Online Appendix
Word Embeddings for the Analysis of Ideological Placement in Parliamentary Corpora

Ludovic Rheault[†] and Christopher Cochrane[‡]

Additional Information on the Corpora

This section provides details on the three corpora used in the main text. Table A1 reports summary statistics on each corpus. The Canadian Hansard corpus is a public resource accessible on the www.lipad.ca website and described further in Beelen et al. (2017). The resource is enriched with metadata about speakers and attributes such as party affiliations and functions. It covers a time-period ranging between February 6, 1901 and October 30, 2018. Before fitting our models, we removed procedural interventions from the corpus, which are not associated to any politician. As explained in the text, we exclude speeches made by the Speakers of the House or the acting speaker. Speakers in Britain and Canada are presiding the parliamentary proceedings, and do not take part in substantive or partisan debates. This position differs markedly from that of Speaker of the House of Representatives in the United States, who is normally the highest ranking member of the House and the parliamentary leader.

The British Hansard corpus can also be accessed online via the Political Mashup website, and covers a period from November 26, 1935 to March 11, 2014. As we did for Canada, we removed speeches from the Speaker. For Britain, we investigated the accuracy of party affiliations, which are based on historical databases of members of parliament. To avoid a source of error, we removed cases of members who crossed the floor and for which the dates of the change in party affiliation was missing. We also manually corrected the affiliation of a number of cabinet members after inspection of cross-tabulations between party affiliations and parliamentary functions. The final corpus comprises speeches made by the three major parties in the British parliament, excluding the Speaker of the House.

[†]Assistant Professor, Department of Political Science and Munk School of Global Affairs and Public Policy, University of Toronto.

[‡]Associate Professor, Department of Political Science, University of Toronto.

Table A1: Size of the Preprocessed Corpora

Corpus	Time-Range	Speeches	Sample Size (Words)	Vocabulary Size
Britain	1935–2014	3.4M	224M	93,919
Canada	1901–2018	3.0M	196M	78,856
US House	1873–2016	6.8M	339M	109,967
US Senate	1873–2016	6.3M	305M	115,178

The corpus statistics are computed after performing the preprocessing steps described in this section.

Finally, the US corpus is the version released by Gentzkow, Shapiro, and Taddy (2018).[1] From the 43rd to the 111th Congress, the data come from the bound edition of the Congressional Record, whereas the last three Congresses are taken from a compilation of the daily edition. A detailed codebook describing the resource is available with the release. We use the corpus as distributed, and provide a replication script to recreate the format necessary to run our models. Our version of the corpus is limited to speeches from Democrat and Republican voting members, which represents the large majority of all available speeches in both chambers. As explained in the text, we fitted models separately for the Senate and the House. They can also be used in combination.

Scholars can reuse the source code used to implement our analysis. A Python module is released publicly on the GitHub website (https://www.github.com/lrheault/partyembed) and the results presented in this paper can be reproduced with the materials available in the Political Analysis Dataverse (Rheault and Cochrane 2019).

Phrase Detection

We fit our models after performing phrase detection on each corpus, using pointwise mutual information scoring and the threshold using in Mikolov, Sutskever, Chen, Corrado, and Dean (2013). We rely on the implementation from the gensim library for Python, and run two passes of the algorithm on the corpus. While this step is not strictly necessary to use the proposed methodology, we find that it facilitates interpretation. To illustrate phrase detection, Table A2 reports the 20 most frequent phrases for each of the three parliamentary corpora.

[1] https://data.stanford.edu/congress_text

Table A2: Most Common Phrases

USA		Canada		Britain	
Phrase	Count	Phrase	Count	Phrase	Count
united states	2080539	united states	306360	local authorities	212686
unanimous consent	759100	british columbia	115674	united kingdom	157496
new york	623229	years ago	84620	northern ireland	111289
fiscal year	283949	great deal	65044	local authority	103896
years ago	249396	nova scotia	55670	great deal	97110
health care	238898	income tax	55663	right learned	87797
supreme court	197027	national defence	53014	united states	82524
conference report	187871	health care	50308	white paper	76964
social security	184999	province quebec	46955	select committee	73220
printed record	168608	public works	46932	years ago	72275
district columbia	152233	great britain	40841	chancellor exchequer	62441
new jersey	143234	united nations	40646	young people	62249
north carolina	132844	post office	40492	long term	58142
joint resolution	131569	wheat board	40208	past years	55609
majority leader	130438	unemployment insurance	38373	second reading	54659
great deal	128001	auditor general	38128	make statement	53543
soviet union	120680	bloc quebecois	37628	private sector	43884
small business	110462	human rights	36488	united nations	42912
south carolina	108605	new brunswick	36076	post office	41582
united nations	101989	standing committee	35570	home office	36389

