Constituency influence in the Canadian House of commons: A Language Based Approach

POL 2500 Government of Canada - University of Toronto

Justin Savoie

2019-04-10

Abstract

Are officeholders in single-member electoral districts responsive to the preferences of their constituents. Rather, does party discipline overshadows such a possibility? This research paper asks if textual scaling and natural language methods can uncover this latent spatial dimension in legislative speech? In line with previous research, we find that ideology in speech is difficult to observe. The main reason is that other dimensions like the government-opposition dimension are more salient. When the government-opposition dimension is modelled, the provincial and the ministerial dimensions are allowed to dominate. As a result, it is difficult to isolate and measure ideology from speech.

Are officeholders in single-member electoral districts responsive to the preferences of their constituents, or does party discipline overshadows such a possibility?[1] Traditionally, roll call voting have been used to measure and scale the ideological placement of legislators. Legislators voting together are assumed to have similar preferences. Their scores can then be compared to results from representative polls from the electoral districts. More simply, it can be determined whether legislators who were elected in strongholds behave differently than those who were elected in competitive riding. For instance, a MP from the most ideologically progressive party elected in a landslide in his own riding might exhibit more progressive tendencies than one in a very competitive riding.

It is difficult to study roll call voting in Canada because party discipline is very high. Politicians from a given party usually vote in block. An alternative empirical strategy is to examine patterns of speech in Parliament. It can be hypothesized that MPs from very progressive districts will adopt an accordingly fitting pattern in speech. Assuming that an objective measure of spatial (left-right) positioning exists for electoral districts, can we

[1] Replication material is available online: https://github.com/justinsavoie/canadian_constituencies

expect textual scaling and natural language processing methods to uncover this latent spatial dimension? This research paper uses a recent methodology, word embeddings, and two recently made available data sources to test this hypothesis. The paper is organized as follow. First, we discuss the tenets of democratic theory and dyadic representation. Then we discuss methodological approaches to the study of roll call voting and legislative text. After presenting the data and our method based on word embeddings, we present results from various models. Our main conclusion is that modelling ideology is difficult. We find that a government-opposition is dominant. When this dimension is considered, the provincial and ministerial dimensions become dominant.

Democratic theory and dyadic representation

Political actors respond to the preferences of citizens. Otherwise, talking about democracy does not make sense (Key 1961). Public opinion has an assumed importance in shaping policymaker's view (Shapiro 2011). This, of course, is not a new argument. It is as old a polling methods, and (Gallup 1938); possibly older. Even if it is not perfect, nor always easy to implement, democracy allows to protect and advance the interest of the body of citizen (Dahl 1989). In this sense, the work of Pitkin (1967) for whom the fundamental role of representatives is to act in the interest of the represented as become a defining feature of empirical work (Soroka, Penner, and Blidook 2009).

In electoral systems where elected official represent citizens from a geographical district, one fundamental question is the degree to which the policy positions of legislators reflect the policy preferences of their constituents (Hanretty, Lauderdale, and Vivyan 2017). On the one hand, internal cohesion in a political party body is absolutely crucial (Bowler, Farrell, and Katz 1999). Citizens cannot expect their representatives to defend a position, whatever the cost to the party at the national level. Even if members of parliament are ideologically heterogenous, high party discipline is expected, especially in Canada (Malloy 2003). On the other hand, in those electoral system, members of parliament are dependent on support in their constituency for reelection. Moreover, there is some evidence that there exists some patterns of association between voting behaviour and the average voter (Clinton 2006; Kastellec, Lax, and Phillips 2010; Krimmel, Lax, and Phillips 2016). The evidence, though, is mixed (some type of elected officials are more responsive; some persistent bias exists, i.e. elected officials are not responsive on LGB rights) and mostly the result of American scholarship. Weissberg (1978) defines this opposition as a collective vs dyadic representation. In this context, dyadic representation simply means the relation between an elected official and member of its consituency, while

collective representation is the relation between the elected body as a whole and the body of citizen. In the case at hand, it is possible to add the relation between the political party and its supporters. Reformulating Dahl's argument, democracy allows to protect and advance the interest if elected officials are responsive to the preferences of their consitutents, parties are responsive their partisans and the system as whole, to the collectivity. The focus of this paper is on dyadic representation.

In Canada, it is uncertain that representatives can act independently, for the sake of their constituents. Party leader's power is massive with respect to backbenchers (Docherty 1997; Savoie 1999). MPs are not viewed as representing their constituency. Institutional constraints are important, careers are short and it is difficult to estable a reputation (Soroka, Penner, and Blidook 2009). There is evidence, however, that MPs are interested in such representation, but it is difficult in practice (Clarke and Kornberg 1992). There is also some evidence that citizens be more sympathetic to responsive candidates (Blais et al. 2003). This can translate in vote for a local candidate, even if the citizen prefers another party (Blais and Daoust 2017). Furthermore, Canada's single member plurality system provides very strong incentive to allow for dyadic representation (Soroka, Penner, and Blidook 2009).

