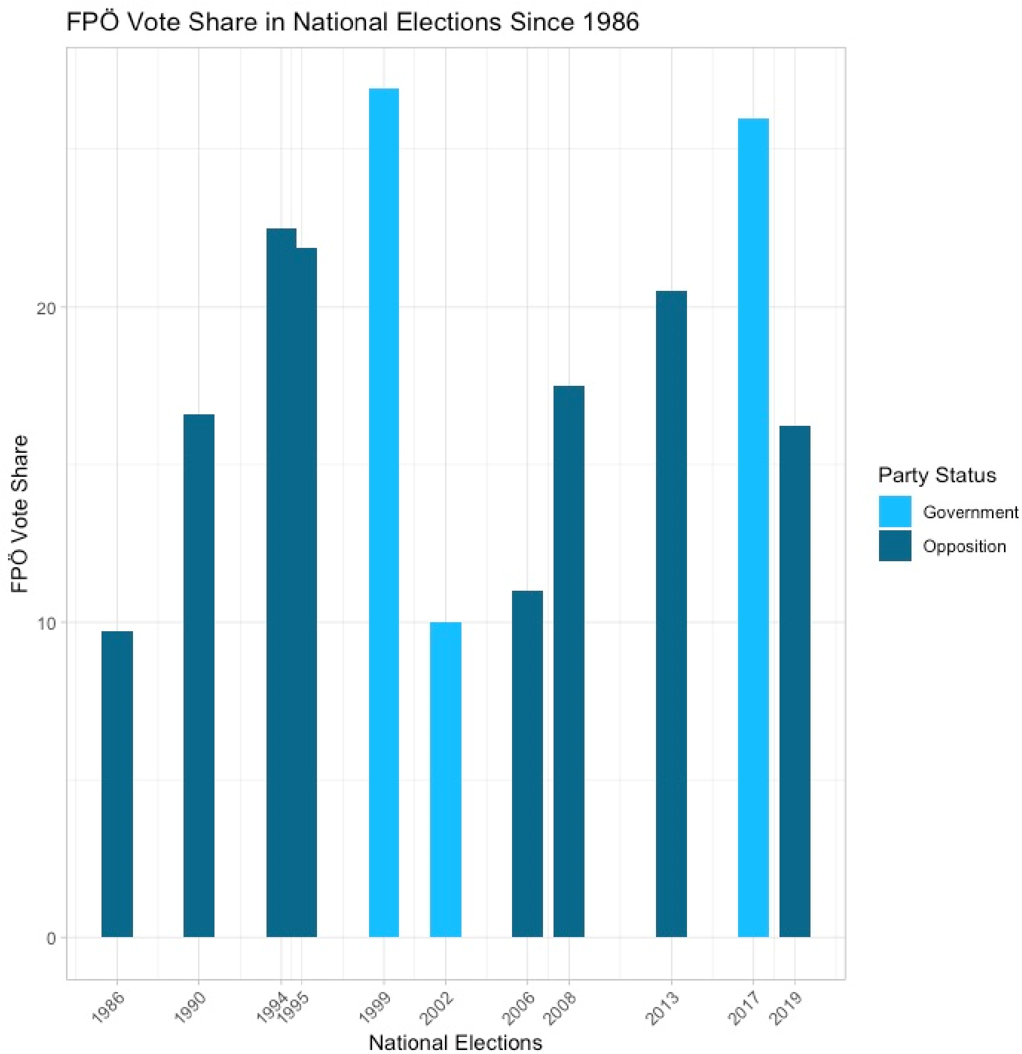
This paper focuses on studying Austrian populism through time to reveal more about their thin-centered ideology and their possible chameleon-like nature. When studying European populism in the post-war period, there are few better countries than Austria suited for analysis. Austria has not only a long standing presence of right-wing populism but also multiple occasions in which a populist party has been a part of a coalition government. This means that populism in Austria has succeeded in not only driving national political discourse by being a part of governing coalitions, but also by returning to government after voted out. In short, Austrian right-wing post-war populism has an important history for dynamic analyses as it has been prominent for a long period of time and has often bounced back to prominence after an unsuccessful stint in government.

