The Effect of Elections and Changed Governments on Discussion in the Australian Federal Parliament (1901–2017) *

Monica Alexander   *University of Toronto*
Rohan Alexander   *Australian National University*

We systematically analyse how discussion in the Australian Federal Parliament changes in response to elections and changed governments. We first create a dataset of what was said in the Australian Federal Parliament from 1901 through to 2017 based on available public records. To reduce the dimensionality of this dataset we use a correlated topic model, and then analyse the effect of various events using a Bayesian hierarchical Dirichlet model. We find that: changes in government tend to be associated with topic changes even when the party in power does not change; and elections that do not result in a change in government are rarely associated with topic changes. The effect of events is much more pronounced in the second half of our sample, and especially in the past two decades. This has implications for how we think about the longer-term trajectory of government policymaking as the media and political cycles becomes increasingly focused on short-term events.

Keywords: text-as-data, Australian political history, unsupervised machine learning, Bayesian hierarchical Dirichlet model

1   Introduction

Government policy is partly driven by parliamentary discussion. Conversely, that same discussion can indicate a government's priorities. But major events – be they somewhat expected, such as an election, and unexpected, such as a recession or terrorist attack – can affect the course of parliamentary discussion. For instance, think of how a new government often goes to some trouble to seem different to the one they replace, or how events such as the 9/11 attacks altered government priorities.

In this paper we examine text records of what was said in the Australian Federal Parliament. We create a dataset based on the sitting days from 1901 through to 2017. We have records for 7,869 days in the House of Representatives and 6,682 days in the Senate. Our dataset consists of counts of every word that occurs at least thrice, after we convert all words to lower-case; remove names, numbers, and punctuation; join commonly co-located words; and remove stop words.

We systematically analyse the text in two stages, each of which takes advantage of a different statistical technique. First, to reduce the dimensionality of our dataset we use a topic model that allows correlation between topics. Second, we analyse those topics using a Bayesian hierarchical Dirichlet model to examine changes at various types of events. We explicitly consider two types of events: changes in government and elections. After taking into consideration these effects, we can then explore the effect of other significant events such as an economic recession or terrorism.

We find that: firstly, changes in government tend be associated with topic changes even when the party in power does not change. For instance, the change from Hughes to Bruce in 1922, Menzies to Holt in 1966; Hawke to Keating in 1991; and Rudd to Gillard in 2010 are all associated with significant changes in topics despite no change in the party in power. Secondly, as expected, elections where the party in power also changes, such as Fisher in 1910 and 1914, Menzies in 1949, Whitlam in 1972, Fraser in 1975, Hawke in 1983, Howard in 1996, Rudd in 2007, and Abbott in 2013 are associated with topic changes, but the 1974 Whitlam, 1980 Fraser, and the 1998 and 2004 Howard re-elections, stand out as elections where the government was returned yet there was a significant change in topics. Finally, economic events, such as financial crises, have less significant effects than other events such as terrorism.

As our dataset covers 117 years we are able to see how the effect of events changes over time. We find that the events that we are interested in are rarely significant in the first half of our dataset. With a small number of exceptions, even changes of government

2

where the party in power also changed were not associated with overly large changes in parliamentary discussion. If governments try to more thoroughly distinguish themselves from their predecessor, then longer-term policy may be more difficult to enact. Similarly, there is the danger that larger-scale signature-accomplishments of a government may be neglected by their successor for solely political reasons.

Our work contributes to the growing literature in quantitative social science that considers text as an input to statistical methods, rather than requiring separate analysis. This literature sits across and draws from various historically-separate disciplines including applied statistics, economics, and political science. We contribute to this literature in terms of both data and methods. From a data perspective we bring to bear an essentially-complete record of what was said in the Australian Federal Parliament on a daily basis, and we make our dataset available to other researchers via the R package `hansardr`. From a methods perspective, our analysis model has several advantages over existing methods. These include: allowing the events to have more-complicated auto-correlated functional forms; implementing pooling across groups of similar documents; and identifying outlying topic distributions without the need to pre-specify the event of interest. As the digitisation of historical sources continues and computational power becomes cheaper, we expect the popularlity of quantitative approaches to text to increase further and the data and methods in our paper to be extended in a variety of ways.

2 Data

Following the example of the UK a daily text record called Hansard of what was said in the Australian Federal Parliament has been made available since Federation in 1901.[1] Analysing Hansard records and their equivalents directly is increasingly viable as new

---

[1]While Hansard is not necessarily verbatim, it is considered close enough for text-as-data purposes. For instance, Mollin (2008) found that in the case of the UK Hansard the differences would only affect specialised linguistic analysis. Edwards (2016) examined Australia, New Zealand and the UK, and found that changes were usually made by those responsible for creating the Hansard record, instead of the parliamentarians. As those who create Hansard are tasked with creating an accurate record of proceedings, this suggests the records should be fit for the purpose of our analysis.

methods and reduced computational costs make it easier.

The recent digitisation of the Canadian Hansard, Beelen et al. (2017), has allowed increased analysis of Canadian parliament text records. For instance, Rheault and Cochrane (2018) examined ideology and party polarisation in Britain and Canada. A digitised version of the UK Hansard has been available for some time. For instance, Duthie, Budzynska and Reed (2016) examined which politicians made supportive or aggressive statements toward other politicians, and Peterson and Spirling (2018) examined polarisation. One exciting aspect of research using the UK Hansard has been linking text records with other datasets. For instance, Slapin et al. (2018) linked votes and speeches to examine grandstanding within parties. Another exciting aspect is that as digitisation methods improve, increasingly older UK records can be analysed, for instance, Dimitruk (2018) considered the effect of estate bills on prorogations in seventeenth century England. In New Zealand, Curran et al. (2018) modelled the topics discussed between 2003 and 2016, and Graham (2016) examined unparliamentary language between 1890 and 1950. And in the US, Gentzkow, Shapiro and Taddy (2018) examined congressional speech records from 1873 to 2016 to find that partisanship has risen in the past few decades.

Parts of Australian Hansard records have been analysed for various purposes. For instance, Rasiah (2010) examined Hansard records for the Australian House of Representatives to examine whether politicians attempted to evade questions about Iraq during February and March 2003. Gans and Leigh (2012) examined Australian Hansard records to associate mentions by politicians of certain public intellectuals with neutral or positive sentiment. Salisbury (2011) examined unparliamentary behaviour. And Fraussen, Graham and Halpin (forthcoming) examined Australian Hansard records to assess the prominence of interest groups. The closest research to ours that we have found is Boulus (2013) who examined parliamentary debate in Australia between 1946 and 2012.

The Australian Federal Parliament makes daily Hansard records available online as PDFs and these are considered the official release. Although we do not use them here,

4

XML records are also available in most cases.[2] We provide an example of a Hansard PDF page in Appendix A.1. There are 14,551 days of publicly available Hansard records across the two chambers of the Australian Federal Parliament for which we have PDFs and further summary statistics for this are provided in Appendix A.2. Our data cleaning process indicates concerns with a small number of PDFs and these are detailed in Appendix A.3.

We use the official PDF release and the formatting of the Hansard records changes over time. We use scripts written in R, R Core Team (2018), to convert the PDFs into daily text records. An example of the workflow and some example scripts are provided in Appendix A.4. Some error is introduced at this stage because many of the records are in a two-column format that need to be separated, and the PDF parsing is not always accurate especially for older records. An example of the latter issue is that 'the' is often parsed as 'thc'. These errors are corrected when they can be identified.

The percentage of stop-words in each record is reasonably consistent over time. This suggests that there is no significant difference in the quality of the parsing over time. Details of this check are provided in Appendix A.5. We use Hansard records on a daily basis in this paper. We pre-process our text before applying a topic model. The specific steps that we take are to: remove numbers and punctuation; change the words to lower case; and concatenate multi-word names titles and phrases, such as new zealand to new_-zealand. Then the sentences are de-constructed and each word considered individually. The resulting dataset used for analysis contains counts of words by day for each of the 14,551 days between 1901 and 2017.

3 Model

The goal of our modelling strategy is twofold. Firstly, we want to use topic modelling (Blei, Ng and Jordan, 2003) to summarise the Hansard text into meaningful topics that

---

[2]Tim Sherratt makes Commonwealth XML records for 1901 to 1980 available as a single download at: http://historichansard.net/. Commonwealth XML records from 1998 to 2014 are available from Andrew Turpin's website, and from 2006 through to today from Open Australia's website. The records can also be downloaded from the Australian Hansard website or the website can be scraped.

reduce the dimensionality of the text data and capture the main themes discussed in parliament over time. Secondly, we want to relate the resulting topic distributions to temporal trends, changes, and events, such as a change in government, elections, and other external events. The output of the first stage is the input to the second stage.

We first use a Correlated Topic Model (CTM) (Blei and Lafferty, 2007) to obtain estimated topic distributions over time. We consider these topic distributions as inputs that can be analysed by another model. Thus, the second modelling step involves using a Bayesian hierarchical Dirichlet model to analyse changes in the topic distributions (obtained from the first step) in relation to events of interest.

In the following section, we briefly describe the topic modelling approach, before discussing the Bayesian hierarchical Dirichlet analysis model used to investigate changes in topics. Background detail on topic modelling is available in Appendix B.

3.1 Overview of topic modelling

Although more- or less-fine levels of analysis are possible, here we are primarily interested in considering a day's topics. This means that each day's Hansard record needs to be classified by its topics. Sometimes Hansard records includes titles that make the topic clear. But not every statement has a title and the titles do not always define topics in a well-defined and consistent way, especially over longer time periods.

Other work such as Baumgartner and Jones (1993) and Dowding et al. (2010) addressed this problem by creating a standardised codebook of policy categories and subcategories and then manually assigning text to topics as appropriate. This approach ensures the categorisation is reasonable but as it is a manual process the size of the text that can be categorised is limited.

In order to effectively categorise the topics of the entire Hansard, we chose to use topic modelling, a statistical technique which aims to extract the underlying or latent 'topics' from a collection of texts. In particular, we use a method which is similar to

Latent Dirichlet Allocation (LDA), first developed by Blei, Ng and Jordan (2003).

The key assumption behind LDA is that each day's text, 'a document', in Hansard is made by speakers who decide the topics they would like to talk about in that document, and then choose words, 'terms', that are appropriate to those topics. A topic could be thought of as a collection of terms, and a document as a collection of topics, where these collections are defined by probability distributions. The topics are not specified *ex ante*; they are an outcome of the method, and it is in this sense that this approach can be considered unsupervised machine learning. Terms are not necessarily unique to a particular topic, and a document could be about more than one topic. The goal is to have the words found in each day's Hansard group themselves to define topics. This can provide more flexibility than other approaches such as a strict word count method, but can require a larger dataset and make interpretation more difficult. A

An overview and simplified example of how topic modelling works is available in Appendix B.1. The underlying document generation process is discussed in Appendix B.2, and Appendix B.3 explains how this is reversed to generate topics. Finally, the choice of the number of topics has a large effect on the composition of the topics, and some diagnostics for selecting the appropriate number are discussed in Appendix B.4.

One notable limitation of LDA is that the model assumes that the presence of one topic is not correlated with the presence of another topic. However, in reality topics are often related. For instance, in the Hansard context, we may expect topics related to the army to be more commonly found with topics related to the navy, but less commonly with topics related to banking.

As such, we chose to use a Correlated Topic Model (CTM) (Blei and Lafferty, 2007) to obtain topic distributions. CTM is a slight extension to LDA which allows for correlations between topics. More details about CTM can be found in Appendix B.5.

3.2   Analysis model

The CTM output of interest is the proportion of each topic appearing in each document. The aim of this stage of the modelling process is to analyse how the distribution of topics changes in relation to different types of events. But with many topics for each of the roughly 14,551 chamber-sitting-days the data are still too noisy to easily visualise changes around events.

One option for relating the topic distributions to events would be to use the Structural Topic Model (STM) of Roberts, Stewart and Airoldi (2016). The distinguishing aspect of the STM is that it considers more than just a document's content when constructing topics. For example, we may believe a document's author, or the time at which it was written, or, in the case of Hansard, the government or election period, may affect the topics within that document. The STM allows this additional information, or metadata, to affect the construction of topics, though influencing either topical prevalence or topical content. The assumption that there is some document generation process is the same as in LDA, it is just that this process now includes metadata.

However, the STM covariate framework has several limitations in terms of our goal to assess the relationship between topics and events:

1. There is no way of specifying more complicated auto-correlated functional forms of the effects of events over time. For example, we believe that the effect of an election would peak at the time of the election, then gradually decay as a function of days since election. In the STM framework, it is possible to specify a constant or linear effect of elections over time, or a spline relationship over elections, but it is not possible to restrict the effect of a specific election over time to be monotonically decreasing.

2. There is no way to implement partial pooling across groups of similar documents. The STM framework assumes that documents are independently and identically

distributed, conditional on the model covariates. However, it could be expected that topic distributions within a particular government, for example, may be more- or less-likely to contain certain topics for reasons that are not reflected in the topic prevalence covariates. To account for this, we would like a covariate model that allows for the partial pooling of variance in topic distributions by group, such as sitting period.

3. There is no way of identifying 'outlying' topic distributions – and therefore events that had an important effect – without pre-specifying the event of interest in the model. For example, if we think that the 9/11 attacks had an effect on parliamentary discourse, then a dummy for 9/11 would have to be included in the STM framework, but the specifics of the dummy construction affect the results. Instead we would like to identify important events based on different-to-expected topic distributions, after accounting for time trends, government and election effects.