Votes as data

Examining dyadic representation through the study of roll call voting is common in the United States and elsewhere (Miller and Stokes 1963; Mayhew 1974; Poole and Rosenthal 1985, 2000, 2001). A logistic regression models binary votes as a function of legislator and item parameters simultaneously. Methodological extensions have allowed measure uncertainty with precision using the parametric bootstrap (Lewis and Poole 2004; Carroll et al. 2009) or a more comprehensive Bayesian framework (Clinton, Jackman, and Rivers 2004; Bafumi et al. 2005).

Studying dyadic representation through voting patterns, however, requires variation among elected officials of a given party. In the United Kingdom, where party discipline is high and such variation is relatively low, Hanretty, Lauderdale, and Vivyan (2017) have used Early Day Motion signatures, votes of conscience and unwhipped votes to estimate a similar model. They found evidence in favour of moderate responsiveness to constituency opinion[2]. In Canada, however, such free votes are scarce (Overby, Raymond, and Taydas 2011); virtually every vote is considered potentially one of non-confidence Kilgour and Kirsner (1988).

[2]Hanretty, Lauderdale, and Vivyan (2017) also offer an interesting literature review of the use of roll-call vote to study dyadic representation in a comparative context.

Methodology: text as data

When roll-call analysis is of limited use, legislative speech can be used to study sptial positions [Proksch and Slapin (2010); Proksch2015]. Apart from computer assisted dictionary methods (Laver and Garry 2000), two different approaches on word frequency have dominated the field of text scaling. Both Wordscores (Laver, Benoit, and Garry 2003) and Wordfish (Slapin and Proksch 2008) use a bag-of-words model. In these models, each text is represented as a set (bag) of its words. Grammar, syntax and everything related to word order generally is discarded. With Wordscores, the researcher provides extreme texts. The frequencies of all the words in the texts to be evaluated will be assessed a through a statistical model against the extreme reference texts. Using Wordfish, the words counts from all documents will be modelled using a Poisson-Item Response Theory. Such a model allows to place texts or legislators along a single spatial dimension. If the researcher wants to estimate multiple dimensions, Slapin and Proksch (2008) simply suggest to run the analysis multiple times on subsets of the data that discuss each of the different dimensions. In a complex political space, it therefore makes sense to estimate multiple wordfish models. The *Wordschoal* generalization of this method is presented by Lauderdale and Herzog (2016). In a first stage, the authors estimate a Wordfish model on the debates for each separate bills. This allows to satisfy Wordfish's most stringent condition, that all texts discuss one single dimension. In a second step, the authors use Bayesian factor analysis to reduce the all the individuals scores by each legislator on various bills to one single dimension. Bayesians methods allow to model missing data seamlessly which is important as not all legislators will discuss each bill. The author's main substantive finding is that some patterns can be observed, but are about the a government-opposition dimension (Ireland) or about major political events (U.S. Senate), rather than left-right ideology.

Recent breakthroughs in machine learning research suggest that word embeddings can replace bag-of-words approach. In a major paper Mikolov, Chen, et al. (2013) show how continuous vector representations of words can be estimated in large data sets. In a nutshell, using backpropagation, an optimization method used in artificial neural network, each word in a corpus is modelled as a function of the words just before and just after. Because of the number of hidden layers between the input and the predictors is low, this is an artificial shallow network; i.e. word embeddings are not deep learning (Goldberg and Levy 2014). An extension allows to also model document metadata (Mikolov, Sutskever, et al. 2013). These methods implemented for general use (Řehůřek and Sojka 2010) in the Python language are called word2Vec and doc2Vec.

Word2vec and the associated doc2vec allow to create vectors of n dimensional that will represent each word or each document. In word2vec, each word is associated with one vector. In doc2vec each word and each document is associated with of n dimensional vector. We choose n = 200 following (Rheault and Cochrane 2018). We have also tested with n = 50, 100, and 300. The results are stable. Hence, when using word2vec, the positio of a document can be studied by looking at the average of the value of the vectors associated with it's content. A document with 100 words will be described by a 100 x 200 matrix. Averaging over words will give a 1 x 200 vector. Standard data reduction techniques such as PCA can then be used to put many documents on a few dimensions. In the case of a corpus of 100,000 documents, this 100,000 x 200 matric can be reduced through PCA or t-sne (Barkan and Koenigstein 2016). This can be useful in political science where the political space is also conceived in a few number of dimensions.