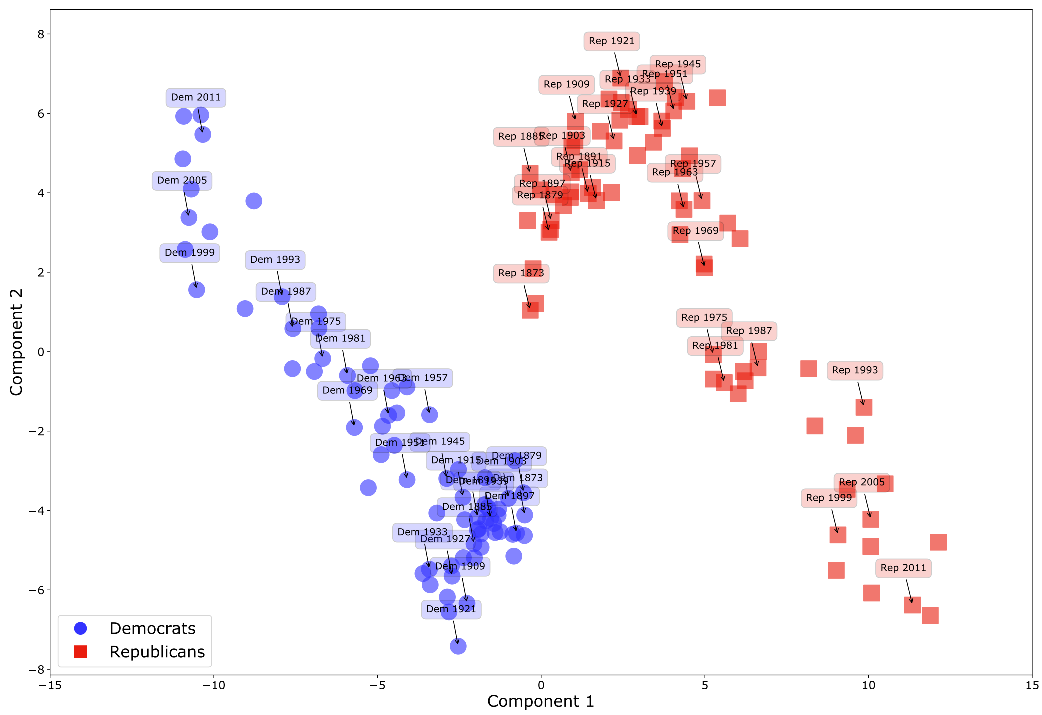
Count of the 20 most frequent phrases (collocations) automatically detected in the three main corpora.

Senate Corpus

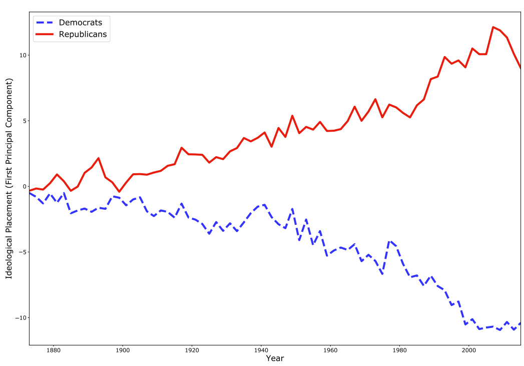
Figure A1 reproduces with the US Senate corpus the visualizations presented in the main text for the US House of Representatives. As can be seen, the patterns are virtually identical to those discussed in the text for the House. Our party embeddings capture an increasing polarization on the x-axis, which we interpret as the left-right (or liberal-conservative) dimension. Meanwhile, the parties switch places on the South-North dimension on the y-axis, with the Republicans becoming closer to the South edge of the spectrum in the recent era.

Figure A1: Party Placement in the US Senate (1873–2016)

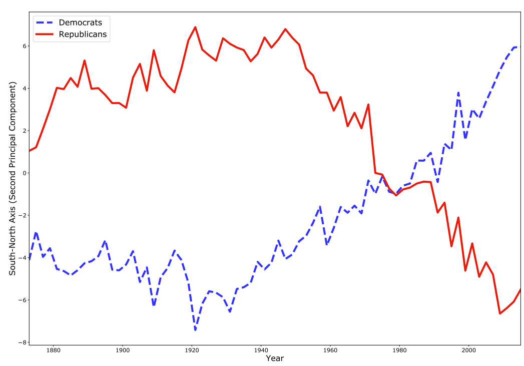
(a) Two-Dimensional Projection



(b) First Dimension



(c) Second Dimension



The figure shows a 2-dimensional projection of the two principal components of party embeddings for the US Senate (a), and time-series plots for each of the two components separately in (b) and (c).

Linguistic Specificity of US Political Parties

Another way to interpret the party placement derived from our methodology consists of retrieving concepts that are semantically associated with specific parties. For instance, we can readily identify the expressions closest to the position of the Democrats in the vector space for the 114th Congress, by retaining word embeddings having the highest cosine similarity with that specific party embedding. This does not require any technique for dimension reduction, as the similarity scores can be computed from the original, M-sized embeddings. We report in Table A3 the top 20 words ranked as most similar to each party for the House of Representatives, searching within the 20,000 most frequent terms in the corpus vocabulary. The top words for the Democrats contain relevant hints at a liberal stance, with concepts such as "gun violence" and "environmental protection". On the other hand, the discourse of Republicans is semantically closer to concepts such as bureaucracy and ideologically-laden expressions such as "overregulation". The lists also contain named entities. These could be pruned out, although the locations and persons mentioned may themselves have a substantive interest in applied research.

It is important to note the difference between this approach and methods for identifying linguistic specificity based on actual word occurrences (e.g. Monroe, Colaresi, and Quinn 2008). Word embeddings are models of meaning representation, which implies that expressions appearing in the Table may not have been uttered during that Congress per se. The ranking reflects that a party's speeches are semantically similar to the listed words and phrases. Like other methods such as latent semantic analysis, the calculation of similarity scores does not rely on string matches.