Over the course of the dataset, the FPÖ has experienced many different election results, including being a part of government on three different occasions. The success of the FPÖ over the length of the dataset is summarized below by Figure 1. The FPÖ enjoyed successes in the first three elections during the period of analysis and reached the goal of being in government in

[5] Taggart, P. (2002). Populism and the Pathology of Representative Politics. In: Mény, Y., Surel, Y. (eds) Democracies and the Populist Challenge. Palgrave Macmillan, London.
[6] Ibid.
[7] Mudde, Cas. "The Populist Zeitgeist." *Government and Opposition*39, no. 4 (2004): 541–63. http://www.jstor.org/stable/44483088.
[8] Ibid.

1999 after securing almost 30% in the national election. After two stints in government, the FPÖ faced sharp decreases in their popularity and had to rebuild the party. This rebuild proved to be successful and after three further national elections it secured another spot in government after winning over 25 percent of the vote share in 2017.

Figure 1:



Literature Review:

In the recent years, the use of dynamic topic models has become more prominent in political science research. Originally proposed by Blei in the early 2000s, topic modelling is a probabilistic model that seeks to discover latent topics within a large corpus of texts.[9] They allow for much larger analyses of data as humans are not required to read and label documents and their topics. Therefore, the creation of these models has paved way for the ability to conduct textual analyses on databases that are far too large for human capability.

In Greene et al. (2017), they analyze European Parliament plenary texts using a dynamic topic model. Specifically, they use a new corpus of all English speeches from 1999-2014. The authors use a dynamic topic model that is based on two layers of non-negative Matrix Factorization (NMF). They find that the political agenda of the European Parliament evolves significantly over time.[10] In Müller-Hansen et al. (2021), they investigate 70 years of German parliamentary debates on coal using dynamic topic modeling to analyze the changes in the political debate on coal through time. The authors here sort on only the speeches made mentioning coal over the time window and find that topics evolved over time very similarly to how domestic energy policy evolved.[11] In Béchara et al. (2021), they use dynamic topic modelling on a corpus of all U.K. House of Commons speeches from 1935-2014.[12] Finally, in Quinn et al. (2010), they use a dynamic topic model to analyze speeches in the U.S. senate from 1997-2004 to examine the agenda of the Senate.[13] From this literature review it is evident that dynamic topic models are widely used in the political science field. This paper contributes empirically to the study of populism as it is – to my knowledge – the only paper to use dynamic topic modeling on a database of populist speeches.

Methods:

Data collection

[9] David M. Blei. 2012. Probabilistic topic models. <i>Commun. ACM</i> 55, 4 (April 2012), 77–84. https://doi.org/10.1145/2133806.2133826

[10] Greene, Derek, and James P. Cross. “Exploring the Political Agenda of the European Parliament Using a Dynamic Topic Modeling Approach.” *Political Analysis* 25, no. 1 (2017): 77–94. doi:10.1017/pan.2016.7.

[11] Müller-Hansen, Finn & Callaghan, Max & Lee, Yuan & Leipprand, Anna & Flachsland, Christian & Minx, Jan. (2021). Who cares about coal? Analyzing 70 years of German parliamentary debates on coal with dynamic topic modeling. Energy Research & Social Science. 72. 101869. 10.1016/j.erss.2020.101869.

[12] Béchara, Hannah, Alexander Herzog, Slava Jankin, and Peter John. “Transfer Learning for Topic Labeling: Analysis of the UK House of Commons Speeches 1935–2014.” Research & Politics, (April 2021). https://doi.org/10.1177/20531680211022206.

[13] Quinn, Kevin M., Burt L. Monroe, Michael Colaresi, Michael H. Crespin, and Dragomir R. Radev. “How to Analyze Political Attention with Minimal Assumptions and Costs.” *American Journal of Political Science* 54, no. 1 (2010): 209–28. http://www.jstor.org/stable/20647980.

First, both PDF and HTML documents were downloaded from the Austrian Parliamentary website.[14] The HTML files were then put aside until the PDFs were further pre-processed. The workflow of converting the PDF files into a workable textual dataset was the following: convert the PDFs into text files, clean the text files, and split them into a workable data frame through the use of various regular expressions.