Data

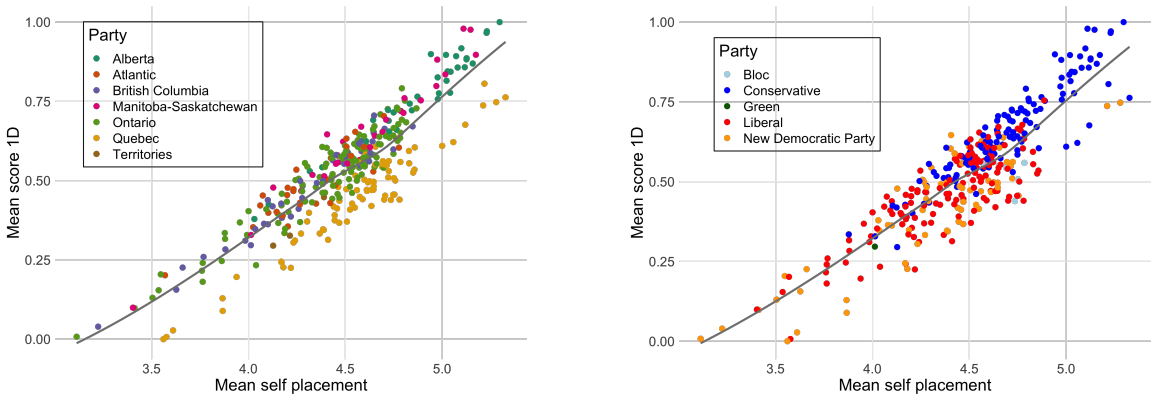
This paper makes use of two very recent and underexploited data sources. Two recent ideological scaling methods are used and compared. The English Debates from the Canadian House of Commons were recently made available online (Beelen et al. 2017). These Parliamentary Debates transcript provide the raw material to estimate ideological positions. Scores are calculated for all the elected MP sitting in the 42nd Canadian Parliament. The researcher analyzing text data has some choices to make in terms of how the corpus is defined. In the Canadian House of Commons, LIPAD separates debate in different topics. The main ones are the government orders, the oral question period, routine proceedings, statements by members, private member's business and ajournment proceedings. The oral question period data is more appropriate for Wordschoal as it is broken down in substantice subcategories; justice, taxation, foreign affairs, national defence, etc. Statements by members are most appropriate for an unsupervised ideological as MP are free "to address the House for up to one minute on virtually any matter of local, provincial, national or international concern" ("Compendium of Procedure," n.d.). Futher choice must but made in regards with the time period of study. One possibility would be to use only speech produced during the 42nd parliament because we are studying MP during this period. We can, however, train the model on more data to obtain more precise word embeddings. All in all, we ran multiple models on different time periods. We ran models both on Statements by members, Oral questions and both. When running a doc2vect model on data coming from both statements by members and the oral questions, we added a metadata variable indicating wheter the text was a statement or a

question[^3].

[^3] If this paper is to be presented or submitted, I will add a section discussing the exact number of politcians in the data set. I also noticed some unexpected behaviour in the speaker riding variable in LIPAD between the 2015 election on October 19th and January 1st 2016. Some ridings from the 2003 representation order were still present in the data after October 19th. This has to be examined qualitatively, but ran out of time before the oral presentation and the due date for this paper in POL 2500. Similarly, I will add a section discussing all the choices made concerning the time period, the type of topic, etc. In other words, word2vec or doc2vec is not one single model. The use of word2vec and doc2vec involves choices and these choices must be justify, something we did not really have time do to as we modified signficantly the methods (from word2vec only to doc2vec) between the oral presentation and the paper's due date.

We use data from Vote Compass Canada 2015 to estimate the objective ideological positions in each of Canada's 338 ridings.[3] The Vote Compass data set asked 30 policy question to over 1.8 million. This cleaned version of the data set has information about around 1.6 million respondents in the 338 canadian ridings. The median number of respondents per riding is 4241, from 558 in Nunavut to 15,501 in Laurier-Sainte-Marie. We use this data as an objective of the ridings' spatial ideology i.e. which ridings are left and which are right. Some have argued that the political value space in Canada is not unidimensional (Héroux-Legault 2016). There would be economic, social, moral, etc. dimensions. For simplicity we use one single dimension. 1 shows strong evidence that citizens think of the spatial left and right in a somewhat unidimensional manner. The figure plots the average self placement in each of the 338 ridings vis-à-vis a one dimensional factor analysis reduction using regression scores. The same results is presented twice, once by province and once by party (the party who ultimately won during the 2015 election). The objective measure based on dimensionality reduction from policy question and the subjective self placement measure are strongly correlated. This suggest that a unidimensional measure captures adeqautely some polarization in the Canadian context. Since this unidimensional axis seem to make sense objectively and subjectively, it therefore makes sense to use it as ideological positional indicator.