Table A3: Words and Phrases Most Similar to Democrats and Republicans in the 114th Congress

Democrats		Republicans	
Expression	Cosine Similarity	Expression	Cosine Similarity
gentlewoman california	0.383	overregulation	0.352
congressional black caucus	0.348	obamacare	0.350
latinos	0.345	nebraska	0.322
latino	0.330	chris	0.320
black caucus	0.307	troops forget september	0.314
protections	0.307	bureaucrats	0.304
progressive caucus	0.305	bureaucracy	0.302
gun violence	0.282	job creators	0.298
oakland	0.280	regulates	0.298
decent housing	0.278	bureaucratic	0.267
congresswoman	0.272	nelson	0.266
houston texas	0.267	overreach	0.266
gonzalez	0.265	big brother	0.262
ryan	0.264	mentioned earlier	0.261
environmental protection	0.264	checkbook	0.259
san francisco	0.263	headquartered	0.256
stocks bonds	0.262	overzealous	0.252
brooklyn	0.261	southeast	0.250
los angeles	0.260	rein	0.248
slashing	0.259	bureaucracies	0.248

Note that the corpus excludes common stop words, which facilitates the identification of phrases. The words "troops forget september", for instance, was detected as being part of a common utterance that usually takes the form "God bless our troops, and we will never forget September 11."

Guided Projections

Instead of interpreting dimensions ex-post, researchers may also choose to define axes of interest. In this section, we briefly illustrate how the proposed methodology can be used in such fashion. We start by choosing expressions representative of opposite ideological stances on economic and social issues (see Table A5 for the full list). When more than one term is used to anchor a position, we can take the centroids of each group of words and phrases, by averaging their word embeddings. Finally, axes are created by taking the difference between the right and left centroids, for each dimension of interest. We project party embeddings onto the customized space by taking dot products:

$$\zeta \cdot \left(\frac{\sum_{i \in L_{\text{Right}}} \boldsymbol{\beta}_i}{V_{\text{Right}}} - \frac{\sum_{i \in L_{\text{Left}}} \boldsymbol{\beta}_i}{V_{\text{Left}}} \right)$$

where L_{Left} is the chosen lexicon for words identifying the left-wing, and V_{Left} the size of that lexicon (and similarly for the Right).[2]

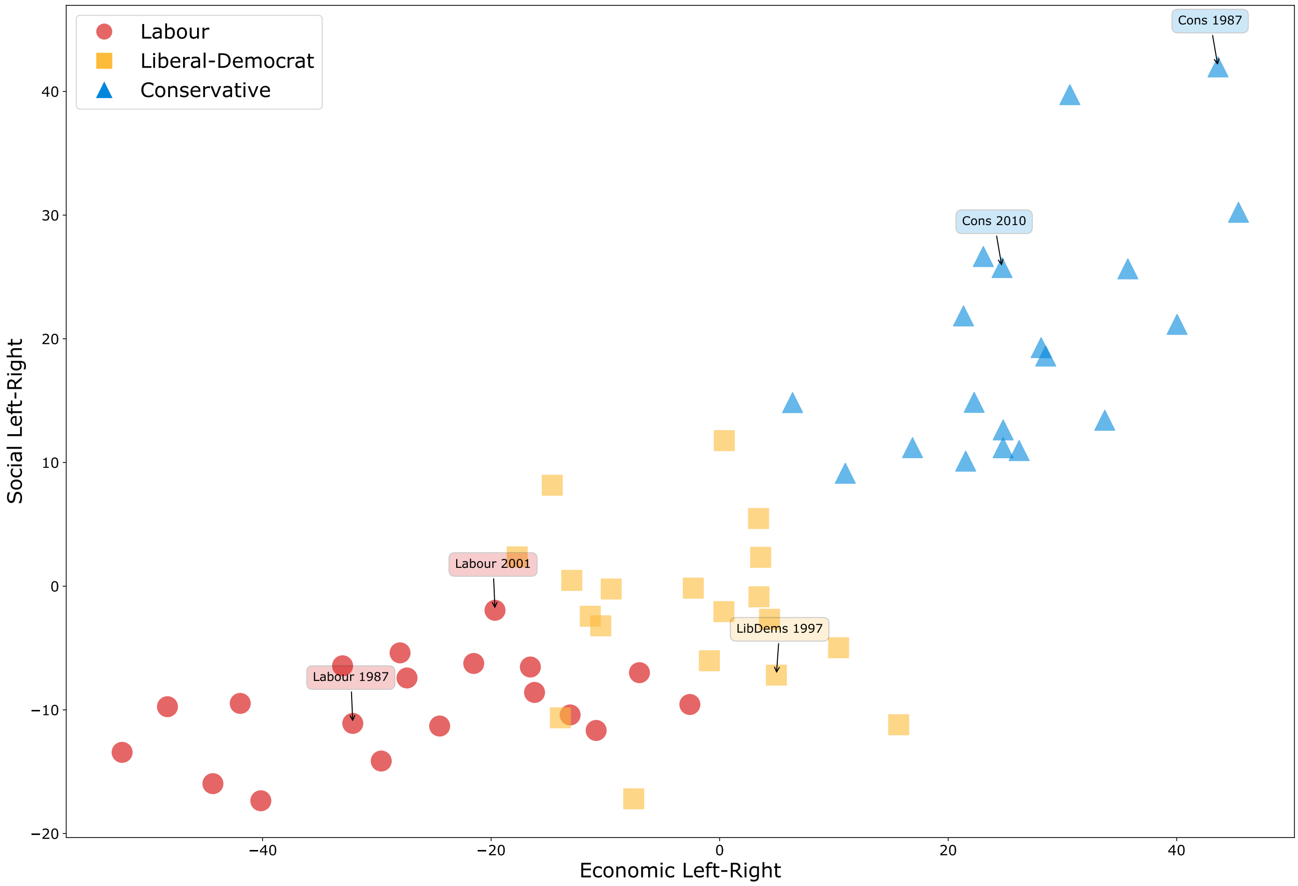
Figure A2 illustrates such a linear projection of party embeddings in a two-dimensional space for the British corpus. The neural network model is the same as that used in the main text. The social dimension (y-axis) uses expressions such as "civil rights" and "traditional values" to represent left and right, respectively. For the economic dimension (x-axis), we use concepts related to workers and redistribution for the left, and expressions such as "businesses", "taxpayers" and "free enterprise" for the right. Consistent with expectations, the figure suggests that the Conservative party is located to the right on both the economic and social dimensions in recent decades. Labour and Liberal-Democrats, on the other hand, appear both socially on the left, but the Labour party is further to the left on the economic axis.

