Saying It How It Is? Assessing member, constituency and party alignment in parliamentary debate contributions

Nikhil Woodruff

Contents

1 Introduction **3**

2 Data sources and models **3**

3 Method **4**

3.1 Cleaning the dataset 4

3.2 Topical matches 4

4 Evaluation **5**

4.1 Party alignment 6

4.2 Constituency alignment 7

4.3 Assumptions 8

5 Conclusions **8**

1 Introduction

Much of political analysis is underpinned by the degree to which MPs, political parties and constituents are aligned in their views. While conventional wisdom suggests that parties and constituencies are aligned with MPs (who must be selected by both), the existence of both party deselection and re-election failure (combined with the fact that both of these are slow processes unlikely to coincide exactly with changes in members' views) suggests that this assumption is false at least some of the time. A key question therefore lies in the grey area between the two extremes: how often and to what extent are MPs, parties and constituents aligned?

This project aims to answer these questions by analysing the language used by MPs in parliamentary debate contributions. Using a cleaned and processed corpus of 180,000 words spoken by Members of Parliament in the House of Commons in 2022, I use topic modelling to determine the degree to which each spoken contribution matches with a particular category. I then combine this with demographic and socio-economic data on UK parliamentary constituencies to determine the degree to which MPs are aligned with the spoken constributions made by MPs with similarly characterised constituencies.

2 Data sources and models

This project uses a combination of data sources:

Hansard speech contributions Hansard is a UK government service maintaining a record of all parliamentary debate contributions made in the House of Commons and House of Lords, annotated with various features describing the speaker and the context. However, Hansard does not provide a data download service, and deploys anti-scraping measures on its website to prevent overuse. Therefore, I used datasets provided by TheyWorkForYou.com,[1] which makes available a set of XML datasets containing parliamentary debate contents - one file for each day.

UK parliamentary constituency demographics The Office for National Statistics provides a range of demographic and socio-economic data for each UK parliamentary constituency. This includes a dataset containing average hours worked, average income and house prices, percent of residents non-UK born and other similar demographic features.[2]

UK constituency election results Linking the parliamentary debate contributions to constituency demographics is not automatically possible, because MPs must be linked to their constituencies by name. To achieve this, I used the House of Commons Library's dataset of UK parliamentary constituency election results, which contains the name of the MP for each constituency in the 2019 election, alongside the ONS identifier for their constituency.[3]

BART BART is a language model developed by Meta AI, modelling the likelihood and distribution of sentence constructions in natural language in a transformer model. Pre-trained on a large corpus of text, it can be used on unseen examples with strong performance at various tasks (e.g. classification, extension, etc.).[4] Within this project, I used BART pre-trained with initial weights to procedurally extract topic matches from each spoken contribution in the Hansard dataset.

3 Method

The general method for this project is as follows:

1. Clean the speech contributions dataset. This involves combining all daily contribution datasets, extracting only speech by Commons MPs, standardising MP names and writing the resulting contributions to a single file.

2. Extract topical matches from the contributions dataset. Using pre-trained BART, identify a set of topics over which to match, and use the model to extract the probability distribution of each spoken text. This is primarily done by using the model to predict the likelihood of the "*this text is about X*", where X is a topic, immediately after the original contribution. This is then repeated for each topic, and the resulting probabilities softmax-weighted are used to determine the most likely topic for each contribution.

3. Analyse member alignment against political parties by comparing member topical matches against party averages. By calculating the average topical alignment score for every political party and topic, we can compare, on a given topic, the alignment scores of each MP (averaged over their contributions) against the average alignment score for their party. This allows us to determine the degree to which MPs are aligned with their party on a given topic.

4. Analyse member alignment against constituency averages. Using a similar method, we can examine to what extent MP scores are similar to those produced by a predictive model (for example, by feature correlation) trained on constituency demographics. This allows us to determine the degree to which MPs are aligned with their constituency on a given topic.