[3] Vote Compass is an interactive, online application developed by political scientists and run during election campaigns. http://votecompass.com/



(a) By province

(b) By party

Figure 1: Each point represents a single riding / member of parliament during 42nd Canadian Parliament. Mean score 1D is the score obtained from a unidimensional factor analysis of the spatial positions of voters in the riding. It is seen that conservative ridings are more to the right. Liberal and neo democrat ridings are more to the left. On the y axis, theta is the legislator point estimated by Wordschoal. There is a correlation along a government-opposition dimension, rather than an ideological dimension.

Results

Is there a pattern differentiating speech from a MP representing a riding on the left and a MP representing a riding on the right? Figure 2 shows the results of a wordschoal analysis. Each point represents a single riding / MP during 42nd Canadian Parliament. Mean score 1D is the score obtained from a unidimensional factor analysis of the spatial positions of voters in the riding. It is seen that conservative ridings are more to the right. Liberal and Neo Democrat ridings are more to the left. On the vertical y axis, *theta* measures the legislator ideological point estimated by wordschoal. There is a correlation along a government-opposition dimension, rather than an ideological dimension. On this dimension, conservative members of parliament are more vigorous on their attacks on the liberal dimension. Neo-democrats are situated midway on this government-opposition dimension. Furthermore, within party correlation between constituents' ideology and candidate's wordschoal position is not observed. Between party variation is more salient than within party variation as is made clear from the different color clusters along the y axis. Using wordschoal, the government-opposition dimension overshadows an ideological dimension.

Figure 3 presents the results of two word2vec models. With this model, only word usage is taken into account, and averaged for each MP over the period. The left panel shows a results from a model trained on data from the oral question period. The panel to the right shows results from a model trained on statements by members. The results are similar to one another and to the results from wordschoal model (the direction of the primary component

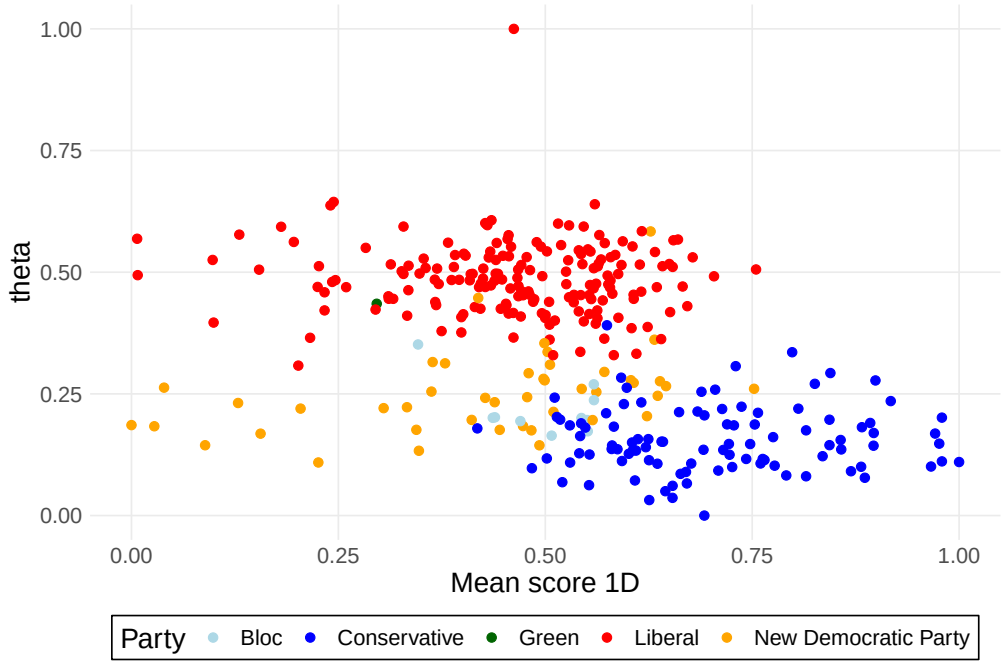
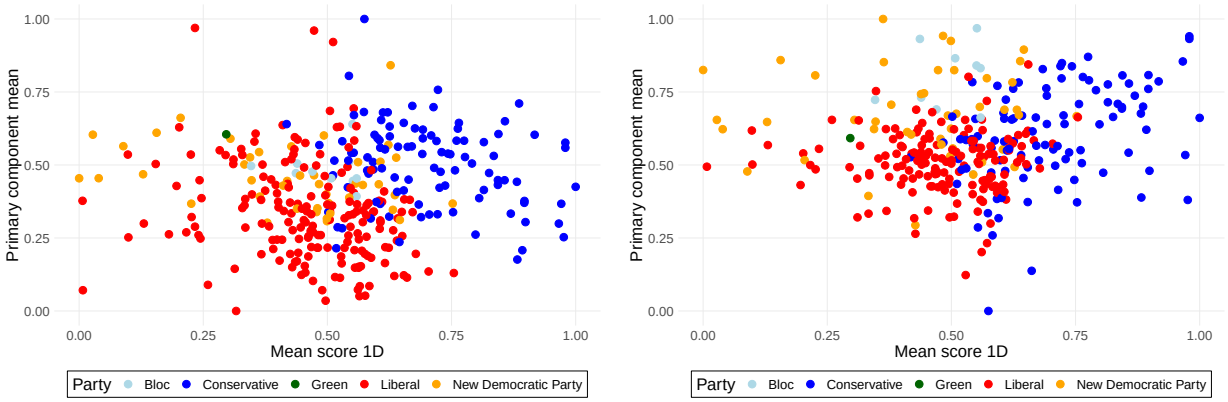