Validation Tests: Guided Approach

Table A4 reports accuracy results for our main models when using the guided method. We rely on a common list of expressions to define the left and right (Table A5). Our two-dimensional projection in Figure A2 considered the economic and social dimensions separately. For the accuracy tests in this section, we combine words from both dimensions on a single ideological axis. We use the same expressions for all countries. These expressions were manually chosen by us, and correspond to concepts that we theoretically expect to be associated with the language of left-wing and right-wing parties. As argued in the main text, ideology cannot be easily reduced to a group of words or phrases. Choosing expressions on the basis of one's judgment entails a risk of leaving out important components of ideology. But the guided approach may have practical

[2] This approach expands on a standard visualization technique for the analysis of word embeddings; for instance, a similar implementation is included in Google's *TensorBoard* tool.

Figure A2: Party Placement in a 2D Space using Customized Ideological Axes (Britain)



applications. In particular, researchers could use other types of lexicons representing political concepts of interest, not just ideology (for instance sentiment, specific issues, and so on).

Table A4 relies on some of the the same gold standards used in the main text, and can be compared with the accuracy of the PCA approach from Table 2 in the main text. In nearly all cases, the fit is not as accurate as the one reported previously. We also tested economic and social dimensions of left and right separately, but doing so does not improve relative to the results in Table 2. We conclude that a completely unsupervised method relying on principal component analysis produces results that are probably as accurate, if not more, than a search for the "correct" list of ideological words.

Table A4: Accuracy of Guided Party Placement against Gold Standard

Gold Standard	Metric	US House	US Senate	Canada	Britain
Voteview	Correlation	0.831	0.759		
	Pairwise Accuracy	85.37%	80.26%		
rile	Correlation	0.600	0.582	0.733	0.715
	Pairwise Accuracy	72.16%	67.29%	75.13%	75.19%
vanilla	Correlation	0.678	0.596	0.723	0.782
	Pairwise Accuracy	75.35%	69.59%	76.30%	78.83%
legacy	Correlation	0.806	0.767	0.855	0.790
	Pairwise Accuracy	84.22%	81.65%	81.68%	77.71%

The guided approach relies on the expressions for the in Table A5. For the United States, we use the average party score on the first dimension of the Voteview DW-NOMINATE estimates (1921-2016). Accuracy is assessed against the same three measure based on the CMP data from the main text (1945-2015 for UK and Canada; 1920-2012 for the USA).

Table A5: Words and Phrases for Guided Ideological Placement

Economic Left	affordable housing, decent housing, eradicate poverty, poverty, gap rich poor, wealthiest, low income, inequality, unequal, workers, minimum wage, unemployment, unemployed, protective tariff, redistribution, redistribution wealth, safety net, social security, homelessness, labor unions, labour unions, trade unions, working classes
Economic Right	decentralization, bureaucracy, business, businesses, creating jobs, job creators, free enterprise, free trade, private enterprise, private sector, debt relief, debt reduction, taxpayers, taxpayers money, taxpayer money, commerce, privatisation, privatization, competitive, industry, productivity, deficit reduction, hard working, hardworking, home owners, homeowners, open market, free market, private enterprise, private sector, property rights, property owners
Social Left	minority rights, gay lesbian, affirmative action, employment equity, pay equity, racial minorities, racism, gun control, minorities, prochoice, pro-choice, civil rights, environment, greenhouse gas, pollution, climate change, child care, childcare, planned parenthood, access abortion
Social Right	law enforcement, moral fabric, social fabric, moral decay, moral values, sentences, tougher sentences, traditional values, tradition, secure borders, illegal immigrants, illegal immigration, criminals, fight crime, prolife, pro-life, sanctity life, unborn child, abortionist, church

The Table reports custom lists of words to define a left-right (liberal-conservative) ideology on two dimensions, economic and social. These words were used to produce the two-dimension projection in Figure 4. To compute accuracy tests, we collapse economic and social categories into two lexicons, for the left and right. We use the list as is for each country and deliberately include alternative spellings. When fitting the models, words absent from the vocabulary are automatically excluded. For instance, "labour unions" does not appear in the American corpus, but "labor unions" does.

Example Application: Party Polarization

Another benefit of having estimates of party placement in a vector space is the possibility of computing quantities of interest based on metrics for vector distances. An obvious application of such metrics is the measurement of the degree of party polarization in a legislature over time. A recurrent finding in the American politics literature is the increasing level of ideological polarization of political parties in the modern era (for discussions, see McCarty, Poole, and Rosenthal 2006; Layman, Carsey, and Horowitz 2006; Abramowitz and Saunders 2008; Dalton 2008; Lee 2015). In Canada, signs of expanding levels of ideological diversity can also be found in the party platforms at least since the 1980s (Cochrane 2015). In Britain, however, previous research suggests that parties have depolarized since the Thatcher era (Clarke et al. 2009; Adams, Green, and Milazzo 2012). Below, we reassess the hypothesis of polarization in the three countries, using the same models previously introduced.