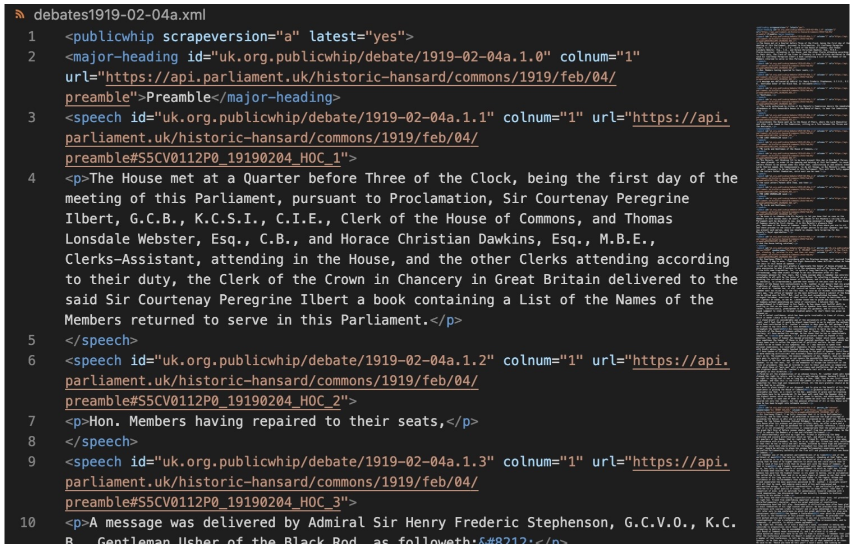
3.1 Cleaning the dataset

The first step in the process is to clean the dataset. Since the original dataset is split into daily files (an example in Figure 1), these must be concatenated into a single file. Each XML file contains a tag for speech, annotated with the author name and the actual speech contents. These can be extracted relatively simply.

Once extracted, these can be saved in a single CSV file containing a row for each contribution and the associated member. The code for this part is in the `fetch_speech_contributions.py` file (the `fetch_members.py` file is a separate script used to extract the list of all MPs directly from Hansard).

3.2 Topical matches

BART is a large model, and therefore each input takes around 10 seconds to process - longer depending on the number of possible categories, since each category needs to be run separately. To balance computation time against accuracy (and unbiased evaluation of individual members), I selected up to five random contributions from each member, and computed their categorical scores against the following categories:



```
debates1919-02-04a.xml
<publicwhip scrapeversion="a" latest="yes">
<major-heading id="uk.org.publicwhip/debate/1919-02-04a.1.0" colnum="1" url="https://api.parliament.uk/historic-hansard/commons/1919/feb/04/preamble">Preamble</major-heading>
<speech id="uk.org.publicwhip/debate/1919-02-04a.1.1" colnum="1" url="https://api.parliament.uk/historic-hansard/commons/1919/feb/04/preamble#S5CV0112P0_19190204_HOC_1">
<p>The House met at a Quarter before Three of the Clock, being the first day of the meeting of this Parliament, pursuant to Proclamation, Sir Courtenay Peregrine Ilbert, G.C.B., K.C.S.I., C.I.E., Clerk of the House of Commons, and Thomas Lonsdale Webster, Esq., C.B., and Horace Christian Dawkins, Esq., M.B.E., Clerks-Assistant, attending in the House, and the other Clerks attending according to their duty, the Clerk of the Crown in Chancery in Great Britain delivered to the said Sir Courtenay Peregrine Ilbert a book containing a List of the Names of the Members returned to serve in this Parliament.</p>
</speech>
<speech id="uk.org.publicwhip/debate/1919-02-04a.1.2" colnum="1" url="https://api.parliament.uk/historic-hansard/commons/1919/feb/04/preamble#S5CV0112P0_19190204_HOC_2">
<p>Hon. Members having repaired to their seats,</p>
</speech>
<speech id="uk.org.publicwhip/debate/1919-02-04a.1.3" colnum="1" url="https://api.parliament.uk/historic-hansard/commons/1919/feb/04/preamble#S5CV0112P0_19190204_HOC_3">
<p>A message was delivered by Admiral Sir Henry Frederic Stephenson, G.C.V.O., K.C.B., Gentleman Usher of the Black Rod, as followeth:&#8212;</p>
```


Figure 1: Example of a Hansard XML file containing a single speech contribution.