Figure 2: Each point represents a single riding / member of parliament during 42nd Canadian Parliament. Mean score 1D is the score obtained from a unidimensional factor analysis of the spatial positions of voters in the riding. It is seen that conservative ridings are more to the right. Liberal and neo democrat ridings are more to the left. On the y axis, theta is the legislator point estimated by Wordschoal. There is a correlation along a government-opposition dimension, rather than an ideological dimension.

is flipped, i.e. the government party is below the opposition). On the y axis, both the Conservatives and the Neo Democrats are above the Liberals. Again, the polarization is along a government-opposition dimension.

We turn our attention to results from doc2vec models. We recall that doc2vec allows to model metadata such a the identity of the legislator and party affiliation. Figure 4 presents the results of two doc2vec models. The panel to the left presents results incorportating legislator identity and party affiliation. Two separate clusters are observed. At the top, there are members of parliament from the province of Quebec. At the bottom, there are members of parliament from the rest of Canada. Hence, once we incorporate the government-opposition dimension in the model, the dimension that is allow to rise to prominence is the Quebec-ROC difference. We incorporate this second dimension in the meta data passed to the model in the right panel of figure 4 (the direction of the y axis is flipped). An interesting phenomenon can be noted. When party affiliation and geography are incorporated in the model, the highlighted dimension identified is one between ministers (at the bottom) and other (backbenchers) members of parliament. This model includes both the oral question period and statements by members as ministers are not allowed such statements. Figure **??**



(a) Oral question period

(b) Statements by members

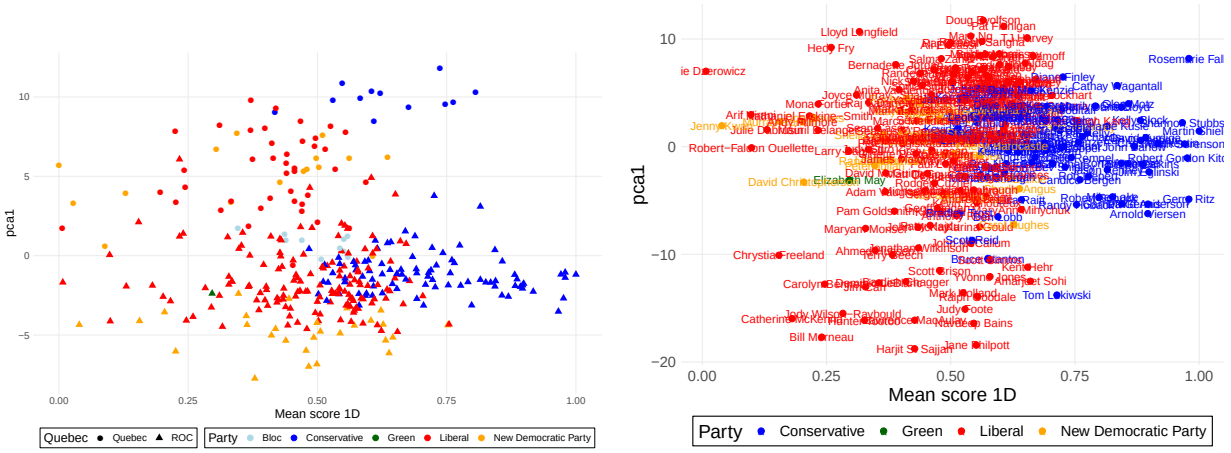
Figure 3: Each point represents a single riding / member of parliament during 42nd Canadian Parliament. Mean score 1D is the score obtained from a unidimensional factor analysis of the spatial positions of voters in the riding. The primary component mean is obtained by average the scores of each word in a text on its 200 dimensions calculated with word2vec. Then PCA is used to reduce again all the MP to one single dimension.

shows the results of such an analysis when looking at statements by members only. No clear pattern is identifiable.