To convert the PDFs into text files, each individual page was converted into a jpeg file. Then, Tesseract's OCR package in python was used to covert the numerous jpeg images into text files. OCR stands for optical character recognition and refers to a computer reading text from an image. The next step was to clean these large text files. The goal was to keep what each speaker said during a plenary sitting, each sitting was filled with parts that were not needed. For example, each sitting had an introduction before the first speaker spoke, headings and footers from each PDF page, and sometimes the text went longer than the closure of the sitting. Below are some examples of these from the PDFs. (insert images). Lastly, after removing the unwanted material from the text files, the text data needed to be split into a dataset in which each frame corresponded to a speaker and their speech. Numerous different regular expressions were implemented to tell the computer to split the text after a new speaker began talking. Regular expressions were also used to extract the various metadata points used for this dataset. These included the: date of the speech, speaker name, party affiliation and text of speech. In comparison to some other databases published, this is a relatively few number of metadata points to include in a textual database, however, due to the: time/scope of the project, the length of the dataset and the differing qualities of the texts over the 35 years only the essential metadata was collected.[15]

Dynamic topic modeling

[14] See: https://www.parlament.gv.at/PAKT/STPROT/

[15] Note on Validity: As this dataset begins in the 1980s and is using scanned original PDFs, there are bound to be some errors in the converting of these PDFs to text files. These could come in the form of a falsely read umlaut (i.e. fur instead of für) or a falsely read punctuation mark (I.e. { instead of (). Some errors are harmless while others negatively affect the accuracy of the dataset as they could be located in places that my regular expressions are looking. If my regular expressions find a { in a place where a (should be, the text will not be split and therefore make the dataset overall less accurate. As the quality of the PDFs increases with each legislative period, it can be assumed that a higher fraction of the plenary sittings in the earlier legislative periods will be less accurate.

This paper uses a dynamic topic model to analyze speeches made by populist politicians in the Austrian national council. Dynamic topic models are a variation of topic models; these are hierarchical probabilistic models that find the latent topics in a collection of texts. Topic models, the most common being latent Dirichlet allocation models, assume that each document is made up of a set of topics and each topic is made up of a set of words. In this sense, documents can be reduced to their topics by the make-up of their words. Dynamic topic models differ from their counterparts by relaxing the assumption that documents are not ordered through time. This allows researchers to look at how topics evolve, in their prevalence and composition, over time.[16] For example, when Blei et al. introduce their dynamic topic model, they also analyzed a subset of journal articles over the time span of 1881 to 1999. With 250 articles from each year, they build a 20 term dynamic topic model that shows how topics such as Atomic Physics or Neuroscience adjust in their composition through time.[17] By keeping the topics the same through time but allowing their parts to be interchanged, this allows for an analysis of how topics have evolved, and in this paper, how political discourse has changed through time. For example, when thinking about the topic of immigration, the discourse changes through time as the topic not only changes in public debate but also because the geographical sources of immigration change.

For this paper, I use a model package called BERTopic. BERTopic is a topic model algorithm that uses state of the art embeddings and a unique class-based version of Term Frequency-Inverse Document Frequency (TF-IDF) to generate topics. [18] TF-IDF measures how important a term is to a document in a corpus. The more often a word appears in a document, the higher it’s term frequency. However, this value is offset by the word frequency in the corpus.[19]

According to the creator of the model, BERTopic is “competitive” when comparing the model with other topic models. The topic model and dynamic topic model are compared to other models’ topic coherence and topic diversity measures, two indicators of model performance that

---

[16] Müller-Hansen, Finn & Callaghan, Max & Lee, Yuan & Leipprand, Anna & Flachsland, Christian & Minx, Jan. (2021). Who cares about coal? Analyzing 70 years of German parliamentary debates on coal with dynamic topic modeling. Energy Research & Social Science. 72. 101869. 10.1016/j.erss.2020.101869.