Several metrics can be used with the embeddings to measure the distance between the language of political actors. One of the simplest is the Euclidean distance d_{ij} between two vectors $\boldsymbol{\zeta}_i$ and $\boldsymbol{\zeta}_j$, which is obtained as:

$$d_{ij} = \sqrt{\sum_{m=1}^{M} (\zeta_{im} - \zeta_{jm})^2} \quad (1)$$

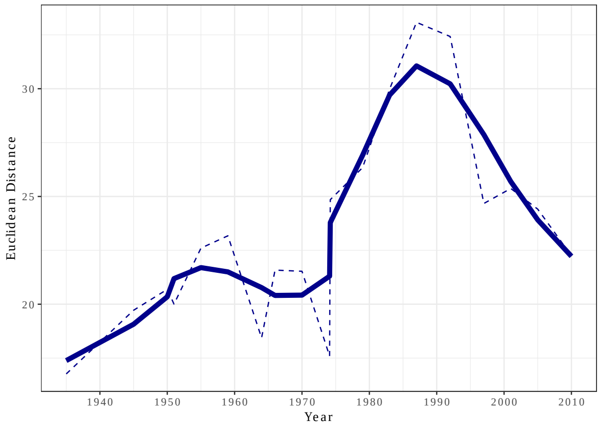
For instance, we can use the party embedding for the Republicans ($\boldsymbol{\zeta}_i$), and measure its Euclidean distance with the corresponding vector for the Democrats ($\boldsymbol{\zeta}_j$) in a given Congress. Other distance metrics have gained in popularized for the analysis of word embeddings, such as Word's Mover Distance (WDM) (Kusner et al. 2015). WMD measures the shortest path required to transform the words of a first document into the words of another document. The metric could be utilized for a variety of analyses using our model's word embeddings. For simplicity, we focus on Euclidean distance in what follows.

To examine party polarization in the United Kingdom, we adopt a definition similar to Peterson and Spirling (2018). That is, we assess polarization as the distance between the ideological positions of the two parties having formed the government since the mid 20th Century, Labour and Conservatives. Figure A3a plots the Euclidean distance between the two party embeddings over time. The pattern is consistent with the expectation of a depolarization, and reflects some of the findings introduced earlier in Figure 3a of the main text. We observe that speeches in the House are most distinct in the period starting with the Parliament after the second general election of 1974, during the Thatcher governments. The gap between the two major parties' ideological placement is emphasized clearly until the 1997 election that brought the Labour party back in power.

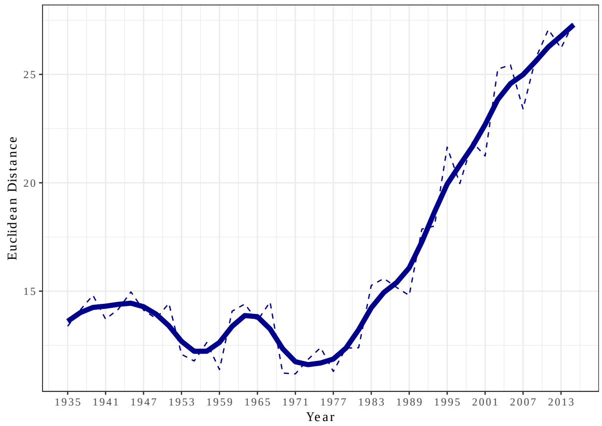
For the US House of Representatives (Figure A3b), we find clear evidence of ideological po-

Figure A3: Party Polarization in Britain, Canada, and the United States (1935–2015)

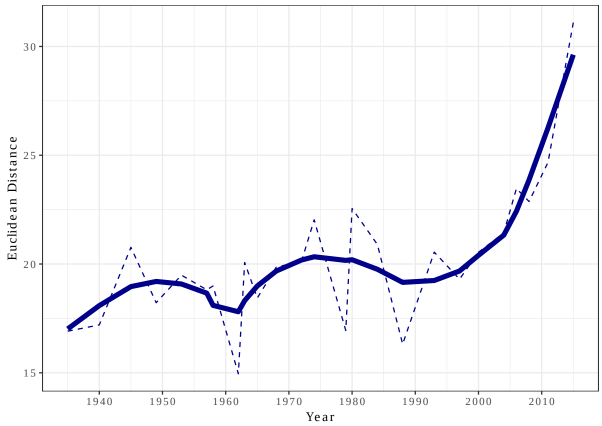
(a) Britain



(b) US House



(c) Canada



Polarization is measured using the Euclidean distance between the party embeddings of the Labour and Conservative parties for Britain, Liberal and Conservative parties for Canada, and Democrat and Republicans in the US House. The thick lines are smooth splines of the raw Euclidean distances.

larization in the recent decades, as was also apparent by observing the trajectory of parties in Figure 2 in the main text. Our results are consistent with other findings from the literature. For instance, our model captures a dip in the levels of partisanship during the 1970s, before the current period of polarization, a trend also discussed in Levendusky (2009). Turning to the Canadian case, we measure polarization as the Euclidean distance between the Liberal and Conservative party embeddings across parliaments, the two parties having formed the government. Figure A3c depicts the trend in party polarization in Canada. The data support the claim of an increasing polarization in recent decades.