To overcome these challenges, we formalise a statistical framework that allows us to systematically identify significant changes in topic distributions over time. Specifically, we use the estimated topic distributions from the CTM described in the previous section as an input into a Bayesian hierarchical Dirichlet regression framework, which relates the proportions of each topic to underlying time trends, changes in governments and elections. This set-up also allows us to identify 'outlying' topic distributions and relate these to other events.

Define θ_{dp} to be the proportion of topic p on day d. Note that the $\theta_{d,1:P}$ for $p = 1, 2, \ldots P = 80$ are equal to the estimated values of θ_d from the CTM. We assume that the majority of variation in topics is across sitting periods, s, where a sitting period is defined as any group of days that are less than one week apart. Using this definition, there are a total of 822 sitting periods over the period 1901 to 2017 inclusive.

The topic proportions on day d are modelled in reference to their membership of a particular sitting period s. Firstly, we assume that each distribution of topics, $\theta_{d,1:P}$ for

9

each day is a draw from a Dirichlet distribution with mean parameter $\mu_{s[d],1:P}$:

$$\theta_{d,1:P} \sim \text{Dirichlet}(\mu_{s[d],1:P})$$

where the notation $s[d]$ refers to the sitting period s to which day d belongs. This distributional assumption accounts for the fact that on any given day, the sum of all proportions in each topic must be one.

The goal of the model is to relate these proportions to government g at time d, and also the days since the most recent election, e, while account for underlying time trends. The mean parameters $\mu_{s,p}$ are modelled on the log scale as:

$$\log \mu_{s,p} = \alpha_{g[s],p} + \cdot \alpha_{e[s],d,p} + \sum_{k=1}^{K} \beta_{p,k} \cdot x_{s,k} + \delta_{s,p}$$

where: $\alpha_{g[s],p}$ is the mean effect for government g (which covers sitting period s) and topic p; $\alpha_{e[s],d,p}$ is the effect of election e (which occurs in sitting period s) for topic p on day d since the election; $\sum_{k=1}^{K} \beta_{p,k} \cdot x_{s,k}$ is the underlying time trend, modelled using splines: $x_{s,k}$ is the kth basis spline in sitting period s and $\beta_{p,k}$ is a coefficient on the kth basis spline; and $\delta_{s,p}$ is a structured random, or levels, effect for each sitting period and topic.

The government term, $\alpha_{g[s],p}$, assumes there is some underlying mean effect of each government on the topic distribution. We place uninformative priors on each of these parameters:

$$\alpha_{g[s],p} \sim \text{Normal}(0, 100).$$

The election term, $\alpha_{e[s],d,p}$, assumes there is an initial effect of an election on the topic distribution, which then decays as a function of days since election, d. In particular, we model this as an AR(1) in d:

$$\alpha_{e[s],d,p} = \rho_{e[s],p} \cdot \alpha_{e[s],d-1,p}.$$

One advantage of our model over using the STM is that we can restrict the effect of an election to be monotonically decreasing. This allows us to identify differences between government and election effects even when there is a one-term government. The value of the initial effect, $\alpha_{e[s],0,p}$, and the AR(1) term, ρ, both have non-informative priors:

$$\alpha_{e[s],0,p} \sim \text{Normal}(0, 100) \text{ and}$$

$$\rho_{e[s],p} \sim \text{Uniform}(0, 1).$$

We model the underlying time trend in topics using splines regression. The intuition behind this term is to capture the underlying non-linear trend in topic distributions over time, which is caused by large-scale structural changes to Australian society, culture, and the economy. The $x_{s,k}$ for $k = 1, 2, \ldots, K$ are the value of cubic basis splines for sitting period s at knot point k. We place knot points every five sitting periods as this is the average length of time for a government to sit. Non-informative priors are placed on the splines coefficients:

$$\beta_{p,k} \sim \text{Normal}(0, 100).$$

Finally, the sitting-period-specific random effect $\delta_{s,p}$ allows for the topic distributions in some sitting periods to be different than expected based on government and election effects. This allows us to identify large deviations away from the expected distribution, thus helping to identify the effect of other, non-government and non-election events. In addition, this set up also partially pools effects across sitting periods. The $\delta_{s,p}$ values are modelled as:

$$\delta_{s,p} \sim \text{Normal}(0, \sigma^2_{e[s],p}).$$

The variance parameters $\sigma^2_{e[s],p}$ give an indication of the how the variation in topics is changing over election periods. If the estimates of the variance are larger, then there is more variation in the topics discussed within an election period. Non-informative priors

are placed on the variance parameters:

$$\sigma_{e[s],p} \sim \text{Uniform}(0,3).$$

We run the model in JAGS using the `rjags` package of Plummer (2018).

4 Results

Firstly, we describe the results of the CTM approach, which defined 80 unique topics over the period 1901-2017. We then describe the results of the Bayesian analysis model, which identified governments, elections and other events that were associated with a change in the topics discussed.

4.1 Topic modelling

We applied the CTM approach discussed in Appendix B.5 on the processed Hansard text database outlined in Section 2. The main output of interest are the types of topics identified by the model, and the prevalence of each topic for each day of parliamentary discussion.

Our results are for a topic model that had 80 distinct topics. With almost 15,000 days and 80 topics, the analysis model is being fit to more than a million observations. The choice of 80 topics was made as a trade-off between standard diagnostic tests that suggested a larger number of topics would be more appropriate, and the need for the analysis model to be tractable. Those diagnostic tests are detailed in Appendix B.4.

Table 1 lists the top ten words associated with each of the 80 topics. The topics cover areas such as budgets, transport and infrastructure, war and conflict, health, education, agriculture, and trade. There are also some topics that appear to be about procedural or day-to-day, such as Topics 7 or 9. Note that, as expected, some topics seem to somewhat overlap with their content: for instance, Topics 4, 26, 28, 30 and 66 all relate to war and

conflict.

Topic	Terms
1	women, rights, marriage, human, discrimination, law, equal, community, society, support
2	death, compensation, injury, estate, abolition, accident, injured, deaths, loss, died
3	constitution, parliament, power, powers, constitutional, referendum, convention, representatives, proposal, section
4	defence, forces, personnel, army, military, defence_force, equipment, base, aircraft, air
5	party, communist, matter, time, mckenna, communists, organization, country, position, henty
6	vietnam, countries, south, china, united_states, world, aid, asia, country, foreign
7	petition, petitioners, citizens, pray, parliament, assembled, representatives, duty, bound, undersigned
8	sugar, industry, bounty, queensland, growers, production, fruit, cotton, ton, paid
9	na, senate, president, question, greens, time, committee, support, australians, country
10	service, public, board, officers, officer, department, salary, commissioner, appointment, salaries
11	senators, ill, time, gardiner, measure, western_australia, collings, position, read, leader
12	television, broadcasting, service, stations, radio, post, services, commercial, abc, telephone
13	senate, senators, chamber, representatives, representing, business, party, week, position, public
14	workers, employees, relations, industrial, employers, workplace, employment, employer, union, business
15	president, sympathy, public, word, world, personal, regret, presiding, standing, bias
16	commission, report, royal, royal_commission, inquiry, evidence, commissioner, commissions, body, appointed
17	tax, income, taxation, sales, treasurer, per_cent, pay, revenue, rate, taxes
18	fl, senate, question, greens, time, carbon, president, support, change, move
19	industrial, union, arbitration, workers, unions, trade, court, industry, conciliation, employers
20	matter, labor_party, question, debate, situation, time, organisation, lo, cavanagh, greenwood
21	department, matter, time, question, regard, money, expenditure, connection, business, information
22	per_cent, ad, val, subitem, item, omitting, inserting, exceeding, duty, intermediate
23	court, law, high_court, justice, federal, courts, attorneygeneral, legal, judge, tribunal
24	debate, time, labor_party, issue, deputy, political, question, matter, country, process
25	life, superannuation, fund, insurance, scheme, funds, national, contributions, retirement, age
26	security, iraq, support, detention, international, intelligence, time, world, australias, terrorism
27	education, schools, students, school, university, universities, training, student, funding, children
28	war, production, country, matter, control, governments, industry, time, prices, services
29	health, medical, hospital, private, insurance, medicare, hospitals, scheme, services, public
30	defence, military, naval, training, navy, forces, officers, time, force, service
31	matter, information, letter, department, evidence, document, documents, report, statement, office
32	british, great_britain, germany, empire, trade, country, canada, new_zealand, imperial, conference

Topic	Terms
33	committee, report, parliament, committees, public, parliamentary, recommendations, time, joint, inquiry
34	immigration, country, migration, migrants, citizenship, policy, immigrants, citizens, english, countries
35	question, time, debate, standing_orders, matter, business, parliament, standing, chairman, chair
36	tasmania, queensland, western_australia, new_south_wales, south_australia, victoria, federal, south, tasmanian, premier
37	rules, december, service, hon, association, january, november, october, july, june
38	housing, building, homes, home, capital, site, houses, new_south_wales, construction, canberra
39	question, department, answer, notice, provided, services, total, staff, nil, ii
40	shipping, ships, ship, vessels, line, trade, port, vessel, ports, sea
41	amendments, subsection, section, schedule, omit, item, line, substitute, person, title
42	prime_minister, party, country, leader, parliament, policy, opposite, political, time, election
43	aircraft, aviation, air, airport, airlines, transport, civil, qantas, airline, services
44	duty, per_cent, item, duties, imported, committee, revenue, industry, article, manufacturers
45	main, electorate, million, committee, community, regional, services, per_cent, time, program
46	late, parliament, loss, time, lost, public, memorial, friend, passing, regret
47	report, senate, matter, democrats, governments, leave, per_cent, aboriginal, program, button
48	roads, road, water, railway, line, transport, construction, river, country, money
49	development, time, service, national, programme, overseas, field, matter, department, country
50	world, nations, international, united_nations, countries, treaty, peace, japan, united_states, japanese
51	tariff, industry, trade, industries, board, customs, protection, country, duties, duty
52	pension, pensions, pensioners, week, social, age, benefits, service, services, repatriation
53	community, electorate, na, time, local, support, australians, day, ms, national
54	northern_territory, territory, regulations, regulation, parliament, council, governor_general, ordinance, territories, administration
55	ill, money, position, country, time, amount, financial, matter, treasurer, dr
56	environment, heritage, nsw, environmental, conservation, community, project, management, forest, council
57	energy, gas, nuclear, fuel, change, industry, emissions, power, climate, carbon
58	question, department, matter, answer, time, notice, information, report, questions, national
59	per_cent, budget, increase, economic, country, unemployment, economy, inflation, increased, time
60	care, aged, veterans, community, services, support, home, nursing, child_care, childcare
61	bank, commonwealth_bank, banking, private, credit, money, savings, treasurer, board, trading
62	research, tobacco, scientific, fishing, disease, quarantine, science, fisheries, health, fish
63	law, person, offence, criminal, police, crime, offences, evidence, penalty, attorneygeneral
64	agreement, trade, local, grants, governments, council, development, financial, national, conference
65	democrats, issue, committee, question, time, million, pmi, report, issues, process

Table 1: Top words associated with each topic *(continued)*

Topic	Terms
66	war, soldiers, service, country, returned, time, ill, forces, military, soldier
67	electoral, vote, election, voting, votes, system, party, electors, elections, candidates
68	land, settlement, property, lands, country, lease, leases, pastoral, money, acres
69	information, amendments, support, ensure, report, services, national, review, financial, provide
70	wheat, wool, growers, industry, farmers, board, prices, scheme, marketing, production
71	industry, export, meat, market, dairy, farmers, producers, levy, wine, rural
72	expenditure, loan, increase, revenue, amount, total, budget, money, financial, estimated
73	question, time, matter, desire, regard, learned, opinion, deal, position, party
74	budget, tax, billion, million, per_cent, business, economy, support, jobs, governments
75	millen, senators, question, mcgregor, de, dobson, givens, lt, clemons, time
76	company, oil, companies, industry, coal, profits, capital, business, private, mining
77	per_cent, industry, tax, policy, time, governments, economic, system, program, national
78	clause, provision, section, proposed, agreed, committee, words, provisions, person, matter
79	aboriginal, per_cent, program, governments, commission, question, assistance, funds, development, deputy
80	family, children, families, parents, income, welfare, time, poverty, allowance, parent

Figure 1 illustrates the CTM output based on a sample from the Hansard. It shows how each day's parliamentary discussion can be apportioned to a topic and highlights how these proportions change over time. The highlighted topics are those related to war and conflict. As expected, the two chambers of the Australian Federal Parliament are fairly similar.