[17] David M. Blei and John D. Lafferty. 2006. Dynamic topic models. In <i>Proceedings of the 23rd international conference on Machine learning</i> (<i>ICML '06</i>). Association for Computing Machinery, New York, NY, USA, 113–120. https://doi.org/10.1145/1143844.1143859

[18] Grootendorst, Maarten. "BERTopic: Neural topic modeling with a class-based TF-IDF procedure." *arXiv preprint arXiv:2203.05794* (2022).

[19] TDIDF.com

emulate human judgement.[20] In the six total measurements between six different models, BERTopic places in the top two models five of the six times, coming in third once. The strengths of BERTopic are that the model remains competitive regardless of the language model used and the class-based TD-IDF model that allows for topics to be represented as a distribution of words. Whereas the weaknesses of the model largely center around the assumption that each document only contains a single topic.[21] However, when considering the plenary sitting dataset at hand, this does not seem to be too inhibiting of an assumption as most speeches address a single law, discussion, question etc. When considering long speeches made by party members, the model only considers the main topic of the speech was and leaves the rest out.

The model functions largely in three steps: first, using a pre-trained language model document embeddings are created. Document embeddings transform the dictionary of words in a text to a vector of numbers. This is a first step in most text-based machine learning models as algorithms function with numbers. The primary function of these embeddings is to cluster semantically similar documents by reducing documents to a vector space. Next, semantically similar clusters are created from the reduced document embeddings. Each of these clusters represent a topic. Finally, a unique class-based TF-IDF is used to extract the topic representations for each topic.[22]

A normal TD-IDF measures the importance of a word to a document. The equation in BERTopic is modified so that the importance of terms is measured in relation to their topic instead.[23] The class-based TD-IDF is shown below:

$$\mathrm{W}_{x,c} = \ tf_{x,c} \times log(1+\tfrac{A}{f_x})$$

Where $tf_{x,c}$ is the frequency of word x in class c, f_x is the occurrence of word x across all clusters, and A is the average number of words of across all clusters. The positive one is included to make the score a positive integer. Whereas the normal TD-IDF equation measures terms in relation to documents, here the documents are switched to classes, thus modeling the importance

[20] Grootendorst, Maarten. "BERTopic: Neural topic modeling with a class-based TF-IDF procedure." *arXiv preprint arXiv:2203.05794* (2022).
[21] Ibid.
[22] Ibid.
[23] Ibid.

of words in clusters or topics. [24] When running a dynamic topic model, first the class-based TD-IDF measure above is calculated to create a global view of topics. Then, topic representations for year i are created by multiplying the term frequency of documents at a certain timestep with their pre-calculated global IDF values.[25][26]

The data for the model was collected from the dataset described earlier. This includes all the national parliament plenary sittings in Austria from 1986-2021. The dataset was then filtered to only include speeches from FPÖ party members. The total number of speeches is 30,010, which is an average of 834 speeches per year. After filtering the data to only include the FPÖ, the text needed to be further cleaned before running the topic model. This began by taking the roots of the words in the speeches; therefore, reducing the overlap of common words being used more than once in the same topic. This process is called lemmatizing. For example, verbs are reduced to their indefinite form to create one entry for all forms of the words.[27] After lemmatizing the dataset, I created a set of stop words to eliminate from the dataset. These are words that are commonly used and included in topics, yet do not reveal anything about the true meaning of the topic at question, such as: word articles, prepositions, conjunctions. Additionally, the date column in the original dataset was changed to only include the year in analysis. This means that there are 36 iterations in the dynamic topic model instead of the total number of plenary sittings over the 36 years. This larger, more coherent time window allows for better visualization of the model as parliamentary schedules stay relatively similar through the years.[28] [29]

After preprocessing the filtered dataset, I run a normal topic model without calculating the dynamic nature of the dataset. This creates various topics from the data. In the topic model I set various other parameters to increase the coherence of the topics created. This includes cutting out words that appear in over 95% of all documents, setting the minimum number of topic occurrences to the number of years in the dataset, namely 36. This means that a topic has to

[24] Ibid.
[25] Ibid.