Discussion

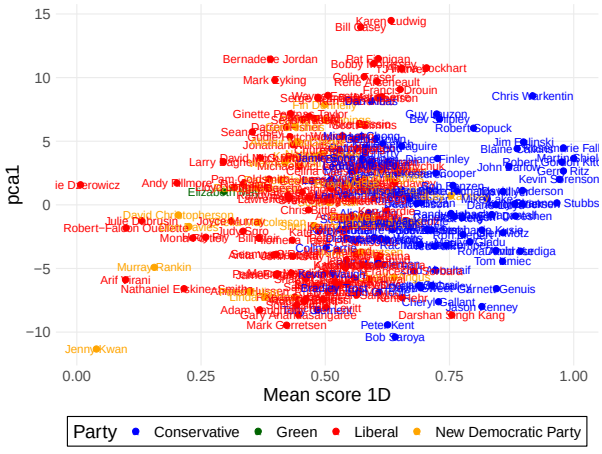
In a similar (using a support-vector machine classifier instead of artificial neural networks) study of ideology measured by speech in the Canadian House of commons, Hirst et al. (2014) note that "language of attack and defence, of government and opposition, dominates and confounds any sensitivity to ideology in these kinds of classifiers." As also shown by Hanretty, Lauderdale, and Vivyan (2017), the government-opposition dimension dominates. Furthermore, this study has shown that when the government-opposition dynamic is modelled, other dynamics are brought to the forefront. Specifically, we found that when the government-opposition dimension is modelled, the provincial and the ministerial dimensions are allowed to dominate; not ideology.

Text analysis using word embeddings in political science is still in its infancy. Much can still be achieved. Two points can be discussed. First, we believe that the unsupervised methods used in this paper are prone to allow confounding. We recall, doc2vec models documents and the output is a high dimensional vector. A supervised approach with a high quality data set could enable us to take better advantage of the richness of such high dimensional vectors. By definition, high dimensional vectors are complex and will lead to the identification of patterns



(a) Modelling party as part of doc2vec (b) Modelling party and province as part of doc2vec

Figure 4: Each point represents a single riding / member of parliament during 42nd Canadian Parliament. Mean score 1D is the score obtained from a unidimensional factor analysis of the spatial positions of voters in the riding. The primary component mean is obtained by average the scores of each word in a text on its 200 dimensions calculated with word2vec. Then PCA is used to reduce again all the MP to one single dimension. Data used is from statements by members and the oral question period.



(a) Modelling party

Figure 5: Each point represents a single riding / member of parliament during 42nd Canadian Parliament. Mean score 1D is the score obtained from a unidimensional factor analysis of the spatial positions of voters in the riding. The primary component mean is obtained by average the scores of each word in a text on its 200 dimensions calculated with word2vec. Then PCA is used to reduce again all the MP to one single dimension. Data used is only from statements by members.

that might not have been expected when reduce to a single dimension using PCA. A high quality *labeled data set* used as explanandum could likely make the most of such vectors of explanans.

Second, the dataset used is relatively small (2015-2019). Although we have considered alternative data sets ranging from 1993 to 2019, these data sets are relatively small compared to the original data sets used to develop and test word2vec and doc2vec. One possibility is to expand the time frame to train a model on data from far in the past (LIPAD goes back to 1901). Additional tests could tell us if this is a good strategy; if data usage is constant over time. Another possibility would be to train the model on other data sources. It is not impossible, for instance, that a large english corpus of political news from google could help us understand patterns of political language more adequately. This trained model could in turn allow us to study ideology in parliaments accross the world.

At the substantive level two further elements should be taken into account. First, it would be interesting also to look at electoral scores to identify strongholds and competitive riding. It can be assumed that secure MP could use speech differently in parliament. Second, high dimensional vectors do not only allow for unidimensional reduction. It can be assumed that in their statements MP discuss issues

Appendix 1 Vote Compass policy questions

- How much should Canada do to reduce its greenhouse gas emissions?
- No new oil pipelines should be built in Canada.
- How much power should unions have?
- The most effective way to create jobs in Canada is to lower taxes.
- How many new immigrants should Canada admit?
- How much should be done to accommodate religious minorities in Canada?
- Terminally ill patients should be able to end their own lives with medical assistance.
- Abortions should be allowed in all cases, regardless of the reason.
- Possession of marijuana should be a criminal offence.
- Longer prison sentences are the best way to prevent crime.