Choosing Hyperparameters

Fitting models of word embeddings requires setting a number of hyperparameters. In this section, we discuss the impact of such parameterization on model accuracy. Overall, we find that using proposed default values from earlier studies leads to reliable results.

The principal decisions in terms of parameterization concern the number of nodes in the hidden layer (denoted M in the text), which determines the vector size for the embeddings, and the size of the context window (which we denoted Δ in the text). We examined models with hidden layers of 100, 200, and 300 dimensions, and report accuracy results based on some of the gold standards already described in the main text (Table A6). These are three values commonly used in the literature on word embeddings. We find that models with 200 dimensions offer a good compromise in terms of accuracy, for the three countries under consideration. This is the vector size used for the models discussed in the main text. Regarding the window size, we rely on values that are slightly larger than usual, a choice driven by the typical length of parliamentary speeches. For Britain, our results indicate that a larger window size (around 30 words) can improve accuracy marginally, whereas the opposite holds with both the Canadian and American corpora, for which windows of 15 or 20 words performed well. We used a value of $\Delta = 20$ for implementations in the main text.

Table A6: Effect of Layer Size on Accuracy

Corpus	M	Pearson Correlation	Pairwise Accuracy
House	100	0.845	86.623%
House	200	0.918	85.658%
House	300	0.931	84.408%
Senate	100	0.887	85.263%
Senate	200	0.919	83.925%
Senate	300	0.869	83.991%
UK	100	0.864	81.831%
UK	200	0.876	82.669%
UK	300	0.870	82.390%
Canada	100	0.814	79.589%
Canada	200	0.855	79.778%
Canada	300	0.856	79.684%

The evaluated placements are obtained using the first principal component of party embeddings, for various hidden layer sizes (M). For Canada and the UK, accuracy is assessed against the Legacy measure based on the Comparative Manifestos Project (CMP) data (1945-2015). For the United States, we use the average party score on the first dimension of the Voteview DW-NOMINATE estimates (1921-2016).

Other hyperparameters involved in the estimation can be modified, in particular the learning rate and the number of epochs—that is, how many times the estimation algorithm cycles through the full set of training examples. To illustrate their impact on the results, we tested a large number of combinations and assessed the impact on the quality of the models. For both the learning rate and epochs, we found evidence of a concave relationship between these parameters and model accuracy based on our gold standards. In simple terms, values of the learning rate set too low or too high tend to reduce accuracy, and similarly for the number of epochs. A learning rate between 0.1 and 0.025 generated better results, and accuracy is only marginally improved by increasing the number of epochs from 5 to 10. The models used in the main text rely on a learning rate of 0.025 and 5 epochs, which are both default values in the implementation of the algorithm used to fit the models. Our conclusion is that modifying these default values is probably not warranted, except when the number of speeches available for each political actor decreases. In the latter case, increasing the number of epochs will improve model accuracy. The last tables in this appendix report an extended set of accuracy results for various combinations of parameters (Tables A9-A11).

Researchers should be wary that points estimates for word embeddings are probabilistic. Reordering the examples and training for a longer period of times (by increasing the number of epochs) will not return identical embeddings. Like other popular approaches such as Bayesian analysis, repeated runs of the models will return slightly different values. Nonetheless, as long as the model is properly parameterized, quantities of interest such as ideological placement and cosine similarities will be very similar from one run to the next.

Evaluating Word Embeddings Trained on Parliamentary Corpora

Table A7 reports accuracy results for the word embeddings contained in our models (based on the specification with $M = 200$). We rely upon public benchmarks commonly used to evaluate the capacity of the methodology to represent semantics. Word embeddings can solve analogies of the type “Ottawa is to Canada as Paris is to...” (France), by subtracting the difference between the two vectors of a known relationship from the query vector for the incomplete one (Mikolov, Chen, Corrado, and Dean 2013). Using a challenging test containing over 3,000 analogies to solve, we obtain satisfying accuracy scores compared to models trained on larger corpora; in particular, the Senate corpus achieves a 67.5% accuracy rate. In comparison, the state-of-the-art achieved by Pennington, Socher, and Manning (2014) with Global Vectors (GloVe) was 75%, using a corpus of 42 billion words. Note, however, that we accounted for the smaller size of our corpora by restricting the tests to analogies containing words among the 10,000 most frequent in our vocabularies, to ensure that the models had a minimal training with the the expressions involved. Overall, the results suggest that our models perform well at capturing semantics, despite the smaller sample size.