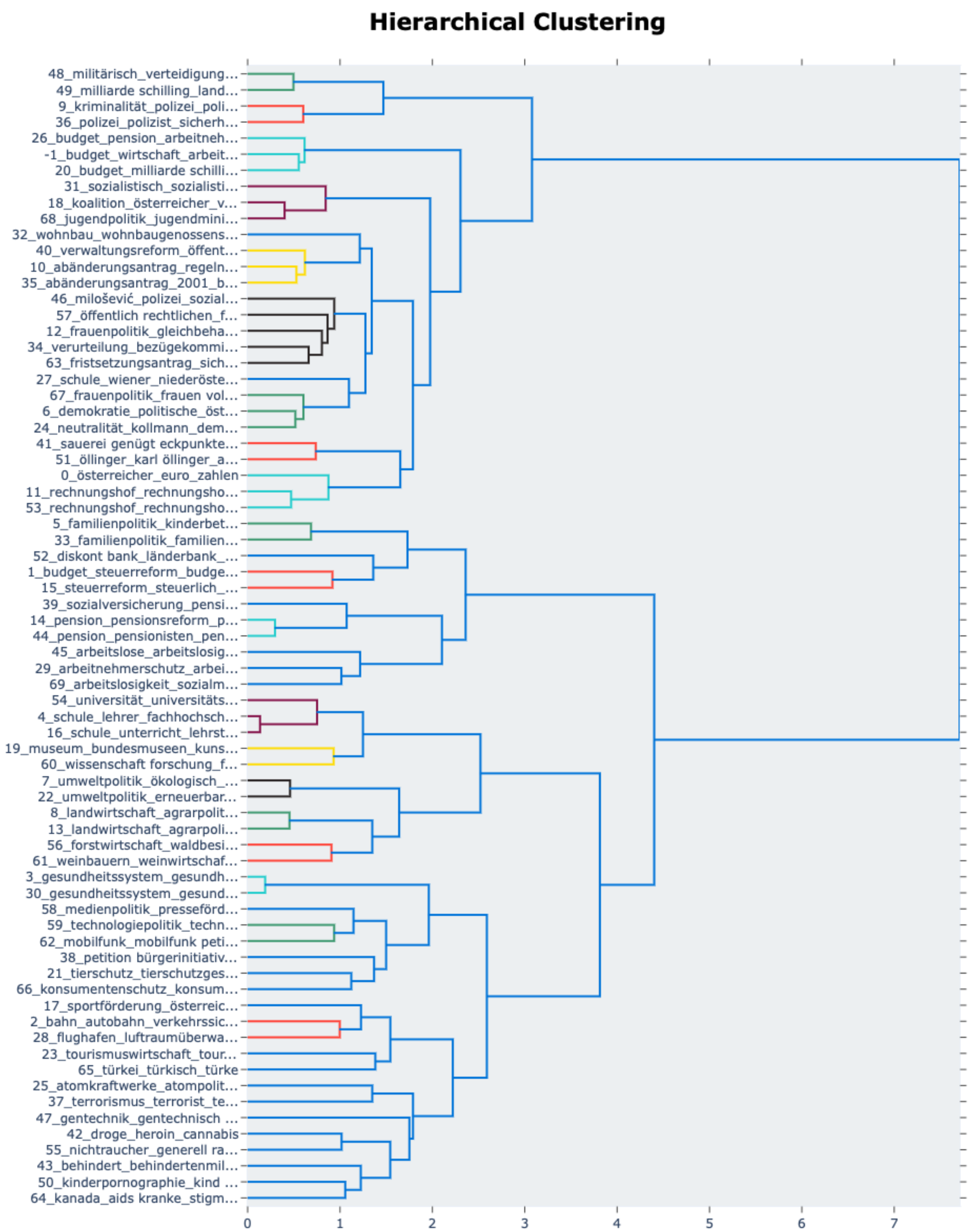
[27] „Sage“ becomes “sagen”, “bin” becomes “sein” etc.
[28] Parliamentary schedules have a couple very long breaks through the year. When looking at a more granular level, these breaks would negatively influence the coherence of the topics and model. By controlling at the year level, these breaks are controlled for.
[29] Additionally, the author of the topic model writes, “Make sure to use a limited number of unique timestamps (<100) as the c-TF-IDF representation will be calculated at each single unique timestamp. Having a large number of unique timestamps can take some time to be calculated. Moreover, there aren't many use-cased where you would like to see the difference in topic representations over more than 100 different timestamps.”

occur at least once a year in order to be included. Additionally, I set a parameter of diversity to .3. This is a [0,1] measure of how much overlap words can have between topics with zero meaning no overlap restrictions and a value of one meaning that words in one topic can't be included in another. The last parameter included is the number of words per topic, set at 6.