4.2 *Analysis model*

4.2.1 *Main results*

We are interested in considering the effect of various political and other events on what is talked about in the Australian Federal Parliament. As discussion in Section 3.2, the next stage of the modelling process takes the topic proportions estimated in the CTM approach, and models the association between these topic distributions and outside events.
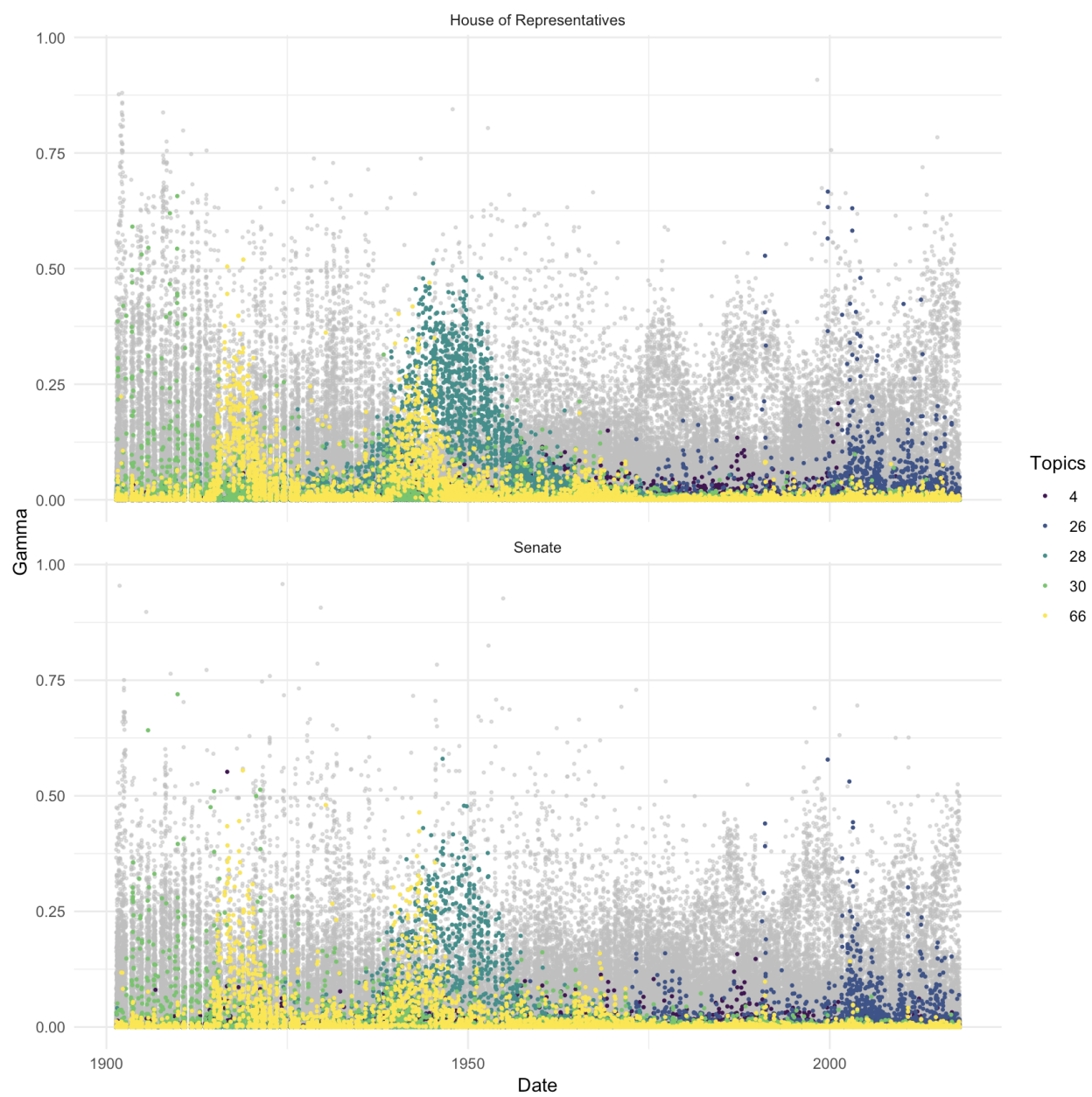
15

Figure 1: Illustrative topic model output, with five topics highlighted

That is, the topic model outputs are inputs for our analysis model.

There are several outputs of interest from this modelling stage. For example, the model provides estimates of topic prevalence by each sitting period. This nicely illustrates how the topics change over time, as the daily estimates tend to be quite variable, but using periods defined by governments or elections tend not to provide enough variability. For instance, examining Topics 26, 28, 30 and 66, which have to do with war and conflict illustrates Australia's involvement in World War I, World War II, the Korean War, the Vietnam War, the First Gulf War, the War in Afghanistan and the Second Gulf War (Figure 2).
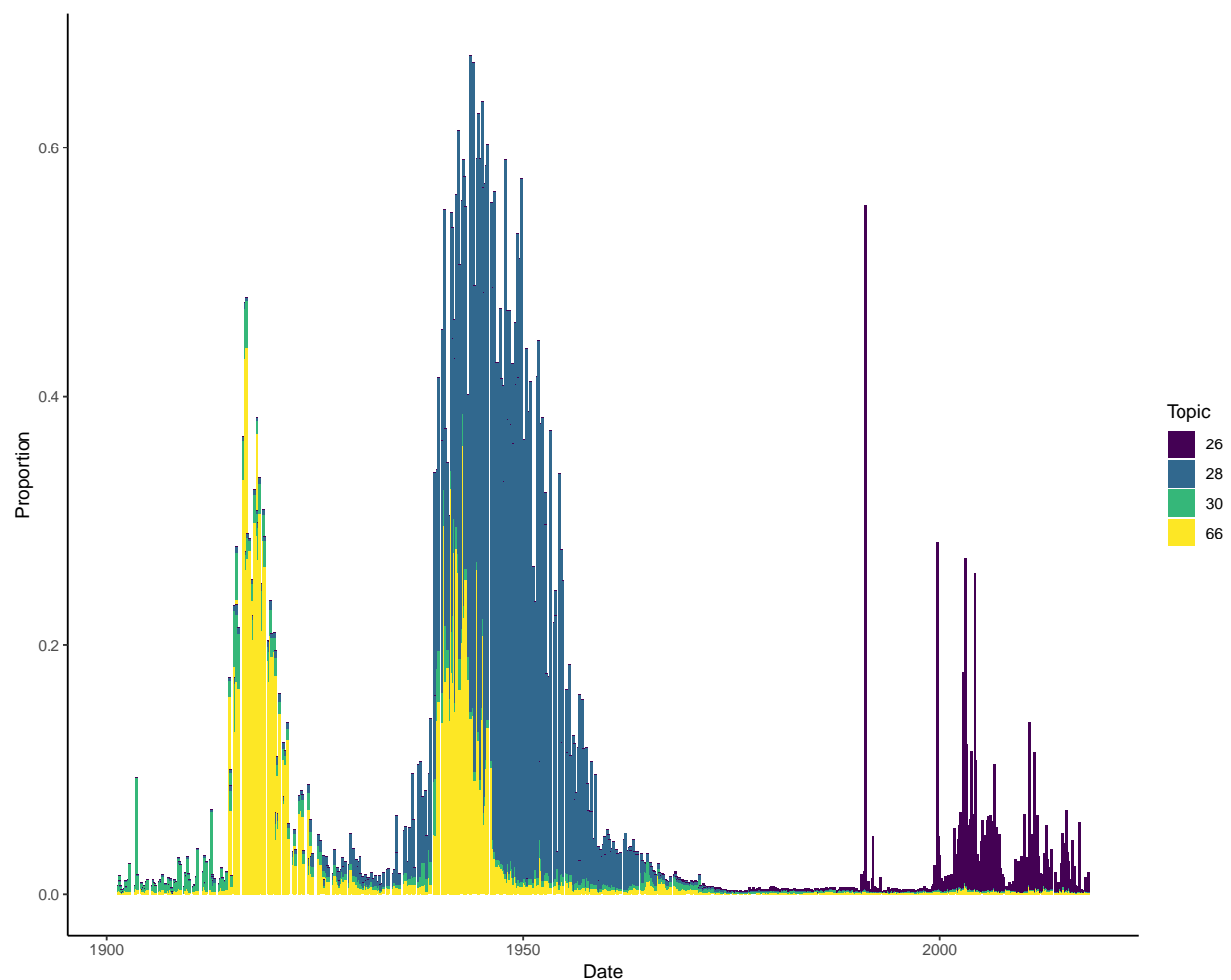

Figure 2: Model estimates of topic prevalance by sitting period for those related to war and conflict

One of the main goals of the analysis model is to see which political events are associated with changes in the prevalence of topics over time. Political events are those related to a change of government or an election. Note that as Australia has a parliamentary system it is possible for the government to change without an election as we define a government based on who is the prime minister, and do not distinguish between terms or cabinet composition as is sometimes done. If a person was prime minister more than once then these periods are considered independently.

As detailed in Section 3.2, the model estimates a mean level-effect for each government, α_g and each election, α_e. We identify differences between neighbouring governments and between neighbouring elections based on calculating 95 per cent credible intervals from posterior samples of these respective mean effects. When these do not overlap we consider that the model finds a significant difference between either the neighbouring governments or elections.

We summarise our results in terms of governments in Table 2 and in terms of elections in Table 3. These tables focus on elections and governments that were different to the ones that preceded them. Complete lists of the Australian elections and governments are available in Appendices C.1 and C.2.

In Figures 3 and 4 we focus on certain topics to illustrate significant differences between governments and elections, respectively. In the graphs, the points show the estimated value of α_g and α_e, respectively, for each of the topics specified. The error bars represent 95 per cent Bayesian credible intervals. We have focused on Topics 4, 26, 28, 30 and 66 here which have to do with war and conflict.

4.2.2   Additional results

Once government and election effects have been taken into consideration, some days stand out compared with others in their sitting period. We do not explicitly include them in the model because of over-fitting and effect-type concerns, but we are interested to

Table 2: Governments that were significantly different to their predecessor

Number	Government	Start	End
6	Fisher 1	1908-11-13	1909-06-02
9	Cook	1913-06-24	1914-09-17
10	Fisher 3	1914-09-17	1915-10-27
12	Bruce	1923-02-09	1929-10-22
14	Lyons	1932-01-06	1939-04-07
15	Page	1939-04-07	1939-04-26
16	Menzies 1	1939-04-26	1941-08-28
20	Chifley	1945-07-13	1949-12-19
21	Menzies 2	1949-12-19	1966-01-26
22	Holt	1966-01-26	1967-12-19
24	Gorton	1968-01-10	1971-03-10
25	McMahon	1971-03-10	1972-12-05
26	Whitlam	1972-12-05	1975-11-11
27	Fraser	1975-11-11	1983-03-11
28	Hawke	1983-03-11	1991-12-20
29	Keating	1991-12-20	1996-03-11
30	Howard	1996-03-11	2007-12-03
31	Rudd 1	2007-12-03	2010-06-24
32	Gillard	2010-06-24	2013-06-27
34	Abbott	2013-09-18	2015-09-15

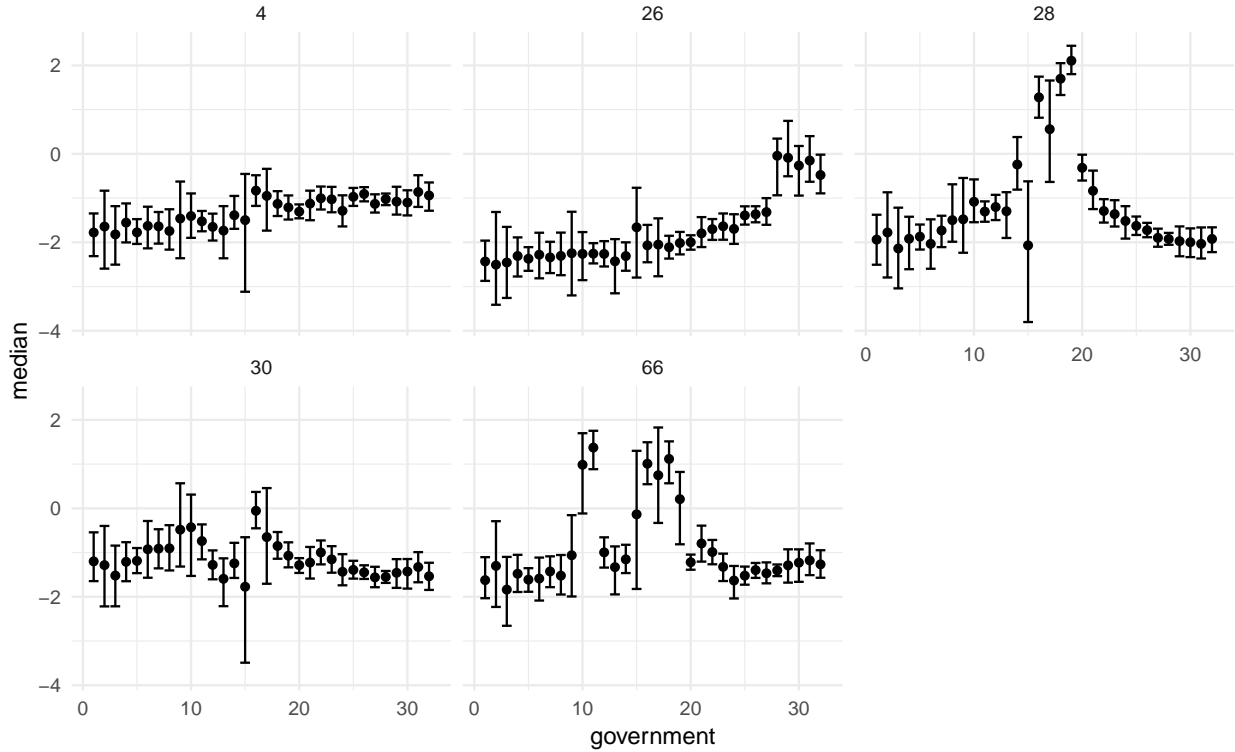Figure 3: Government level effects on selected topics.