1. Health
2. Economy
3. Education
4. Housing
5. Transport
6. Environment
7. Crime

This creates a dataset of around 2,000 contributions, each labelled with the name of the MP, the full text and the probability scores for each of the seven categories. This dataset is stored in the `data/classified_speeches.csv.gz` file.

4 Evaluation

In order to ensure that the model is validly classifying speeches, I manually inspected many of the produced outputs: in particular, the highest-rating speech for each category. Below are some examples:

- Education:

> I was delighted to visit that excellent college in my hon. Friend's constituency and to see the fantastic work being done there. She will be pleased to know that we are investing £450 million of capital funding in higher education providers over the next three years, and that £400 million of that

> will be targeted on strategic priorities such as high-cost science, technology, engineering and maths and degree apprenticeships, for which providers can submit their bids until 27 June.
>
> -Alex Burghart MP

- Economy:

  > I am afraid the hon. Gentleman confuses what he is talking about. The fact that we have not hit the target is precisely a reflection of the fact that the wider economic recovery has been so strong. It is a measure of the success of the wider recovery that we simply do not need to offer those opportunities and that the regular economy is generating them.
  >
  > - Simon Clarke MP

- Transport:

  > In my constituency and across south-east London there is significant concern about the impact of the new December train timetable, which Southeastern drew up without any consultation with passengers, local rail user groups or elected representatives. Can we have a debate in Government time about the role of the Department for Transport in this planned alteration to the services my constituents rely on?
  >
  > - Matthew Pennycooke MP

These examples show that the model is correctly identifying the topic of the speech. However, I did find a peculiarity: the highest-scoring speeches for each category were disproportionately likely to be from Conservative members of Parliament. Upon further investigation, this seems rational: since the current government is Conservative led, all ministers of departments are Conservative, and ministers are (by their brief) more likely to only speak about the departmental category (shadow Ministers, those selected by the Opposition to specialise against a particular department, often do not have the same level of exclusivity for that topic). For example, current and previous Health Secretaries are far more likely than the average MP to talk about health-related topics, especially given that many sittings in Parliament are set up exclusively to discuss health, and ministers are required to answer questions from multiple other members.

4.1 Party alignment

By subtracting each MP's average topical alignment score from the average score for their party, we can determine the degree to which each MP is aligned with their party on a given topic. Table 1 shows the results for five randomly sampled MPs.

Table 1: Topical alignment (over their party) for five randomly sampled MPs.

	crime	economy	education	environment	health	housing	party	speaker	transport
175	-0.059223	0.036948	-0.034855	-0.001007	0.006183	0.008541	Democratic Unionist Party	Jim Shannon	0.043411
378	-0.092109	-0.065141	-0.030603	0.018984	0.204263	-0.025266	Labour	Wes Streeting	-0.010129
19	-0.045081	-0.040032	-0.055748	-0.014195	-0.061154	-0.028559	Conservative	Andrea Leadsom	0.244768
101	-0.036316	-0.027142	-0.016164	-0.032298	-0.012530	-0.000944	Conservative	Dean Russell	0.125392
357	0.184726	-0.073256	-0.056837	-0.039989	-0.052649	-0.040366	Conservative	Suella Braverman	0.078371

The table is encouraging as a sanity check: Wes Streeting MP has been the Shadow Health Secretary (with a brief focused exclusively on health) since 2021. Suella Braverman MP is the current Home Secretary, with a large subset of her jurisdiction over crime.