Table A7: Word Embedding Accuracy - Analogy Tests

	US House		US Senate		British Hansard		Canadian Hansard	
Category	Accuracy	Correct/Subtotal	Accuracy	Correct/Subtotal	Accuracy	Correct/Subtotal	Accuracy	Correct/Subtotal
Capitals: Common Countries	59.5%	(25/42)	54.8%	(23/42)	67.9%	(38/56)	85.0%	(17/20)
Capitals: World	76.5%	(13/17)	58.8%	(10/17)	70.4%	(19/27)	75.0%	(6/8)
Currency	0.0%	(0/6)	0.0%	(0/6)	8.3%	(1/12)	0.0%	(0/2)
City in State	74.3%	(277/373)	91.2%	(302/331)				
Family Relationships	64.3%	(36/56)	57.1%	(24/42)	71.4%	(30/42)	57.1%	(24/42)
Grammar 1: Adjective-to-adverb	33.6%	(170/506)	30.2%	(153/506)	33.4%	(169/506)	35.7%	(150/420)
Grammar 2: Opposite	55.8%	(87/156)	57.1%	(104/182)	50.7%	(138/272)	56.7%	(136/240)
Grammar 3: Comparative	79.3%	(476/600)	82.4%	(455/552)	86.6%	(608/702)	79.2%	(475/600)
Grammar 4: Superlative	79.5%	(105/132)	83.6%	(92/110)	79.5%	(105/132)	87.3%	(96/110)
Grammar 5: Present-participle	79.6%	(191/240)	80.5%	(219/272)	64.1%	(196/306)	75.0%	(180/240)
Grammar 6: Nationality-adjective	89.7%	(208/232)	87.5%	(230/263)	99.6%	(446/448)	85.9%	(177/206)
Grammar 7: Past-tense	58.9%	(445/756)	64.5%	(524/812)	53.8%	(407/756)	57.3%	(433/756)
Grammar 8: Plural	80.8%	(194/240)	75.3%	(137/182)	76.4%	(139/182)	74.2%	(178/240)
Grammar 9: Plural-verbs	47.8%	(87/182)	44.9%	(70/156)	46.2%	(61/132)	42.9%	(90/210)
Total	65.4%	(2314/3538)	67.5%	(2343/3473)	66.0%	(2357/3573)	63.4%	(1962/3094)

Analogy tests based on a benchmark list of word associations for word embeddings, using the models fitted with 200 dimensions. To account for the smaller sample sizes, the models are evaluated by restricting to the vocabulary of the 10,000 most frequently observed words. The test sheet is included in our release package.

The second accuracy test reported in Table A8 is another common benchmark based on a list of word similarities evaluated by humans (Finkelstein et al. 2002). The correlation coefficients measure to which extent the cosine similarities between the two words in our models are associated with human-based similarity scores for the same word pairs. For both models, we achieve positive and statistically significant correlation coefficients, using either Pearson's method or Spearman's rank-order correlation. The test comprises 353 word pairs.

Table A8: Word Embedding Accuracy - Word Similarity Tests

	US House		US Senate		British Hansard		Canadian Hansard	
	Correlation	p-value	Correlation	p-value	Correlation	p-value	Correlation	p-value
Pearson	0.5812	3.35e-31	0.5670	2.74e-29	0.5430	2.09e-25	0.5528	1.11e-26
Spearman	0.6052	2.37e-34	0.5965	5.46e-33	0.5709	1.84e-28	0.5731	5.58e-29

Analogy tests based on a list of 353 human-evaluated word similarities from Finkelstein et al. (2002).

Table A9: Extended Accuracy Results for Various Parameterizations, Part 1 (Vector Size, Window Size, Learning Rate)

Corpus	M	Δ	Learning Rate	Pearson Correlation	Pairwise Accuracy
House	100	15	0.025	0.837	85.329%
House	200	15	0.025	0.931	85.636%
House	300	15	0.025	0.938	85.329%
House	100	20	0.025	0.845	86.623%
House	200	20	0.025	0.918	85.658%
House	300	20	0.025	0.931	84.408%
House	100	30	0.025	0.834	85.921%
House	200	30	0.025	0.921	85.724%
House	300	30	0.025	0.902	84.671%
House	200	20	0.01	0.911	86.996%
House	200	20	0.02	0.925	86.096%
House	200	20	0.03	0.934	86.009%
House	200	20	0.04	0.928	83.750%
House	200	20	0.05	0.898	83.004%
Senate	100	15	0.025	0.859	85.219%
Senate	200	15	0.025	0.912	84.693%
Senate	300	15	0.025	0.909	84.298%
Senate	100	20	0.025	0.887	85.263%
Senate	200	20	0.025	0.919	83.925%
Senate	300	20	0.025	0.869	83.991%
Senate	100	30	0.025	0.885	84.583%
Senate	200	30	0.025	0.874	84.605%
Senate	300	30	0.025	0.844	84.057%
Senate	200	20	0.01	0.864	84.386%
Senate	200	20	0.02	0.910	84.496%
Senate	200	20	0.03	0.908	83.816%
Senate	200	20	0.04	0.920	83.509%
Senate	200	20	0.05	0.863	82.412%

The evaluated placements are obtained using the first principal component of party embeddings, for various parameterizations. Accuracy is assessed against the average party score on the first dimension of the Voteview DW-NOMINATE estimates (1921-2016).

Table A10: Extended Accuracy Results for Various Parameterizations, Part 2 (Vector Size, Window Size, Learning Rate)

Corpus	M	Δ	Learning Rate	Pearson Correlation	Pairwise Accuracy
Canada	100	15	0.025	0.799	78.608%
Canada	200	15	0.025	0.856	80.285%
Canada	300	15	0.025	0.856	79.778%
Canada	100	20	0.025	0.814	79.589%
Canada	200	20	0.025	0.855	79.778%
Canada	300	20	0.025	0.856	79.684%
Canada	100	30	0.025	0.830	79.905%
Canada	200	30	0.025	0.842	79.494%
Canada	300	30	0.025	0.850	79.937%
Canada	200	20	0.01	0.848	79.367%
Canada	200	20	0.02	0.853	79.589%
Canada	200	20	0.03	0.858	79.652%
Canada	200	20	0.04	0.853	79.715%
Canada	200	20	0.05	0.850	79.937%
UK	100	15	0.025	0.877	82.600%
UK	200	15	0.025	0.871	83.159%
UK	300	15	0.025	0.859	81.971%
UK	100	20	0.025	0.864	81.831%
UK	200	20	0.025	0.876	82.669%
UK	300	20	0.025	0.870	82.390%
UK	100	30	0.025	0.862	81.132%
UK	200	30	0.025	0.872	82.460%
UK	300	30	0.025	0.854	81.831%
UK	200	20	0.01	0.858	81.971%
UK	200	20	0.02	0.866	81.621%
UK	200	20	0.03	0.869	82.460%
UK	200	20	0.04	0.857	82.041%
UK	200	20	0.05	0.815	78.966%

The evaluated placements are obtained using the first principal component of party embeddings, for various parameterizations. Accuracy is assessed against the Legacy measure based on the CMP data (1945-2015).