Table 3: Elections that were significantly different to the one before

Number	Year	Date	Total seats	Election winner
9	1922	1922-12-16	75	Non-labor
10	1925	1925-11-14	75	Non-labor
19	1949	1949-12-10	121	Non-labor
21	1954	1954-05-29	121	Non-labor
22	1955	1955-12-10	122	Non-labor
23	1958	1958-11-22	122	Non-labor
28	1972	1972-12-02	125	Labor
30	1975	1975-12-13	127	Non-labor
32	1980	1980-10-18	125	Non-labor
33	1983	1983-03-05	125	Labor
34	1984	1984-12-01	148	Labor
35	1987	1987-07-11	148	Labor
36	1990	1990-03-24	148	Labor
38	1996	1996-03-02	148	Non-labor
39	1998	1998-10-03	148	Non-labor
40	2001	2001-11-10	150	Non-labor
41	2004	2004-10-09	150	Non-labor
42	2007	2007-11-24	150	Labor
43	2010	2010-08-21	150	Labor
44	2013	2013-09-07	150	Non-labor

Note:
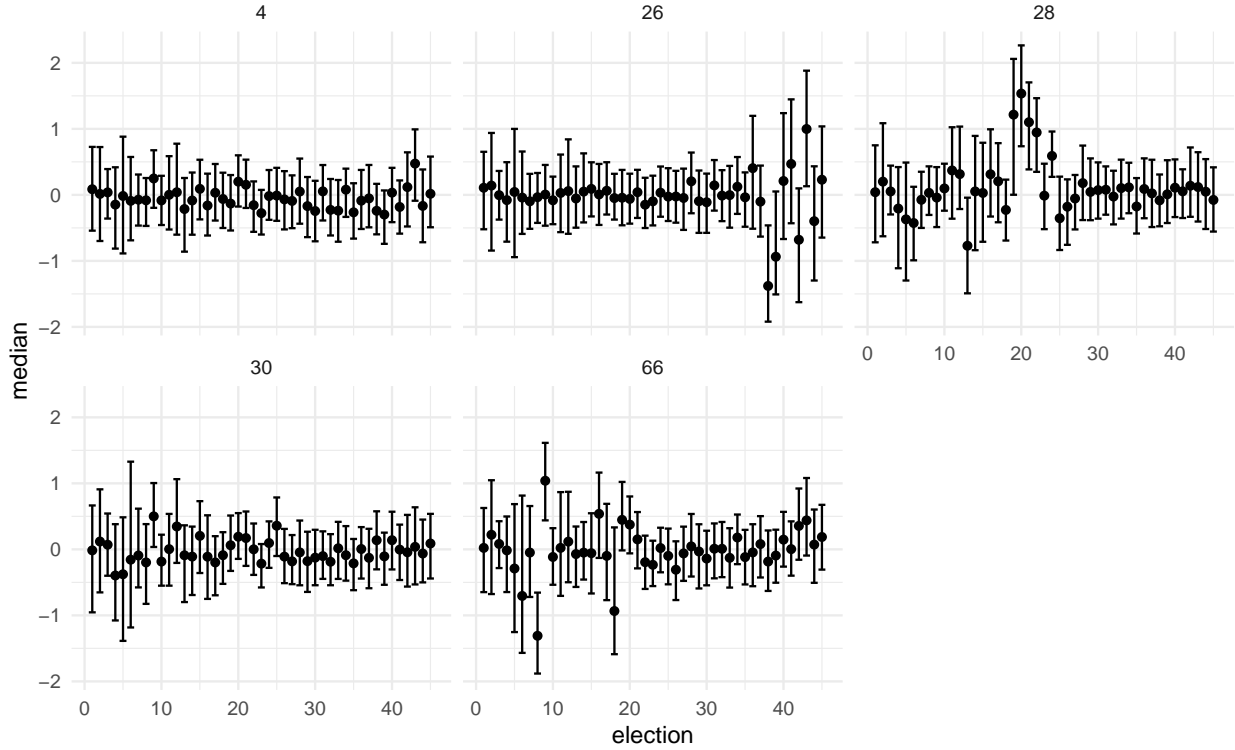Here is a general comments of the table.

Figure 4: Elections level effects on selected topics

see if these can be explained by events that occurred on or before that sitting day. For instance, we may expect that events of a historical magnitude, such as the 9/11 attacks would change the discussion, or that the sitting day when, for instance, the Apology to the Stolen Generation was delivered, or some particularly prominent legislation introduced, would be different to others in that sitting period.

To do this we estimate sitting-period level-effects (essentially a mean for each topic by sitting period). The difference between this mean distribution and a particular day's topic distribution defines a measure that can be thought of as essentially a residual which allows us to identify outlying days. This approach means that the model generates dates that are interesting without us having to specify interesting dates. More specifically, we define a day to be 'outlying' or 'different-to-expected' if the topic distribution on that particular day is more than three standard deviations different to the mean topic distribution for the relevant sitting period. Table 4 summarises the days where parliamentary

Table 4: Days that were significantly different to other days in their sitting period

Dates							
1902-04-16	1913-10-31	1938-05-10	1947-03-19	1951-10-11	1959-09-22	1964-09-16	1997-12-06
1902-04-17	1920-05-11	1938-05-18	1947-04-22	1951-10-16	1960-10-25	1964-10-13	1998-04-09
1903-06-09	1920-05-14	1938-05-19	1947-09-24	1953-11-17	1960-09-21	1965-03-25	1999-03-23
1904-11-18	1920-05-18	1938-11-30	1947-09-25	1955-09-14	1960-09-22	1965-05-04	2000-09-05
1904-11-22	1921-04-15	1941-11-20	1948-04-15	1956-06-06	1960-11-16	1965-05-06	2000-11-30
1907-09-19	1924-06-05	1942-05-19	1948-11-19	1956-09-26	1961-09-07	1965-09-28	2001-09-17
1908-05-20	1926-03-17	1943-03-23	1949-02-09	1957-10-03	1961-09-12	1965-09-29	2002-10-14
1908-11-04	1930-05-15	1945-02-21	1949-02-16	1957-10-08	1961-09-13	1966-08-16	2005-05-10
1908-11-12	1930-06-13	1945-08-30	1949-03-01	1957-11-14	1962-11-14	1967-04-11	2007-09-18
1910-09-02	1931-05-07	1945-09-05	1949-03-02	1957-11-19	1963-09-17	1969-05-20	2009-02-04
1911-09-26	1931-05-01	1945-09-06	1949-07-06	1957-11-21	1964-03-17	1971-10-12	2010-03-16
1911-09-27	1931-05-05	1946-03-26	1950-03-22	1957-11-26	1964-08-13	1978-11-24	2010-09-28
1912-11-12	1932-05-17	1946-07-16	1950-11-23	1958-08-06	1964-09-01	1985-11-29	2013-02-12
1913-09-23	1933-10-24	1946-07-17	1951-07-05	1958-08-07	1964-08-11	1997-06-27	2013-02-13

discussion was significantly different from the rest prevailing in that week.

5 Discussion

Of the 36 governments over this period, we find that 20 of them are significantly different to the government that preceded them. However, four of these results – the significance of the Page Government, the First Menzies Government, the Chifley Government and the Abbott Government – are likely due to the short length of either the government or its predecessor and should be ignored, leaving only 16. The earliest of these are the First Fisher Government, the Cook Government and the Third Fisher Government, which were different to the Second Deakin Government, the Second Fisher Government, and the Cook Government, respectively. To a certain extent this is likely due to changes due to World War I.

The Second Menzies Government, begining in 1949, is the next government that is significantly different to its predecessor. The other governments that are different are concentrated in the second half of our sample, with three of them being in the past twenty years. Similarly, of the 45 general elections that have been held we find that 14 of them de-

fine periods that were significantly different to their immediate predecessor. 1974, 1980, 1990, 1998, 2004, and 2007 stand out as elections where the government did not change, but the model suggests there was considerable change in the topics discussed in parliament.

The Second Menzies Government was associated with a variety of changes compared with the preceding Chifley Government. The Chifley Government had governed through the end of World War II and the difficult economic times that followed. There was also a large increase in the number of seats in the House of Representatives at the 1949 election. Many new politicians entered parliament, and this changed representation may also have been partly to do with the changed distribution of discussion topics. The sixteen-year length of the Second Menzies Government, and better economic conditions over this time make it understandable why parliamentary discussion would have been different. There were six elections within the Second Menzies Government. Three toward the middle of that government were associated with significant changes in the topics discussed.

The Menzies Government was succeeded by the Holt Government in January 1966. This is an example where there was a change in government without an election, as the next election only happened in November 1966. We find that the Holt Government is different to the Menzies Government. In Figures 5 we compare the topics during the final term of the Menzies Government with the topics of the Holt Government.

The Whitlam government is especially interesting as we find a difference in the topics after it was first elected in December 1972, compared with its second election win in May 1974. Figure 5 compares the topics that are significant in the first Whitlam term and then compares them to those in the second Whitlam government.

The Howard Government is also interesting because of the significant differences between elections. For instance, each of the election periods is associated with fairly substantial differences compared with the preceding election periods, and all are actually significantly different at the 95 per cent level. Figure 5 compares the topics that are sig-
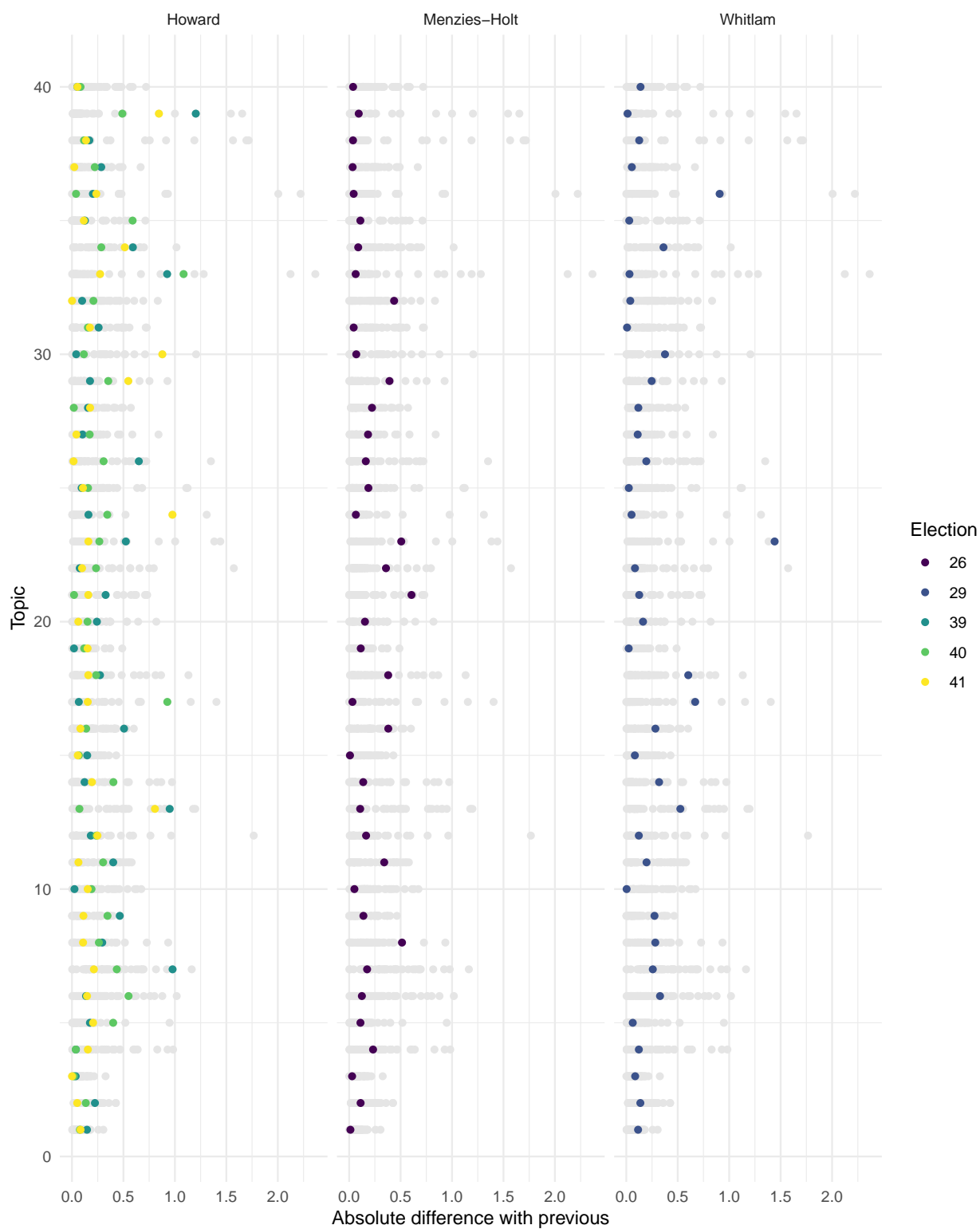
Figure 5: Differences between various elections

nificant in the different Howard terms.

To a certain extent the change after the November 2001 election is expected because of the 9/11 terrorist attacks that had only occurred two months earlier, the Bali Bombings that occurred in October 2002, and the dramatic increase in the discussion related to terrorism and conflict over these years. However, the change in 1998 and 2004 is more unexpected. Although the Howard Government is the second-longest serving government and commonly thought of as a period of stability because the senior ministers were consistent as well, it might be that it is better to think of the Howard Government as a combination of three or four different periods and that the Howard Government reinvented itself over this period.