After running the first model, topics can be further reduced by looking at a topical hierarchy of the model. This shows topical groupings and allows for visualizations of how topics are related to one another. Figure 2 below presents this topical hierarchy. The numbers on the left hand column are the topic numbers while the words to the right are the tops words in the corresponding topic. Topic -1 is considered a cluster of outliers collected by the topic model. These are left out of future analyses. After viewing this topic hierarchy, topics are reduced that have a node before the number one node on the x axis.[30] This reduces the number of topics from 71 to 40.

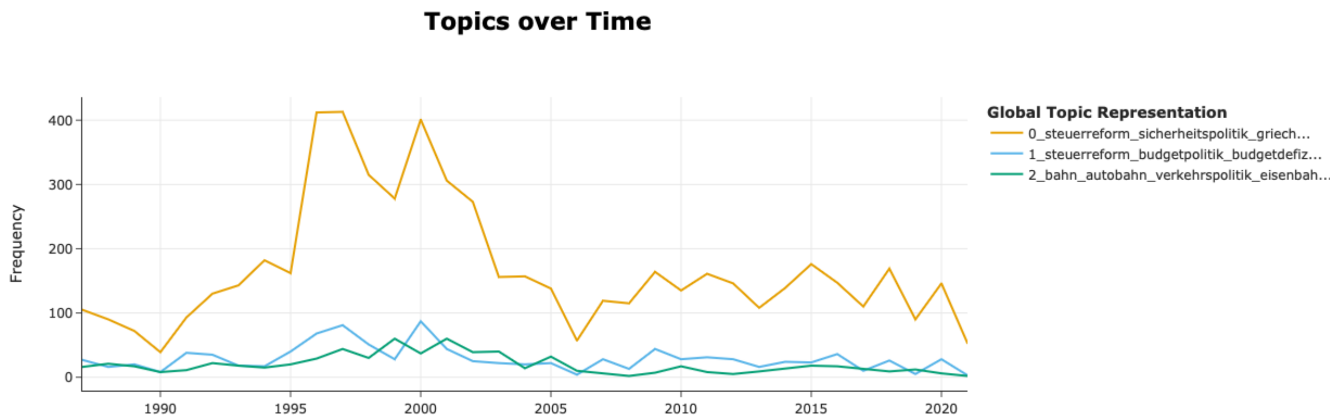
[30] I chose the number one on the X axis as the topics that join before this point appear very similar. By setting this at the number one on the axis I allow the model to be further adjusted by a more quantitative measure.

Figure 2: Topical Hierarchical Clustering



After adjusting the topic model, I rerun the model, this time controlling for the temporal nature of the dataset. This allows for an analysis of how topics change in nature and prevalence over the span of the 36 year dataset. Various topics from the model are visualized below. Topic occurrences is mapped on the Y-axis and refers to how often a topic is occurs in a FPÖ member speech in a given year. Topics are not presented all together due to the large level of noise when viewing a line graph that includes 40 number of topics with a differing range of topic occurrences.

Figure 3: Top Three Topics Mapped Over Time



The figure above plots the occurrences of the top three topics over the length of the dataset. Below are the tables with the corresponding topics listed and their top five terms through the length of the dataset in five year intervals. From the graph above there are visible fluctuations in topic occurrences. When viewing the corresponding topics and their makeup in the table below, one can see how the topics evolve through time. For example, when looking at topic zero, the topic begins with largely European/EU related words. In the 2011 interval this has been taken over by financial and sovereign debt crisis terms. This is also evident in topic one, a topic centered around budgetary terms. In the 80s and 90s this topic is largely made up of terms pertaining to the budget. In the 2000s, this shifts to words more centered around debt.