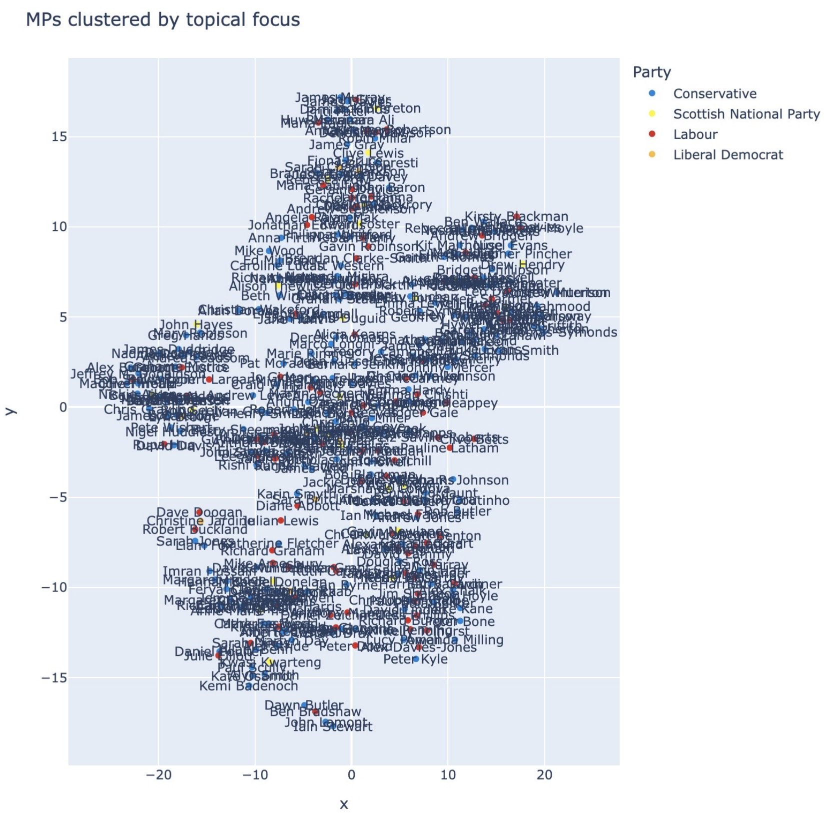


Figure 2: MPs clustered by their topical alignment (coloured by party).

Another way to assess model performance is to examine the patterns in the topical focus scores generated for each MP. Figure 2. Using the t-SNE algorithm[5] to reduce the dimensionality of the data, we can see that MPs fall into distinct groups, with an interesting finding that these groups largely don't correlate with party affiliation. This is unexpected, though one possible reason could be that parties are largely successful at ensuring diversity of focus among MPs in order to adversarially engage with opposition on each individual topic.

4.2 Constituency alignment

To assess how well the model predicts alignment with constituency demographics, we can inspect correlation coefficients within each demographic feature.

Table 2 shows these correlation coefficients. Many of the results are expected: for example, the strongest positive correlation is that salary correlates with degree attainment, and one of the strongest negative correlations is that foreign born percentage correlates

Table 2: Correlation coefficients between demographic features and topical scores.

	salary	degree	age	nonukborn	health	transport	housing	education	crime	economy	environment
salary	1.000000	0.743145	0.012033	0.364434	-0.034027	0.033573	-0.033951	0.066623	-0.046468	-0.011793	0.005955
degree	0.743145	1.000000	0.032226	0.453349	-0.061031	0.046563	-0.060737	0.024258	-0.003536	0.007305	0.025134
age	0.012033	0.032226	1.000000	-0.010166	-0.041228	0.082237	-0.035106	0.034356	-0.046669	0.031511	-0.084613
nonukborn	0.364434	0.453349	-0.010166	1.000000	-0.016459	-0.040405	0.076939	-0.096774	0.142469	0.020454	-0.074010
health	-0.034027	-0.061031	-0.041228	-0.016459	1.000000	-0.300552	-0.036717	-0.081930	-0.171166	-0.240301	-0.120114
transport	0.033573	0.046563	0.082237	-0.040405	-0.300552	1.000000	-0.222865	-0.271599	-0.219707	-0.168405	-0.161178
housing	-0.033951	-0.060737	-0.035106	0.076939	-0.036717	-0.222865	1.000000	-0.172637	-0.125502	-0.096848	-0.024377
education	0.066623	0.024258	0.034356	-0.096774	-0.081930	-0.271599	-0.172637	1.000000	-0.155610	-0.243157	-0.174008
crime	-0.046468	-0.003536	-0.046669	0.142469	-0.171166	-0.219707	-0.125502	-0.155610	1.000000	-0.209344	-0.134598
economy	-0.011793	0.007305	0.031511	0.020454	-0.240301	-0.168405	-0.096848	-0.243157	-0.209344	1.000000	-0.023027
environment	0.005955	0.025134	-0.084613	-0.074010	-0.120114	-0.161178	-0.024377	-0.174008	-0.134598	-0.023027	1.000000

negatively with environmental focus (since foreign-born usually correlates with urban areas, which are less likely to be environmentally conscious than rural areas). However, some results are counter-intuitive: for example, salary negatively correlates with economic focus: a common prior is that higher earners would be more exposed to changes in the economy.