One advantage of our analysis model compared with using the STM approach is that we can create a measure that is equivalent to testing for outliers in a model where the underlying variables were not latent. The results of this reduction in supervision are promising, but suggest the specifics of our process may need further refinement. For instance, our approach appropriately identifies the sitting day that first follows 11 September 2001 and 12 October 2002, which were the dates of the 9/11 Attacks and Bali Bombings respectively. But there are many dates that we would have expected to be identified that were not, and similarly some of the dates that were identified are surprising. When we dig into this we find that some of them, such as 4 February 2009, are associated with significant legislation. However the over-weighting of dates in the first half of our sample in Table 4 highlights further work is needed to improve this approach. For instance, our approach may not be appropriately considering step-changes or it may be that our identification of sitting periods is not appropriate for the entire period.

6 Summary, weaknesses and future work

In this paper we consider what was said in the Australian Federal Parliament between 1901 and 2017. We download and parse PDFs of Hansard to create a new dataset of

text. We use a correlated topic model to group the parliamentary discussion into topics to reduce the dimensionality, and then analyse the effect of various events on the distribution of these topics using a Bayesian hierarchical Dirichlet model. In general we find that changes in government change the distribution of topics discussed in parliament, but that most elections did not. We find that significant events such as 9/11 and the Bali Bombings had a substantial effect, but that with certain exceptions, economic events did not.

By bringing a new dataset of what was said in the Australian Federal Parliament to bear for our analysis we are able to consider events over the full history of the Commonwealth of Australia. However even after cleaning the dataset remains imperfect and is more fit-for-purpose than of broad applicability.

Using text as data allows larger-scale analysis that would not be viable using less-automated approaches and so it can identify associations and patterns that may otherwise be overlooked. That said, text analysis has well-known shortcomings and weaknesses and should be considered a complement to more detailed analysis such as qualitative methods and case studies. For instance, the use of topic models is problematic because of the need to interpret the topics, which at can sometimes feel like interpreting hieroglyphics. This can be difficult especially when the number of topics is large, but in a dataset of the size that we have, a large number of topics is needed. Also, although topic modelling is an unsupervised machine learning technique, the inputs require fine-tuning. For instance, selecting stop words for removal and which word to merge because of common co-location has an impact on the topics. Even after doing this there tend to be topics that are not overly meaningful, especially on their own. One way to get around using topic models is to use the words more directly, for instance word2vec and other approaches. As computational power become cheaper and more appropriate analytical methods, such as Taddy (2015) as applied in Gentzkow, Shapiro and Taddy (2018), are developed this becomes a more feasible options and future research could take that approach.

In this paper we think of events as affecting discussion in parliament. Given the events

that we consider and the broad topics of discussion we consider this the most appropriate direction for this paper. Other papers could consider an alternative specification, especially if they were to focus on particular aspects.

In terms of the analysis model which relates the topic distributions with events, there are several limitations to the model. Firstly, we are assuming that the effect of a particular government is constant across the whole period. In addition, we assume that the effect of elections is monotonically decreasing across days since election. Future work could consider different functional forms on both of these effects, and in particular try to allow for elections to have a 'lead-up' effect.

The way that we identify unusual periods could also be improved. We defined sitting periods in a constant fashion across the whole dataset time frame, but how long the stretches are that parliament sits for has changed over time. In addition, more work needs to be done on how to identify outlying events. For instance, the extent to which an important event that occurs outside a sitting period can be identified has a great deal of uncertainty. And if an event happens in the middle of a sitting period, it may have a large effect on the overall mean, such that specific days are not identified as significantly outlying.

Finally, the current modelling and analysis set-up is a two-stage process: we take the output of a topic model, and use this as the input to a second model. However, this approach does not appropriately propagate the uncertainty of the topic distribution estimation stage. Future methodological work could consider how to combine these two modelling steps, for instance by extending the STM approach into a more flexible framework.

Watching our politicians at work can sometimes be a little disheartening. It can be hard to believe that not only are those in charge shouting insults that would not be tolerated in a schoolyard, but that we voted to put them there. Nonetheless, our work suggests that important topics are discussed in parliament. It is easy to look back and think that we

27

live in uniquely tumultuous times, but our analysis suggests events have always driven debate and that periods of stability may be the exception. However, we do find that since the 1980s differences between government and election periods seem to be more pronounced than they used to be. It is, of course, reasonable that a government should act on the mandate they receive, but it takes a long time to get policy right, and we may find progress being held back if the media and political cycles become too focused on short-term events.

A Hansard details

A.1 Example Hansard page

Figure 6: Example Hansard page – 6 February 1902

The number of sitting days in a year varies considerably. The highest in the House of Representatives was 122 days in 1904, followed by 113 days in 1901 and 1920. The year with the most sitting days in the Senate was 1902 with 93 days, followed by 1989 with 92 days, and 1986 with 86 days (Figure 7).
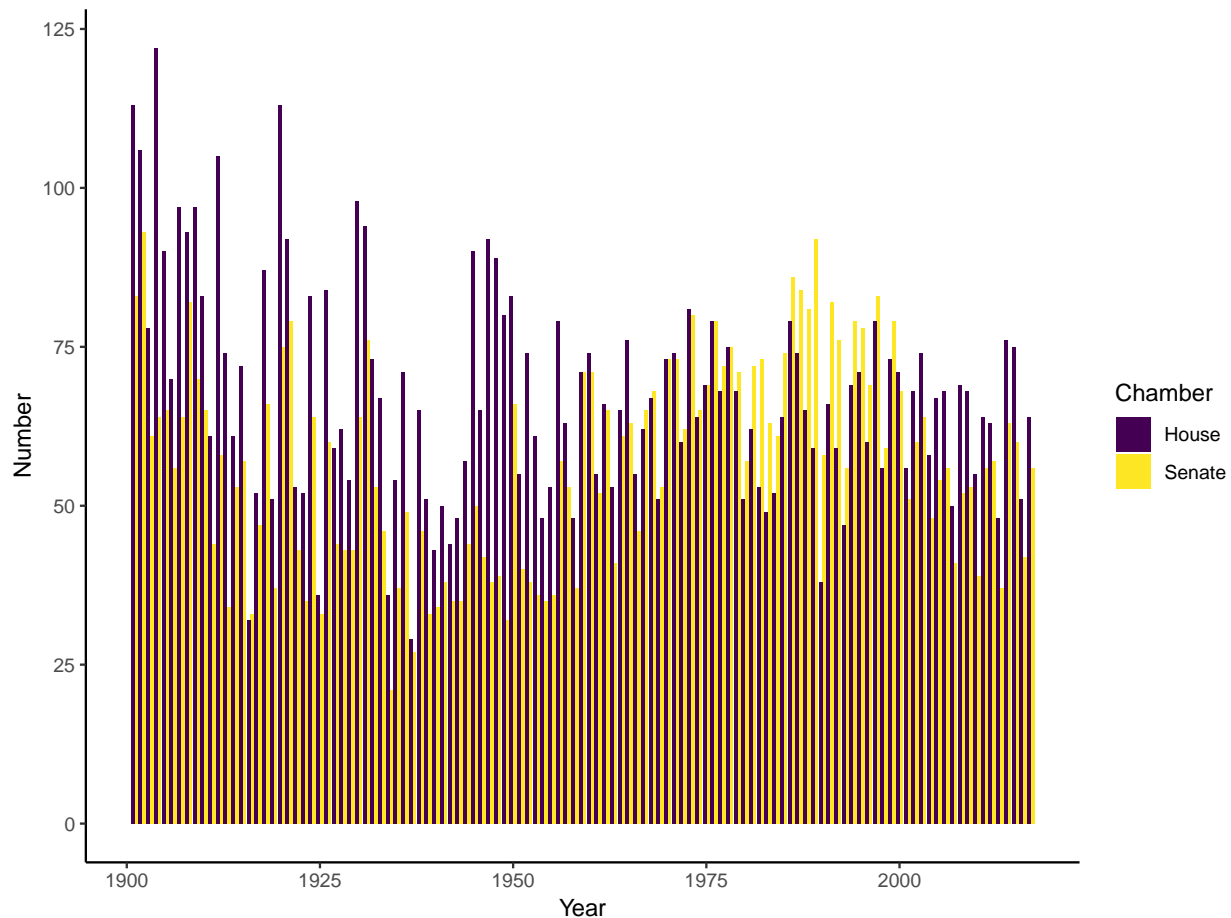

Figure 7: Number of sitting days, by year

Until the 1950s the House of Representatives tended to have more sitting days than the Senate. It was then similar, before the Senate had more days in the 1980s and 1990s. Since the 2000s the House of Representatives again has tended to have more sitting days than the Senate.

These counts of the number of sitting days are based on available PDFs. For this reason

the counts may be slightly different to other counts. An example of one known issue of this type is detailed in the next section.

A.3   Compared with parliamentary website

The parliamentary website provides a summary table of the number of sitting days in each year by chamber.[3] Comparing the numbers provided in that table with number of days that we have provides an indication of how complete our dataset is.

In general the number of sitting days on the parliamentary website summary table is similar to the number of PDFs that we have although it does identify a few particularly concerning years (Figure 8).
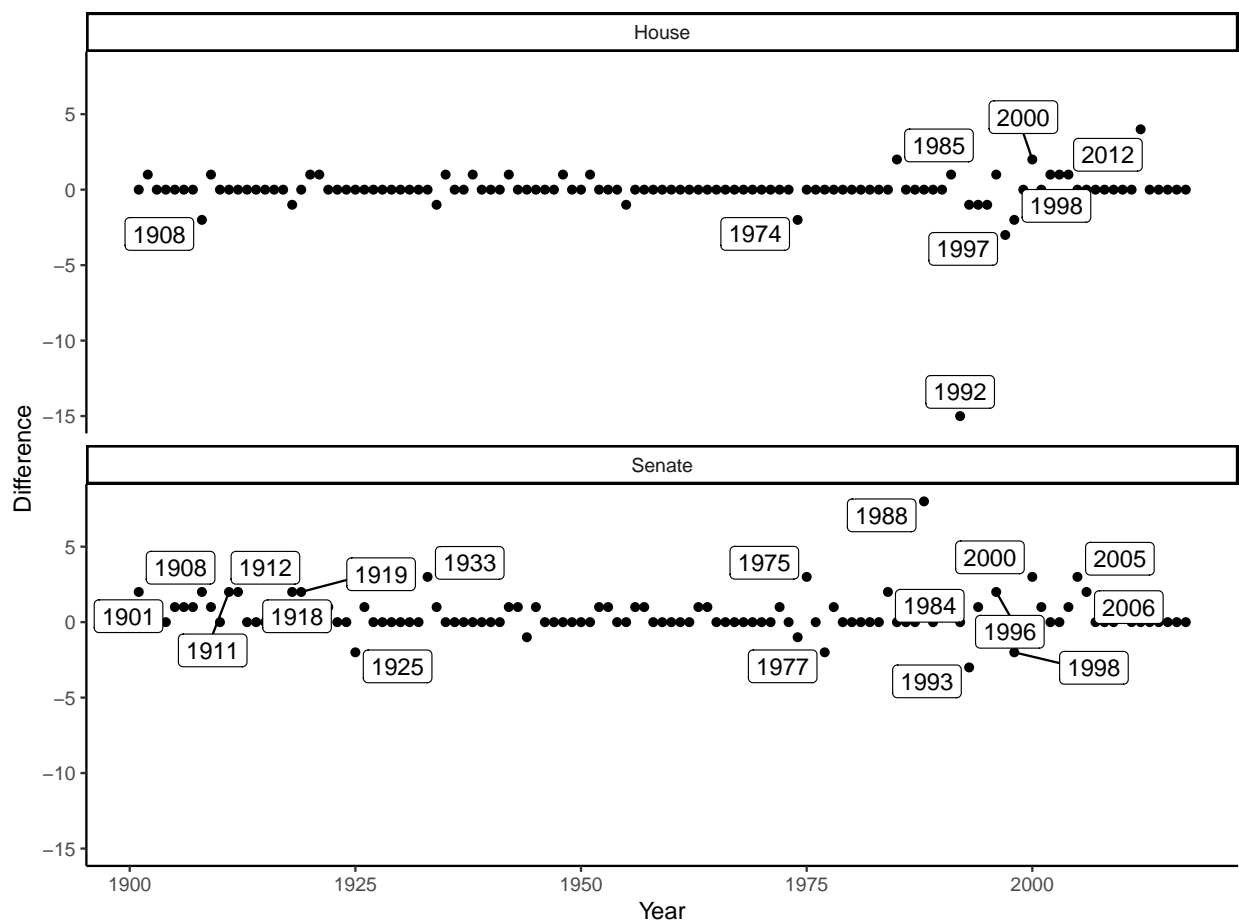

Figure 8: Differences by year between the number of sitting days and our number of PDFs

When the difference is positive, it means that in that year we have fewer PDFs than the parliamentary website claims. For instance, 5 could mean that the parliamentary website

---

[3] As at 5 November, the website was available at: https://www.aph.gov.au/Parliamentary_Business/Statistics/Senate_StatsNet/General/sittingdaysyear.

claimed there were 100 days, but we only had PDFs for 95 days. Similarly, when the difference is negative then we have more PDFs than the parliamentary website claims there were sitting days.