- Handguns should be banned in Canada.
- To what extent should law enforcement be able to monitor the online activity of Canadians?
- Government workers should not be allowed to strike.
- How much tax should corporations pay?
- How much should wealthier people pay in taxes?
- How supportive should Canada be of Israel?
- How much should Canada spend on foreign aid?
- How involved should the Canadian military be in the fight against ISIS?
- How much of a role should the private sector have in health care?
- Illicit drug users should have access to safe injection sites.
- How much should the government do to make amends for past treatment of Aboriginal Peoples in Canada?
- Aboriginal Peoples in Canada should have more control over their ancestral territory.
- Canada should introduce a publicly funded childcare program.
- Canada's budget should be balanced no matter what.
- Quebec should be formally recognized as a nation in the Constitution.
- Quebec should become an independent state.
- Canada should end its ties to the monarchy.
- The Canadian government should put a price on carbon.
- The Senate should be abolished.
- Only those who speak both English and French should be appointed to the Supreme Court.
- In politics people sometimes talk of left and right. Where would you place yourself on the scale below, where 0 is left and 10 is right?

References

Bafumi, Joseph, Andrew Gelman, David K Park, and Noah Kaplan. 2005. "Practical Issues in Implementing and Understanding Bayesian Ideal Point Estimation." *Political Analysis* 13 (2). Cambridge University Press:171–87.

Barkan, Oren, and Noam Koenigstein. 2016. "Item2vec: Neural Item Embedding for Collaborative Filtering." In *2016 Ieee 26th International Workshop on Machine Learning for Signal Processing (Mlsp)*, 1–6. IEEE.

Beelen, Kaspar, Timothy Alberdingk Thijm, Christopher Cochrane, Kees Halvemaan, Graeme Hirst, Michael Kimmins, Sander Lijbrink, et al. 2017. "Digitization of the Canadian Parliamentary Debates." *Canadian Journal of Political Science/Revue Canadienne de Science Politique* 50 (3). Cambridge University Press:849–64.

Blais, André, and Jean-François Daoust. 2017. "What Do Voters Do When They Like a Local Candidate from Another Party?" *Canadian Journal of Political Science/Revue Canadienne de Science Politique* 50 (4). Cambridge University Press:1103–9.

Blais, André, Elisabeth Gidengil, Agnieszka Dobrzynska, Neil Nevitte, and Richard Adeau. 2003. "Does the Local Candidate Matter? Candidate Effects in the Canadian Election of 2000." *Canadian Journal of Political Science/Revue Canadienne de Science Politique* 36 (3). Cambridge University Press:657–64.

Bowler, Shaun, David M Farrell, and Richard S Katz. 1999. "Party Cohesion, Party Discipline, and Parliaments." In *Party Discipline and Parliamentary Government*, edited by Shaun Bowler, David M Farrell, and Richard S Katz, 3–22. Ohio State University Press Columbus, OH.

Carroll, Royce, Jeffrey B Lewis, James Lo, Keith T Poole, and Howard Rosenthal. 2009. "Measuring Bias and Uncertainty in Dw-Nominate Ideal Point Estimates via the Parametric Bootstrap." *Political Analysis* 17 (3). Cambridge University Press:261–75.

Clarke, Harold D, and Allan Kornberg. 1992. "Do National Elections Affect Perceptions of Mp Responsiveness? A Note on the Canadian Case." *Legislative Studies Quarterly*. JSTOR, 183–204.

Clinton, Joshua. 2006. "Representation in Congress: Constituents and Roll Calls in the 106th House." *The Journal of Politics* 68 (2). Cambridge University Press New York, USA:397–409.

Clinton, Joshua, Simon Jackman, and Douglas Rivers. 2004. "The Statistical Analysis of Roll

Call Data." *American Political Science Review* 98 (2). Cambridge University Press:355–70.

"Compendium of Procedure." n.d. https://www.ourcommons.ca/about/Compendium/TypicalSittingDay/c_d_statementsmembers-e.htm.

Dahl, Robert Alan. 1989. *Democracy and Its Critics*. Yale University Press.

Docherty, David C. 1997. *Mr. Smith Goes to Ottawa: Life in the House of Commons*. UBC Press.

Gallup, George. 1938. "Testing Public Opinion." *The Public Opinion Quarterly* 2 (1). JSTOR:8–14.

Goldberg, Yoav, and Omer Levy. 2014. "Word2vec Explained: Deriving Mikolov et Al.'s Negative-Sampling Word-Embedding Method." *arXiv Preprint arXiv:1402.3722*.

Hanretty, Chris, Benjamin E Lauderdale, and Nick Vivyan. 2017. "Dyadic Representation in a Westminster System." *Legislative Studies Quarterly* 42 (2). Wiley Online Library:235–67.

Héroux-Legault, Maxime. 2016. "Substate Variations in Political Values in Canada." *Regional & Federal Studies* 26 (2). Taylor & Francis:171–97.

Hirst, Graeme, Yaroslav Riabinin, Jory Graham, Magali Boizot-Roche, and Colin Morris. 2014. "Text to Ideology or Text to Party Status?" *Kaal, Bertie; Maks, E. Isa; and van Elfrinkhof, Annemarie ME (Editors), from Text to Political Positions: Text Analysis Across Disciplines, Amsterdam: John Benjamins*, 93–115.