Table A11: Extended Accuracy Results for Various Parameterizations, Part 3 (Epochs)

Corpus	Epochs	Pearson Correlation	Pairwise Accuracy
House	1	0.917	75.285%
House	3	0.949	85.263%
House	5	0.918	85.658%
House	10	0.907	86.447%
House	15	0.895	85.263%
Senate	1	0.902	72.873%
Senate	3	0.946	83.333%
Senate	5	0.919	83.925%
Senate	10	0.903	84.145%
Senate	15	0.857	84.276%
Canada	1	0.762	78.323%
Canada	3	0.844	79.778%
Canada	5	0.855	79.778%
Canada	10	0.864	80.696%
Canada	15	0.828	79.810%
UK	1	0.876	82.600%
UK	3	0.866	82.110%
UK	5	0.876	82.669%
UK	10	0.870	83.089%
UK	15	0.861	81.761%

The evaluated placements are obtained using the first principal component of party embeddings, for various epoch lengths. Accuracy is assessed against the Legacy measure based on the CMP data for Canada and the UK (1945-2015), and using the Voteview data for the USA (1921-2016). All models are computed with a 0.025 learning rate, 200-dimensional embeddings and a symmetrical context window of 20 words.

References

Abramowitz, Alan I., and Kyle L. Saunders. 2008. "Is Polarization a Myth?" *The Journal of Politics* 70 (2): 542-555.

Adams, James, Jane Green, and Caitlin Milazzo. 2012. "Who Moves? Elite and Mass-Level Depolarization in Britain, 1987–2001." *Electoral Studies* 31 (4): 643–655.

Beelen, Kaspar, Timothy Alberdingk Thijm, Christopher Cochrane, Kees Halvemaan, Graeme Hirst, Michael Kimmins, Sander Lijbrink, Maarten Marx, Nona Naderi, Roman Polyanovsky, Ludovic Rheault, and Tanya Whyte. 2017. "Digitization of the Canadian Parliamentary Debates." *Canadian Journal of Political Science* 50 (3): 849–864.

Clarke, Harold D., David Sanders, Marianne C. Stewart, and Paul F. Whiteley. 2009. *Performance Politics and the British Voter*. New York: Cambridge University Press.

Cochrane, Christopher. 2015. *Left and Right: The Small World of Political Ideas*. Montreal, Kingston: McGill-Queen's University Press.

Dalton, Russell J. 2008. "The Quantity and the Quality of Party Systems." *Comparative Political Studies* 41 (7): 899-920.

Finkelstein, Lev, Evgeniy Gabrilovich, Yossi Matias, Ehud Rivlin, Zach Solan, Gadi Wolfman, and Eytan Ruppin. 2002. "Placing Search in Context: The Concept Revisited." *ACM Transactions on Information Systems* 44 (1): 116–131.

Gentzkow, Matthew, Jesse M. Shapiro, and Matt Taddy. 2018. *Congressional Record for the 43rd-114th Congresses: Parsed Speeches and Phrase Counts*. Palo Alto, CA: Stanford Libraries [distributor].

Kusner, Matt, Yu Sun, Nicholas Kolkin, and Kilian Weinberger. 2015. "From Word Embeddings to Document Distances." In *International Conference on Machine Learning*. pp. 957–966.

Layman, Geoffrey C., Thomas M. Carsey, and Juliana Menasce Horowitz. 2006. "Party Polarization in American Politics: Characteristics, Causes, and Consequences." *Annual Review of Political Science* 9 (1): 83-110.

Lee, Frances E. 2015. "How Party Polarization Affects Governance." *Annual Review of Political Science* 18 (1): 261–282.

Levendusky, Matthew. 2009. *The Partisan Sort*. Chicago: University of Chicago Press.

McCarty, Nolan, Keith Poole, and Howard Rosenthal. 2006. *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge: MIT Press.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. "Distributed Representations of Words and Phrases and their Compositionality." In *Proceedings of the 26th International Conference on Neural Information Processing Systems*. pp. 3111–3119.

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. "Efficient Estimation of Word Representations in Vector Space." In *Proceedings of Workshop at ICLR*.

Monroe, Burt L., Michael P. Colaresi, and Kevin M. Quinn. 2008. "Fightin' Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict." *Political Analysis* 16 (4): 372–403.

Pennington, Jeffrey, Richard Socher, and Christopher D. Manning. 2014. "Glove: Global Vectors for Word Representation." In *Conference on Empirical Methods in Natural Language Processing (EMNLP 2014).*

Peterson, Andrew, and Arthur Spirling. 2018. "Classification Accuracy as a Substantive Quantity of Interest: Measuring Polarization in Westminster Systems." *Political Analysis* 26 (1): 120–128.

Rheault, Ludovic, and Christopher Cochrane. 2019. "Word Embeddings for the Analysis of Ideological Placement in Parliamentary Corpora." https://doi.org/10.7910/DVN/K0OYQF, Harvard Dataverse.