The two major years of concern are 1992 in the House of Representatives where we have 15 days more than the parliamentary website claims there were, and 1988 in the Senate where we have eight days fewer.

The Parliament website is missing the Hansard PDFs for the following dates in the Senate: 1988-12-21, 1988-12-20, 1988-12-19, 1988-12-16, 1988-12-15, 1988-12-14, 1988-12-13, 1988-12-12, 2000-10-12, 2000-06-19, 2004-08-09.

There are two unaccounted for differences in 2006, one unaccounted for difference in 2001.

The Parliament website is missing the Hansard PDFs for the following dates in the House of Representatives: 1985-08-23, 1992-09-10, 1996-12-13, 2002-05-14, 2000-10-12, and 2000-06-29.

There is one unaccounted for difference in 1920, 1921, 1935, 1942, 1948, 1951, 1991, 2003, 2004, and there are two unaccounted for in 1985 and four unaccounted for in 2012.

In the Senate, the PDF for 10 August 1917 may be the wrong one. It says it is for 10 January 1918 on the cover sheet, but there's not an entry for 10 January 1918.

Senate Hansard - 1918-12-18 has the wrong PDF so does 1917-08-01.

A.4   Example Hansard PDF to text record workflow

The exact scripts that we use are too long to provide here but are available on request. Instead here we provide example scripts that summarises the workflow to convert the PDFs to usable text datasets and may be a useful starting point for other researchers. The scripts are primarily based on: the PDFtools R package of Ooms (2018); the tidyverse R package of Wickham (2017); the tm R package of Feinerer and Hornik (2018); the lubridate R package of Grolemund and Wickham (2011); the tidytext R package of Silge and Robinson (2016); and the stringi R package of Gagolewski (2018). The functions of those packages are augmented by: the furrr R package of Vaughan and Dancho (2018); and the tictoc R package of Izrailev (2014). The hunspell R package of Ooms (2017) is used to help find spelling issues; and the quanteda R package of Benoit (2018) is used to compound multiword expressions.

[Insert the example scripts.]

A.5 Stopwords over time

[Insert the graph of stop words over time.]

B Topic modelling example and details

B.1 Overview and example

As applied to Hansard, LDA considers each statement to be a result of a process where a politician first chooses the topics they want to speak about. After choosing the topics, the politician then chooses appropriate words to use for each of those topics. Statistically, LDA considers each document as having been generated by some probability distribution over topics. Similarly, each topic is considered a probability distribution over terms. To choose the terms used in each document, terms are picked from each topic in the appropriate proportion.

As an example, Figures 9 and 10 illustrate a smaller application with five topics, two documents, and ten terms. In this case, the first document may be comprised mostly of the first few topics; the other document may be mostly about the final few topics (Figure 9).
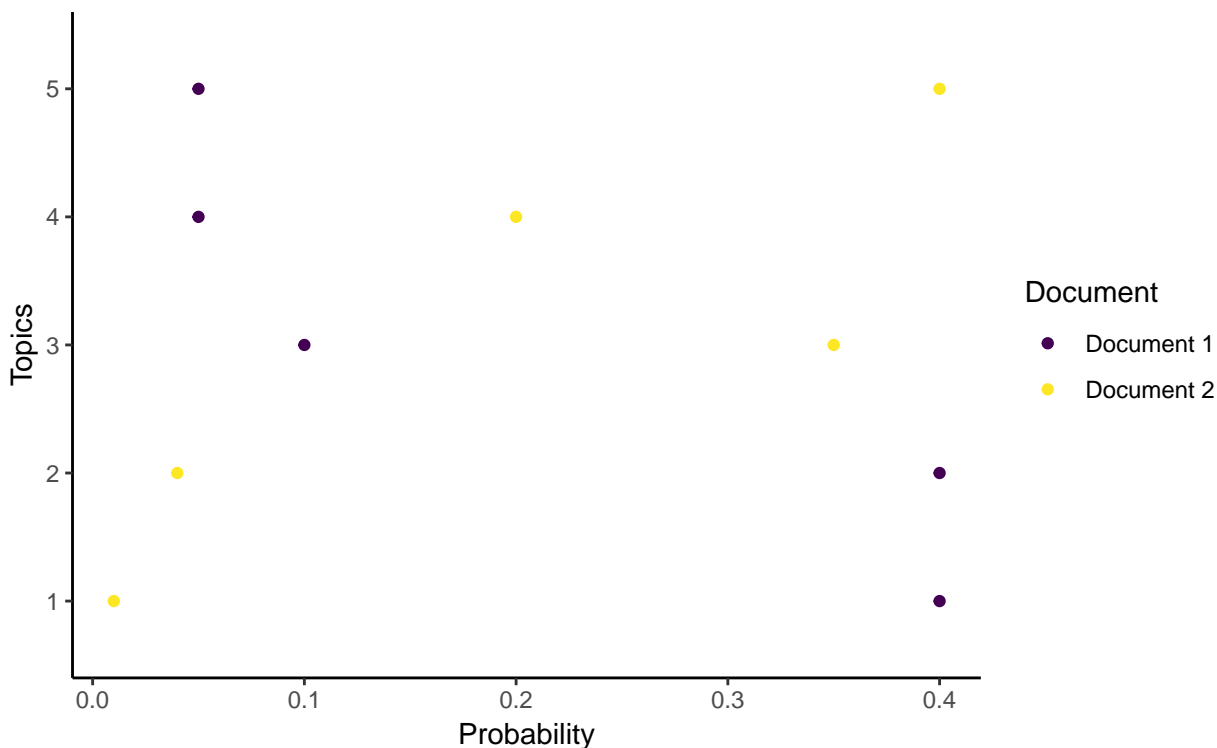

Figure 9: Probability distributions over topics for two documents

For instance, if there were ten terms, then one topic could be defined by giving more weight to terms related to immigration; and some other topic may give more weight to terms related to the economy (Figure 10).
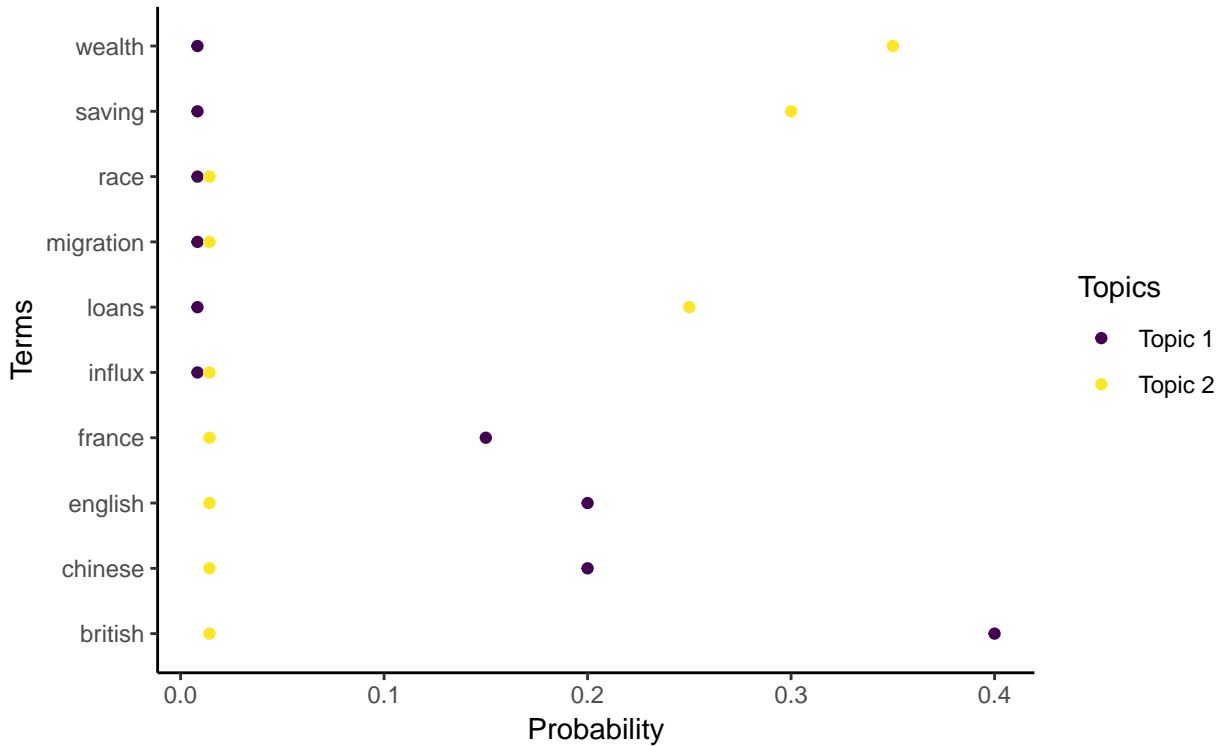

Figure 10: Probability distributions over terms

B.2 Document generation process

Following Blei and Lafferty (2009), Blei (2012) and Griffiths and Steyvers (2004), the process by which a document is generated is more formally considered to be:

1. There are $1, 2, \ldots, k, \ldots, K$ topics and the vocabulary consists of $1, 2, \ldots, V$ terms. For each topic, decide the terms that the topic uses by randomly drawing distributions over the terms. The distribution over the terms for the kth topic is β_k. Typically a topic would be a small number of terms and so the Dirichlet distribution with hyper-parameter $\boldsymbol{\eta}$ is used: $\beta_k \sim \text{Dirichlet}(\boldsymbol{\eta})$, where $\boldsymbol{\eta} = (\eta_1, \eta_2, \ldots, \eta_K)$.[4] In practice, a symmetric Dirichlet distribution is usually used, where all elements of $\boldsymbol{\eta}$ are equal.

2. Decide the topics that each document will cover by randomly drawing distributions over the K topics for each of the $1, 2, \ldots, d, \ldots, D$ documents. The topic distributions for the dth document are θ_d, and $\theta_{d,k}$ is the topic distribution for topic k in document d. Again, the Dirichlet distribution with the hyper-parameter $0 < \alpha < 1$ is used here because usually a document would only cover a handful of topics: $\theta_d \sim \text{Dirichlet}(\boldsymbol{\alpha})$. Again, strictly $\boldsymbol{\alpha}$ is vector of length K of hyper-parameters and they are usually equal.

3. If there are $1, 2, \ldots, n, \ldots, N$ terms in the dth document, then to choose the nth term, $w_{d,n}$:

    a. Randomly choose a topic for that term n, in that document d, $z_{d,n}$, from the multinomial distribution over topics in that document, $z_{d,n} \sim \text{Multinomial}(\theta_d)$.

    b. Randomly choose a term from the relevant multinomial distribution over the terms for that topic, $w_{d,n} \sim \text{Multinomial}(\beta_{z_{d,n}})$.

[4]The Dirichlet distribution is a variation of the beta distribution that is commonly used as a prior for categorical and multinomial variables. If there are just two categories, then the Dirichlet and the beta distributions are the same. In the special case of a symmetric Dirichlet distribution, where all elements of $\eta = 1$, it is equivalent to a uniform distribution. If $\eta < 1$, then the distribution is sparse and concentrated on a smaller number of the values, and this number decreases as η decreases. A hyper-parameter is a parameter of a prior distribution.

Given this set-up, the joint distribution for the variables is (Blei (2012), p.6):

$$p(\beta_{1:K}, \theta_{1:D}, z_{1:D,1:N}, w_{1:D,1:N}) = \prod_{i=1}^{K} p(\beta_i) \prod_{d=1}^{D} p(\theta_d) \left(\prod_{n=1}^{N} p(z_{d,n}|\theta_d) p(w_{d,n}|\beta_{1:K}, z_{d,n}) \right).$$

Based on this document generation process the analysis problem, discussed next, is to compute a posterior over $\beta_{1:K}$ and $\theta_{1:D}$, given $w_{1:D,1:N}$. This is intractable directly, but can be approximated (Griffiths and Steyvers (2004) and Blei (2012)).

After the documents are created, they are all that we have to analyse. The term usage in each document, $w_{1:D,1:N}$, is observed, but the topics are hidden, or 'latent'. We do not know the topics of each document, nor how terms defined the topics. That is, we do not know the probability distributions of Figures 9 or 10. In a sense we are trying to reverse the document generation process – we have the terms and we would like to discover the topics.

If the earlier process around how the documents were generated is assumed and we observe the terms in each document, then we can obtain estimates of the topics (Steyvers and Griffiths (2006)). The outcomes of the LDA process are probability distributions and these define the topics. Each term will be given a probability of being a member of a particular topic, and each document will be given a probability of being about a particular topic. That is, we are trying to calculate the posterior distribution of the topics given the terms observed in each document (Blei (2012), p. 7):

$$p(\beta_{1:K}, \theta_{1:D}, z_{1:D,1:N}|w_{1:D,1:N}) = \frac{p(\beta_{1:K}, \theta_{1:D}, z_{1:D,1:N}, w_{1:D,1:N})}{p(w_{1:D,1:N})}.$$

Gibbs sampling or the variational expectation-maximization algorithm can be used to approximate the posterior. A summary of these approaches is provided next.