Kastellec, Jonathan P, Jeffrey R Lax, and Justin H Phillips. 2010. "Public Opinion and Senate Confirmation of Supreme Court Nominees." *The Journal of Politics* 72 (3). Cambridge University Press New York, USA:767–84.

Key, V. O. Jr. 1961. *Public Opinion and American Democracy*. Knopf.

Kilgour, David, and John Kirsner. 1988. "Party Discipline and Canadian Democracy." *Canadian Parliamentary Review* 11:10–11.

Krimmel, Katherine, Jeffrey R Lax, and Justin H Phillips. 2016. "Gay Rights in Congress: Public Opinion and (Mis) Representation." *Public Opinion Quarterly* 80 (4). Oxford University Press US:888–913.

Lauderdale, Benjamin E, and Alexander Herzog. 2016. "Measuring Political Positions from Legislative Speech." *Political Analysis* 24 (3). Cambridge University Press:374–94.

Laver, Michael, and John Garry. 2000. "Estimating Policy Positions from Political Texts."

American Journal of Political Science. JSTOR, 619–34.

Laver, Michael, Kenneth Benoit, and John Garry. 2003. "Extracting Policy Positions from Political Texts Using Words as Data." *American Political Science Review* 97 (2). Cambridge University Press:311–31.

Lewis, Jeffrey B, and Keith T Poole. 2004. "Measuring Bias and Uncertainty in Ideal Point Estimates via the Parametric Bootstrap." *Political Analysis* 12 (2). Cambridge University Press:105–27.

Malloy, Jonathan. 2003. "High Discipline, Low Cohesion? The Uncertain Patterns of Canadian Parliamentary Party Groups." *The Journal of Legislative Studies* 9 (4). Taylor & Francis:116–29.

Mayhew, David R. 1974. *Congress: The Electoral Connection*. Vol. 26. Yale University Press.

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. "Efficient Estimation of Word Representations in Vector Space." *arXiv Preprint arXiv:1301.3781*.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. "Distributed Representations of Words and Phrases and Their Compositionality." In *Advances in Neural Information Processing Systems*, 3111–9.

Miller, Warren E, and Donald E Stokes. 1963. "Constituency Influence in Congress." *American Political Science Review* 57 (1). Cambridge University Press:45–56.

Overby, L Marvin, Christopher Raymond, and Zeynep Taydas. 2011. "Free Votes, Mps, and Constituents: The Case of Same-Sex Marriage in Canada." *American Review of Canadian Studies* 41 (4). Taylor & Francis:465–78.

Pitkin, Hanna F. 1967. *The Concept of Representation*. Vol. 75. Univ of California Press.

Poole, Keith T, and Howard Rosenthal. 1985. "A Spatial Model for Legislative Roll Call Analysis." *American Journal of Political Science*. JSTOR, 357–84.

———. 2000. *Congress: A Political-Economic History of Roll Call Voting*. Oxford University Press on Demand.

———. 2001. "D-Nominate After 10 Years: A Comparative Update to Congress: A Political-Economic History of Roll-Call Voting." *Legislative Studies Quarterly*. JSTOR, 5–29.

Proksch, Sven-Oliver, and Jonathan B Slapin. 2010. "Position Taking in European Parliament Speeches." *British Journal of Political Science* 40 (3). Cambridge University Press:587–611.

Rheault, Ludovic, and Christopher Cochrane. 2018. "Word Embeddings for the Estimation of Ideological Placement in Parliamentary Corpora." PolMeth.

Řehůřek, Radim, and Petr Sojka. 2010. "Software Framework for Topic Modelling with Large Corpora." In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, 45–50. Valletta, Malta: ELRA.

Savoie, Donald J. 1999. *Governing from the Centre: The Concentration of Power in Canadian Politics*. University of Toronto Press.

Shapiro, Robert Y. 2011. "Public Opinion and American Democracy." *Public Opinion Quarterly* 75 (5). Oxford University Press:982–1017.

Slapin, Jonathan B, and Sven-Oliver Proksch. 2008. "A Scaling Model for Estimating Time-Series Party Positions from Texts." *American Journal of Political Science* 52 (3). Wiley Online Library:705–22.

Soroka, Stuart, Erin Penner, and Kelly Blidook. 2009. "Constituency Influence in Parliament." *Canadian Journal of Political Science/Revue Canadienne de Science Politique* 42 (3). Cambridge University Press:563–91.

Weissberg, Robert. 1978. "Collective Vs. Dyadic Representation in Congress." *American Political Science Review* 72 (2). Cambridge University Press:535–47.