4.3 Assumptions

There are several assumptions involved here:

- *Spoken contributions are an unbiased sample of an MP's inner views.* This is likely to be reasonable, in the absence of any evidence that politicians are more likely to speak about certain topics than others. However, it is possible that some bias is introduced here: MPs could feasibly be withholding contributions that relate to some topics more than others (national security, for example).

- *Context does not affect the topic alignment of a given speech contribution.* By this, I mean that a spoken contribution in one sitting has exactly the same consequence for an MP's topical focus as if the MP had spoken it in any other sitting. This is likely to be a safe assumption, but different sittings might plausibly influence what we should derive from speech: for example, Prime Minister's Question, ministerial questions settings, and Select Committees all have particular purposes and focuses that might influence what is said. That said, MPs do have (relatively) free choice about whether to attend these sittings, and so it is reasonable to assume that they are not being forced to speak about certain topics.

- *The number of individual speeches on a topic is more important than the combined length of the speeches.* This assumption is plausibly likely to be false. It occurs in this implementation because of the random sampling of an MP's speeches for input into the topic alignment model (for example, if an MP just repeated the word 'economy' at various interruptions within a sitting, and also spoke for five hours about transport once, it's far more likely that their inputs for analysis would be overweighted by the economic focus). Most would likely agree that the MP is more focused on transport than the economy. A better approach might have been to use the length of the speech as a weighting factor, but this would have required a more complex implementation.

5 Conclusions

NLP models for topical analysis have often been deployed, but given the relative size of the Hansard database, it's likely the accuracy and coverage of this experiment could be far improved by using older years' contributions, or better pre-processed inputs. However, even the current results suggest interesting insights: while many of the highest-scorers are government ministers in their respective categories, it seems very plausible that topical focus among non-government MPs could be a strong predictor of future appointment to ministerial positions (or, to the shadow cabinet).

The specific NLP model was chosen based on its large training and testing performance in the literature, but other models might have been faster. Future research could examine whether a faster, but less accurate, topical analysis model might produce more valuable insights by nature of being able to evaluate more speech data in the same time period.

Clustering techniques were highly useful here in visualising and understanding the consequence of the topical modelling process. The t-SNE algorithm was able to effectively reduce the dimensionality of the data, making clear the relevant clusters of MPs with a similar topical focus average across their speeches. This could be useful in future research aimed at clustering MPs by other high-dimensional data, such as voting records.

Overall, this project has demonstrated the potential of NLP models to analyse MPs' speech, showing that speech is a key predictor of role (explicit or implicit) within party hierarchies for focus on a particular topic. Future research could further examine this connection between speech and role over a greater period of time, or as an input for other interesting research questions.

References

[1] TheyWorkForYou.com, "Theyworkforyou.com," *TheyWorkForYou.com*, 2023. [Online]. Available: https://www.theyworkforyou.com/

[2] ONS, "Visualising your constituency," *ONS*, 2023. [Online]. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/housing/articles/visualisingyourconstituency/2015-03-26

[3] H. of Commons Library, "Constituency data: election results," *House of Commons Library*, 2020. [Online]. Available: https://commonslibrary.parliament.uk/constituency-data-election-results/

[4] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, and L. Zettlemoyer, "BART: denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension," *CoRR*, vol. abs/1910.13461, 2019. [Online]. Available: http://arxiv.org/abs/1910.13461

[5] L. van der Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of Machine Learning Research*, vol. 9, no. 86, pp. 2579–2605, 2008. [Online]. Available: http://jmlr.org/papers/v9/vandermaaten08a.html