Topic zero	1987	1991	1996	2001	2006
1 2 3 4 5	europäische gemeinschaft österreichischen volkspartei österreichische volkspartei europapolitik europäisch gemeinschaft	europäische gemeinschaft entwicklungshilfe kroatien jugoslawien europäische wirtschaftsraum	pension währungsunion österreichisch wirtschaft wirtschaftsminister austria	kindergeld sicherheitspolitik schulde neutralität militärisch	europarat bulgarien europarates gewerkschaft rumänien
occurrences	105	93	412	306	57

Topic zero	2011	2016	2021
1 2 3 4 5	griechenland banken schuldenbremse volksabstimmung finanzministerin	türkei flüchtling asylwerber außenminister europäisch ebene	israel impfstoff hygiene austria mercosur abkommen corona kreditstundungen
occurrences	161	147	53

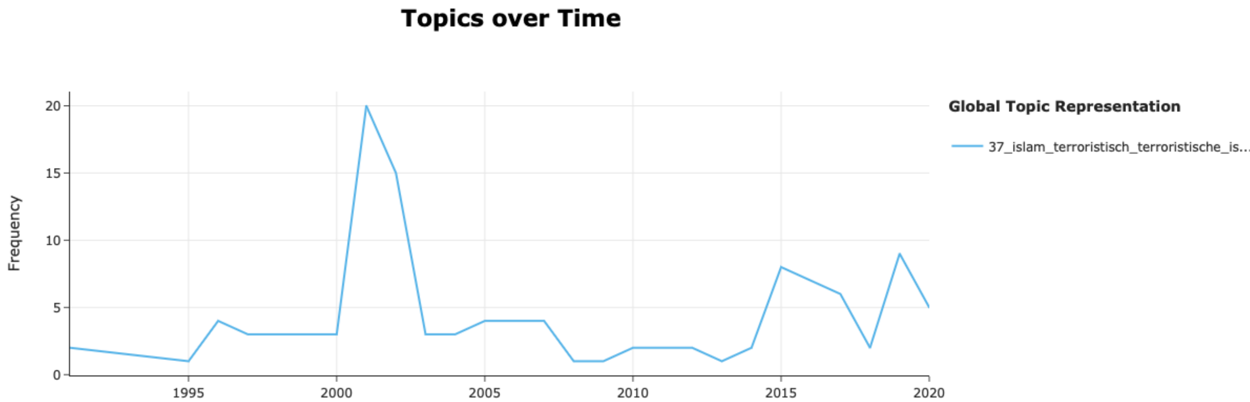
Topic one	1987	1991	1996	2001	2006
1	budgetpolitik	budgetpolitik	privatisierung	neuverschuldung	wirtschaftssprecher
2	budget 1987	budgetdebatte	steuerrecht	finanzpolitik	durchschnitt liegen lohnnebenkosten
3	budgetkonsolidierung	budgetsanierung	steuerlich	budgetpolitik	wirtschaftsthema
4	budgetdebatte	außerbudgetären finanzierung	budgetsanierung	budget 2002	job
5	budgetsanierung	budgetausschuß	parlamentarismus	budgetdefizit	deutsch nachbar
occurrences	27	38	68	44	4

Topic one	2011	2016	2021
1	schulde	schulde abgestuft	milliarde euro
2	schuldenbremse	bevölkerungsschlüssel	euro milliarde euro
3	finanzministerin	strukturell defizit	gemeinde 11
4	milliarde schulde	budgetrede	paar milliarde euro
5	schuldenabbau	budget 2017	milliarde euro milliarde
occurrences	31	36	3

Topic Two	1987	1991	1996	2001	2006
1	autobahn	semmeringbasistunnel	bahn	verkehrspolitik	verkehrssicherheit
2	bahn	bahn	autobahn	autobahn	geschwindigkeit
3	bundesbahnen	semmeringbasistunnels	semmering basistunnel	verkehrssicherheit	210 km gefahren
4	süd autobahn	eisenbahn	stadtautobahnen	straßenverkehr	verkehrspolitik
5	österreichischen bundesbahnen	südbahn	süd autobahn	generalverkehrsplan	zahl verkehrstoten
occurrences	16	11	29	60	10