B.3 Posterior estimation

Following Steyvers and Griffiths (2006) and Darling (2011), the Gibbs sampling process attempts to find a topic for a particular term in a particular document, given the topics of all other terms for all other documents. Broadly, it does this by first assigning every term in every document to a random topic, specified by Dirichlet priors with $\alpha = \frac{50}{K}$ and $\eta = 0.1$ (Steyvers and Griffiths (2006) recommends $\eta = 0.01$), where α refers to the distribution over topics and η refers to the distribution over terms (Grün and Hornik (2011), p. 7). It then selects a particular term in a particular document and assigns it to a new topic based on the conditional distribution where the topics for all other terms in all documents are taken as given (Grün and Hornik (2011), p. 6):

$$p(z_{d,n} = k | w_{1:D,1:N}, z'_{d,n}) \propto \frac{\lambda'_{n \to k} + \eta}{\lambda'_{\cdot \to k} + V\eta} \frac{\lambda'^{(d)}_{n \to k} + \alpha}{\lambda'^{(d)}_{-i} + K\alpha}$$

where $z'_{d,n}$ refers to all other topic assignments; $\lambda'_{n \to k}$ is a count of how many other times that term has been assigned to topic k; $\lambda'_{\cdot \to k}$ is a count of how many other times that any term has been assigned to topic k; $\lambda'^{(d)}_{n \to k}$ is a count of how many other times that term has been assigned to topic k in that particular document; and $\lambda'^{(d)}_{-i}$ is a count of how many other times that term has been assigned in that document. Once $z_{d,n}$ has been estimated, then estimates for the distribution of words into topics and topics into documents can be backed out.

This conditional distribution assigns topics depending on how often a term has been assigned to that topic previously, and how common the topic is in that document (Steyvers and Griffiths (2006)). The initial random allocation of topics means that the results of early passes through the corpus of document are poor, but given enough time the algorithm converges to an appropriate estimate.

The choice of the number of topics, k, drives the results and must be specified *a priori*. If there is a strong reason for a particular number, then this can be used. Otherwise,

one way to choose an appropriate number is to use cross validation. More detail on this process is provided in the next section.

The choice of the number of topics to use in a topic model has a substantial effect on the results of the model. For instance, in our topic model, choosing a smaller number of topics, such as 10 or 20 results in a model that is not all that useful because the topics are so broad.

There are a variety of diagnostic measures that can guide the selection of the topics, but there is rarely an obvious best choice, especially at a finer level such as choosing between 60 and 65 topics. We found it useful to try a few quite different measures before settling on 80 topics. This provided a balance between being granular enough to be informative–anything less than 40 topics tended to be too broad– yet still being tractable for our analysis model in a reasonable amount of time. In addition to looking at the topics and how they changed over time, diagnostic measures that we considered include the held-out likelihood, the lower bound, residuals, exclusivity. and semantic coherence (Figures 11).[5]

Roberts, Stewart and Tingley (2018) provides more detail about the diagnostic tests that we use, but we briefly discuss each here. Exclusivity is a measure of how specific words are to particular topics. It looks at the proportion that a word makes up of a particular topic compared with the proportion that word makes up of the other topics. As the number of topics increases we usually expect exclusivity to increase because the topics become more particular. Higher values are better. The held-out likelihood as described by Wallach et al. (2009) takes a test/training approach to estimate the probability of held-out documents given the training documents. Higher values are better. The lower bound gives some indication of whether the model may have multiple modes and hence the end result be sensitive to the starting position (Roberts, Stewart and Tingley (2016)). Residuals analysis (Taddy (2012)) compares the theoretical distribution of the variance with the actual distribution. It is a test for overdispersion of the variance, and if it is found then this

---

[5]The code for creating the figures in this section is based on Silge (2018).
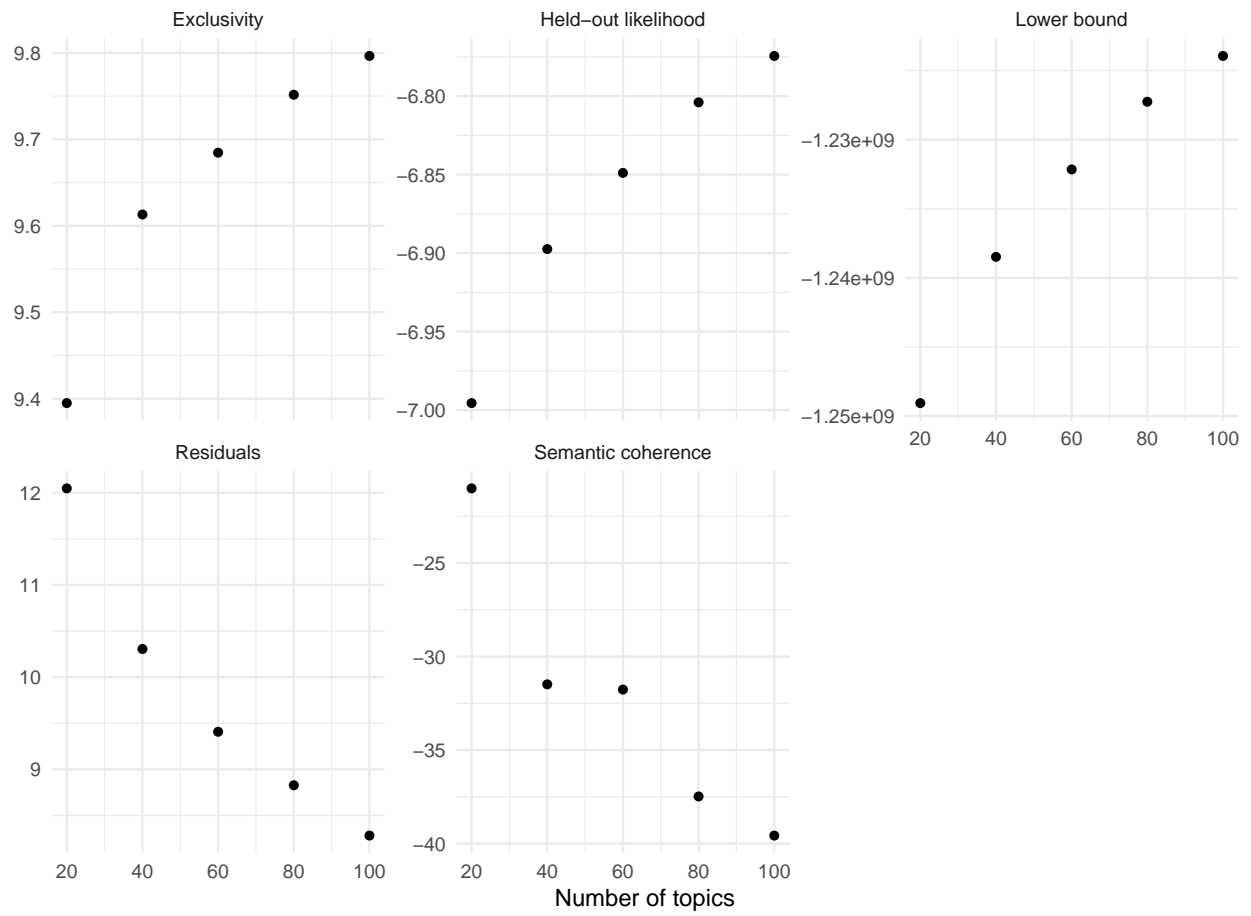
Figure 11: Model diagnostics

can suggest that more topics would be appropriate. Semantic coherence is the trade-off for having topics that are more specific and the subsequent risk that the topics become meaningless. Mimno et al. (2011) define a measure of coherence that is based on ratios of single words compared with pairs of words. The idea is that words that should occur in the one document should be more likely to be in a particular topic than ones that do not occur together. For instance, a topic that has 'wine' and 'cheese' as highly rated words would score better on their measure than another that contained 'cheese' and 'mining'. Lower values are better. Roberts, Stewart and Tingley (2018) recommend examining the trade-off between exclusivity and semantic coherence. This suggests that the magnitude of improvement reduces from about 80 topics (Figure 12).
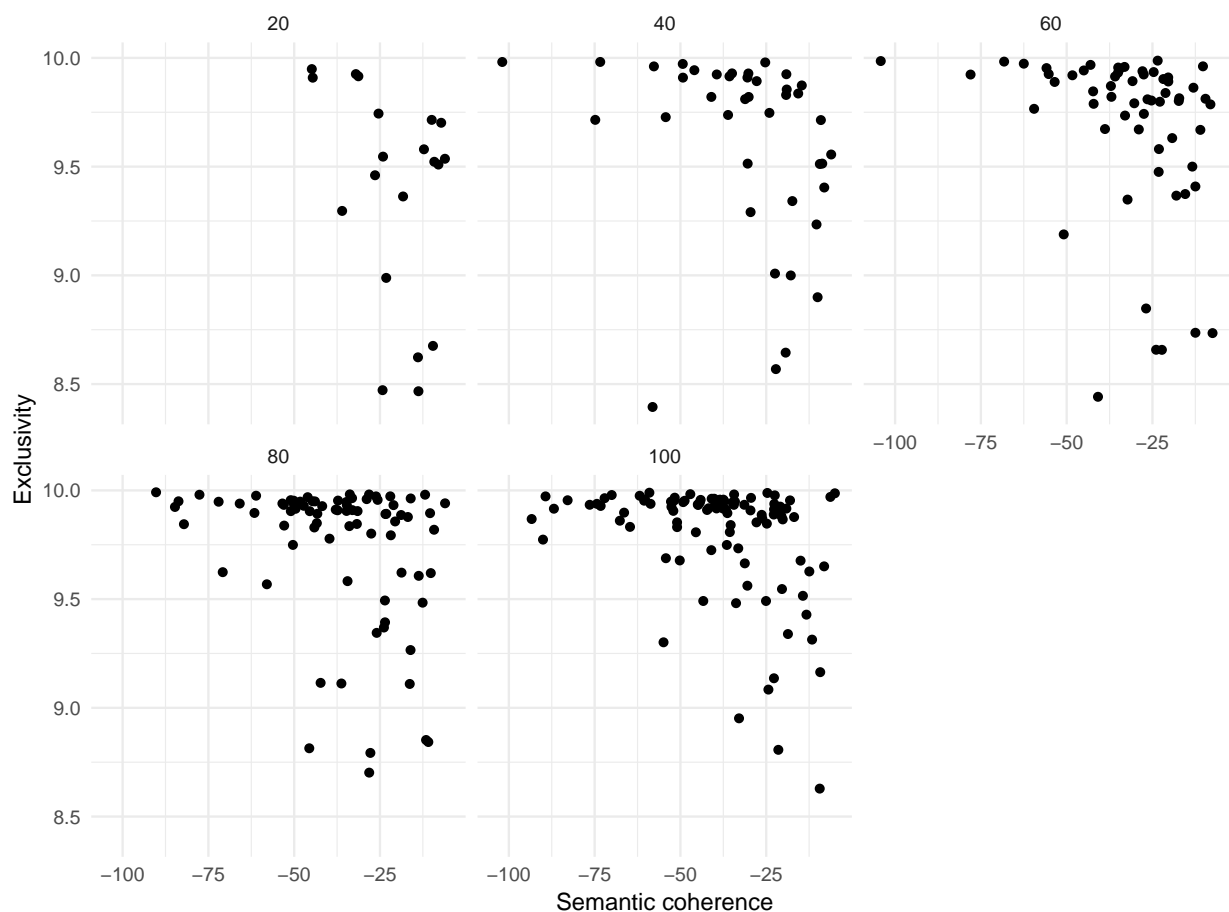

Figure 12: Exclusitivity compared with semantic coherence

B.5 Correlated Topic Model

One of the limitations of LDA is that the model assumes that the presence of one topic is not correlated with the presence of another topic. In reality often topics are related. For instance, in the Hansard context, we may expect topics related to the army to be more commonly found with topics related to the navy, but less commonly with topics related to banking. The goal of the CTM (Blei and Lafferty, 2007) is to account for this correlation between topics, in order to produce more realistic and stable topic distributions over time. The models are very similar, and the key difference is the underlying distributions that are drawn from.

As with LDA, the process assumed to generate the documents is the key aspect as this will be reversed to estimate the topics. The document generation process of Blei, Ng and Jordan (2003), discussed in Appendix B.2, is just slightly modified. Specifically, rather than assuming that the distribution of topics in a document, θ_d, are a draw from a Dirichlet distribution, as in Step 2 of the LDA document generation process detailed in Appendix B.2, CTM assumes:

$$\theta_d \sim \text{Logistic Normal}(\mu, \Sigma).$$

That is, the main difference of CTM over LDA is that it replaces the assumption of the Dirichlet distribution with a more flexible logistic multivariate Normal distribution. This distribution can incorporate a covariance structure across the topics. The remainder of the steps of the document generating process are pretty much the same as LDA.

However, the replacement of the Dirichlet distribution with the logistic multivariate Normal distribution adds a level of computational complexity to CTM. The posterior distributions of the parameters of interest ($\beta_{1:K}, \theta_{1:D}, z_{1:D,1:N}$) can no longer be obtained using standard simulation techniques such as Gibbs Sampling. Blei and Lafferty (2007) develop a fast variational inference procedure for estimating the CTM. CTM itself has also

been extended by Roberts, Stewart and Airoldi (2016) as part of their work on Structural Topic Models. The main difference is to add a covariate to μ which allows consideration of additional information.

C Events

C.1 List of Elections

Table 5: Elections

year	electionDate	electionWinner
1901	1901-03-29	Non-labor
1903	1903-12-16	Non-labor
1906	1906-12-12	Non-labor
1910	1910-04-13	Labor
1913	1913-05-31	Non-labor
1914	1914-09-05	Labor
1917	1917-05-05	Non-labor
1919	1919-12-13	Non-labor
1922	1922-12-16	Non-labor
1925	1925-11-14	Non-labor
1928	1928-11-17	Non-labor
1929	1929-10-12	Labor
1931	1931-12-19	Non-labor
1934	1934-09-15	Non-labor
1937	1937-10-23	Non-labor
1940	1940-09-21	Non-labor
1943	1943-08-21	Labor
1946	1946-09-28	Labor
1949	1949-12-10	Non-labor
1951	1951-08-28	Non-labor
1954	1954-05-29	Non-labor
1955	1955-12-10	Non-labor
1958	1958-11-22	Non-labor
1961	1961-12-09	Non-labor
1963	1963-11-30	Non-labor
1966	1966-11-26	Non-labor
1969	1969-10-25	Non-labor
1972	1972-12-02	Labor
1974	1974-05-18	Labor
1975	1975-12-13	Non-labor
1977	1977-12-10	Non-labor
1980	1980-10-18	Non-labor
1983	1983-03-05	Labor
1984	1984-12-01	Labor
1987	1987-07-11	Labor
1990	1990-03-24	Labor
1993	1993-03-13	Labor
1996	1996-03-02	Non-labor
1998	1998-10-03	Non-labor
2001	2001-11-10	Non-labor
2004	2004-10-09	Non-labor
2007	2007-11-24	Labor
2010	2010-08-21	Labor
2013	2013-09-07	Non-labor
2016	2016-07-02	Non-labor

Table 6: Change in governments

government	primeMinister	party	start	end	diedInOffice
Barton	Edmund Barton	Protectionist	1901-01-01	1903-09-24	No
Deakin 1	Alfred Deakin	Protectionist	1903-09-24	1904-04-27	No
Watson	Chris Watson	Labour	1904-04-27	1904-08-18	No
Reid	George Reid	Free Trade	1904-08-18	1905-07-05	No
Deakin 2	Alfred Deakin	Protectionist	1905-07-05	1908-11-13	No
Fisher 1	Andrew Fisher	Labour	1908-11-13	1909-06-02	No
Deakin 3	Alfred Deakin	Commonwealth Liberal	1909-06-02	1910-04-29	No
Fisher 2	Andrew Fisher	Labor	1910-04-29	1913-06-24	No
Cook	Joseph Cook	Commonwealth Liberal	1913-06-24	1914-09-17	No
Fisher 3	Andrew Fisher	Labor	1914-09-17	1915-10-27	No
Hughes	Billy Hughes	Labor, National Labor and Nationalist	1915-10-27	1923-02-09	No
Bruce	Stanley Bruce	Nationalist (Coalition)	1923-02-09	1929-10-22	No
Scullin	James Scullin	Labor	1929-10-22	1932-01-06	No
Lyons	Joseph Lyons	United Australia (Coalition)	1932-01-06	1939-04-07	Yes
Page	Earle Page	Country (Coalition)	1939-04-07	1939-04-26	No
Menzies 1	Robert Menzies	United Australia (Coalition)	1939-04-26	1941-08-28	No
Fadden	Arthur Fadden	Country (Coalition)	1941-08-28	1941-10-07	No
Curtin	John Curtin	Labor	1941-10-07	1945-07-05	Yes
Forde	Frank Forde	Labor	1945-07-06	1945-07-13	No
Chifley	Ben Chifley	Labor	1945-07-13	1949-12-19	No
Menzies 2	Robert Menzies	Liberal (Coalition)	1949-12-19	1966-01-26	No
Holt	Harold Holt	Liberal (Coalition)	1966-01-26	1967-12-19	Yes
McEwen	John McEwen	Country (Coalition)	1967-12-19	1968-01-10	No
Gorton	John Gorton	Liberal (Coalition)	1968-01-10	1971-03-10	No
McMahon	William McMahon	Liberal (Coalition)	1971-03-10	1972-12-05	No
Whitlam	Gough Whitlam	Labor	1972-12-05	1975-11-11	No
Fraser	Malcolm Fraser	Liberal (Coalition)	1975-11-11	1983-03-11	No
Hawke	Bob Hawke	Labor	1983-03-11	1991-12-20	No
Keating	Paul Keating	Labor	1991-12-20	1996-03-11	No
Howard	John Howard	Liberal (Coalition)	1996-03-11	2007-12-03	No
Rudd 1	Kevin Rudd	Labor	2007-12-03	2010-06-24	No
Gillard	Julia Gillard	Labor	2010-06-24	2013-06-27	No
Rudd 2	Kevin Rudd	Labor	2013-06-27	2013-09-18	No
Abbott	Tony Abbott	Liberal (Coalition)	2013-09-18	2015-09-15	No
Turnbull	Malcolm Turnbull	Liberal (Coalition)	2015-09-15	2018-08-24	No
Morrison	Scott Morrison	Liberal (Coalition)	2018-08-24	NA	NA

C.2   List of Governments

References

Baumgartner, Frank R. and Bryan D. Jones. 1993. *Agendas and Instability in American Politics*. University of Chicago Press.

Beelen, Kaspar, Timothy Alberdingk Thim, Christopher Cochrane, Kees Halvemaan, Graeme Hirst, Michael Kimmins, Sander Lijbrink, Maarten Marx, Nona Naderi, Ludovic Rheault, Roman Polyanovsky and Tanya Whyte. 2017. "Digitization of the Canadian Parliamentary Debates." *Canadian Journal of Political Science* pp. 1–16.

Benoit, Kenneth. 2018. *quanteda: Quantitative Analysis of Textual Data*. R package version 1.3.4.
URL: *http://quanteda.io*

Blei, David M. 2012. "Probabilistic Topic Models." *Communications of the ACM* 55(4):77–84.

Blei, David M, Andrew Y Ng and Michael I Jordan. 2003. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3(Jan):993–1022.

Blei, David M and John D Lafferty. 2007. "A Correlated Topic Model of Science." *The Annals of Applied Statistics* 1(1):17–35.

Blei, David M and John D Lafferty. 2009. Topic Models. In *Text Mining*. Chapman and Hall/CRC pp. 101–124.

Boulus, Paul. 2013. *Ask Keith for title*. ANU Honours thesis.

Curran, Ben, Kyle Higham, Elisenda Ortiz and Demival Vasques Filho. 2018. "Look Who's Talking: Two-mode Networks As Representations Of A Topic Model Of New Zealand Parliamentary Speeches." *PLOS ONE* 13(6):1–16.

Darling, William M. 2011. A Theoretical and Practical Implementation Tutorial on Topic Modeling and Gibbs Sampling. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. pp. 642–647.

Dimitruk, Kara. 2018. "I Intend Therefore to Prorogue: The Effects of Political Conflict and the Glorious Revolution in Parliament, 1660–1702." *European Review of Economic History* 22(3):261–297.

Dowding, Keith, Andrew Hindmoor, Richard Iles and Peter John. 2010. "Policy Agendas in Australian Politics: The Governor-General's Speeches, 1945–2008." *Australian Journal of Political Science* 45(4):533–557.

Duthie, Rory, Katarzyna Budzynska and Chris Reed. 2016. Mining Ethos in Political Debate. In *Computational Models of Argument*, ed. P Baroni, TF Gordon, T Scheffler and M Stede. Vol. 287 pp. 299–310.

Edwards, Cecilia. 2016. "The Political Consequences of Hansard Editorial Policies: The Case for Greater Transparency." *Australasian Parliamentary Review* 31(2):145–160.

Feinerer, Ingo and Kurt Hornik. 2018. *tm: Text Mining Package.* R package version 0.7-5.
  URL: *https://CRAN.R-project.org/package=tm*

Fraussen, Bert, Timothy Graham and Darren Halpin. forthcoming. "Assessing The Prominence Of Interest Groups In Parliament: A Supervised Machine Learning Approach." *Journal of Legislative Studies* .

Gagolewski, Marek. 2018. *R Package stringi: Character String Processing Facilities.*
  URL: *http://www.gagolewski.com/software/stringi/*

Gans, Joshua and Andrew Leigh. 2012. "How Partisan Is The Press? Multiple Measures Of Media Slant." *The Economic Record* 88(280):127–147.

Gentzkow, Matthew, Jesse M. Shapiro and Matt Taddy. 2018. Measuring Group Differences in High-Dimensional Choices: Method and Application to Congressional Speech. Technical report Voldemort's University.
  URL: *http://web.stanford.edu/ gentzkow/research/politext.pdf*

Graham, Ruth. 2016. Withdraw and Apologise: A Diachronic Study of Unparliamentary Language in the New Zealand Parliament, 1890–1950 PhD thesis.

Griffiths, Thomas and Mark Steyvers. 2004. "Finding Scientific Topics." *PNAS* 101:5228–5235.

Grolemund, Garrett and Hadley Wickham. 2011. "Dates and Times Made Easy with lubridate." *Journal of Statistical Software* 40(3):1–25.
  URL: *http://www.jstatsoft.org/v40/i03/*

Grün, Bettina and Kurt Hornik. 2011. "topicmodels: An R Package for Fitting Topic Models." *Journal of Statistical Software* 40(13):1–30.

Izrailev, Sergei. 2014. *tictoc: Functions for Timing R Scripts.* R package version 1.0.
  URL: *https://CRAN.R-project.org/package=tictoc*

Mimno, David, Edmund Talley, Miriam Leenders, Hanna M Wallach and Andrew McCallum. 2011. Optimizing Semantic Coherence in Topic Models. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*. Edinburgh, Scotland, UK pp. 262–272.

Mollin, Sandra. 2008. "The Hansard hazard: Gauging the Accuracy of British Parliamentary Transcripts." *Corpora* 2(2):187–210.

Ooms, Jeroen. 2017. *hunspell: High-Performance Stemmer, Tokenizer, and Spell Checker*. R package version 2.9.
  URL: *https://CRAN.R-project.org/package=hunspell*

Ooms, Jeroen. 2018. *pdftools: Text Extraction, Rendering and Converting of PDF Documents*. R package version 1.8.
  URL: *https://CRAN.R-project.org/package=pdftools*

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.

Plummer, Martyn. 2018. *rjags: Bayesian Graphical Models using MCMC*. R package version 4-7.
URL: *https://CRAN.R-project.org/package=rjags*

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
URL: *https://www.R-project.org/*

Rasiah, Parameswary. 2010. "A Framework for the Systematic Analysis of Evasion in Parliamentary Discourse." *Journal of Pragmatics* 42:664–680.

Rheault, Ludovic and Christopher Cochrane. 2018. Word Embeddings for the Estimation of Ideological Placement in Parliamentary Corpora. In *PolMeth 2018*. Provo, UT: Society for Political Methodology.

Roberts, Margaret E, Brandon M Stewart and Dustin Tingley. 2016. "Navigating the Local Modes of Big Data." *Computational Social Science* 51.

Roberts, Margaret E., Brandon M Stewart and Dustin Tingley. 2018. *stm: R Package for Structural Topic Models*. R package version 1.3.3.
URL: *http://www.structuraltopicmodel.com*

Roberts, Margaret E., Brandon M Stewart and Edoardo M Airoldi. 2016. "A Model of Text for Experimentation in the Social Sciences." *Journal of the American Statistical Association* 111(515):988–1003.

Salisbury, Christopher. 2011. "'Mr Speaker, I Withdraw...': Standards Of (mis)behaviour In The Queensland, Western Australian And Commonwealth Parliaments Compared Via Online Hansard." *Australasian Parliamentary Review* 26(1):166–177.

Silge, Julia. 2018. *Training, Evaluating, And Interpreting Topic Models*. Last accessed: 2018-11-06.
URL: *https://juliasilge.com/blog/evaluating-stm/*

Silge, Julia and David Robinson. 2016. "tidytext: Text Mining and Analysis Using Tidy Data Principles in R." *JOSS* 1(3).
URL: *http://dx.doi.org/10.21105/joss.00037*

Slapin, Jonathan B, Justin H Kirkland, Joseph A Lazzaro, Patrick A Leslie and Tom O'Grady. 2018. "Ideology, Grandstanding, and Strategic Party Disloyalty in the British Parliament." *American Political Science Review* 112(1):15–30.

Steyvers, Mark and Tom Griffiths. 2006. Probabilistic Topic Models. In *Latent Semantic Analysis: A Road to Meaning*, ed. T. Landauer, D McNamara, S. Dennis and W. Kintsch.

Taddy, Matt. 2015. "Distributed Multinomial Regression." *The Annals of Applied Statistics* 9(3):1394–1414.

Taddy, Matthew. 2012. On Estimation and Selection for Topic Models. In *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics (AISTATS)*. La Palma, Canary Islands pp. 1184–1193.

Vaughan, Davis and Matt Dancho. 2018. *furrr: Apply Mapping Functions in Parallel using Futures*. R package version 0.1.0.9002.
URL: *https://github.com/DavisVaughan/furrr*

Wallach, Hanna M, Iain Murray, Ruslan Salakhutdinov and David Mimno. 2009. Evaluation Methods for Topic Models. In *Proceedings of the 26th International Conference on Machine Learning*. Montreal, Canada.

Wickham, Hadley. 2017. *tidyverse: Easily Install and Load the 'Tidyverse'*. R package version 1.2.1.
URL: *https://CRAN.R-project.org/package=tidyverse*