Topic Two	2011	2016	2021
1	ybbstalbahn	autobahn	motorradfahrer
2	partnerschaft	bahn	motorisieren zweirad
3	eisenbahn	autofahrer	motorradfahrern
4	österreich schweiz	straßenverkehr	motorisieren
5	kombination bahn rad	verkehrstote	märz 2021
occurrences	8	17	2

Figure 4: Islam Topic Over Time



The graph presented above maps the topic of Islam over time. There is a sharp increase in topic occurrences after 9/11 and 2015 after the migrant crisis in Europe. Below is the topical makeup:

1991	2001	2009	2015
vereinte nation militärisch	terrorismus	terrorismus	islamist
nation militärisch	terrorist	kriminalität importieren	islam österreich
außenpolitisch grundlinie	kampf terror	kriminell	islamismus
vereinte nation krieg	terroristisch	freiheit sicherheit freiheitsinteressen	islamische
nation militärisch maßnahme	islamisch staat	sicherheitsbedürfnisse	radikal islamist

There are forty topics in the model. I have decided to share a few to show how topics talked about by the FPÖ have changed in prevalence and makeup through the 36 year dataset.

Further discussion:

As with most textual analyses, the dynamic topic model is not a perfect mode of analysis. The dynamic topic model is stochastic in nature and combines a mix of both quantitative and qualitative analysis during the parameter selection process to generate topics and determine their meanings. Additionally, the interpretation of the topic model results may be subject to bias. One labels the topics from their make-up of words and this may be different than if someone else were to do the same. Furthermore, the model assumes that there is only one topic per speech and therefore loses information included in the data. Although the model deduces the most important topic from large speeches, there is still a lot of information that is left out. This also applies when considering what the model does and doesn't understand from a linguistic perspective, such as:

“subtleties in the language such as irony and sarcasm, incomplete references to previous speeches and implicitly stated policy positions, e.g. distancing from another party’s position.”[31] Lastly, topic models could suffer from confirmation bias.[32] “If the results make sense in light of the knowledge that the researcher has about a topic or finding in the literature, they can be interpreted as a confirmation. But if the results suggest counter-intuitive conclusions, this is likely to be blamed on models supposedly not performing well.”[33]

This paper presented the results of a dynamic topic model on a novel dataset of all the Austrian national parliament debates from 1986 onwards. As this is the only paper I could find that uses a dynamic topic model to study populism, I hope that this paper serves as a template for further use of the model in populism studies. Although there are shortcomings in the model as discussed above, it provides researchers with a valuable tool that can read and analyze textual data a lot quicker than human capabilities.

As with the dynamic topic model, the dataset created for this project can play a role in future studies on populism, Austrian politics and other related subjects. As stated in the introduction, Austrian populism is an important subject in the realm of populism studies due to its long-running occurrence. Some future topics could be analyzing how center-right parties adapt to right-wing populist successes, mapping an Austrian political spectrum through the years and analyzing how the rhetoric in parliament has changed over the years.

---

[31] Müller-Hansen, Finn & Callaghan, Max & Lee, Yuan & Leipprand, Anna & Flachsland, Christian & Minx, Jan. (2021). Who cares about coal? Analyzing 70 years of German parliamentary debates on coal with dynamic topic modeling. Energy Research & Social Science. 72. 101869. 10.1016/j.erss.2020.101869.

[32] R.S. Nickerson, Confirmation bias: a ubiquitous phenomenon in many guises, Rev. Gen. Psychol. 2 (2) (1998) 175–220, https://doi.org/10.1037/1089- 2680.2.2.175.

[33] Müller-Hansen, Finn & Callaghan, Max & Lee, Yuan & Leipprand, Anna & Flachsland, Christian & Minx, Jan. (2021). Who cares about coal? Analyzing 70 years of German parliamentary debates on coal with dynamic topic modeling. Energy Research & Social Science. 72. 101869. 10.1016/j.erss.2020.101869.