

# Climate Congressional Record and Climate Hansard Topic Modelling

This R markdown document contains the code to create a topic model of the Climate Congressional Record and Climate Hansard datasets. A topic model is used to investigate the following hypotheses:

- H2a: Between 1997 and 2015, there was a stronger association between party affiliation and pro- or anti-environmental topic prevalence among US legislators climate policy speeches than among UK legislators' climate policy speeches.
- H2b: Between 1997 and 2015, the association between party affiliation and pro- or anti-environmental topic prevalence increased among US legislators' climate policy speeches, while it reduced among UK legislators' climate policy speeches.
- H4: Between 1997 and 2015, the words used in US legislators' climate policy speeches exhibited more variation between elections and non-electoral periods than those used in UK legislators' climate policy speeches.
- H5: Between 1997 and 2015, the association between the topics discussed in US legislators' climate policy speeches and American climate change opinion was stronger than the association between the topics discussed in UK legislators' climate policy speeches and British climate change opinion.

## Setup

```
library(tidyverse)
library(stringr)
library(quanteda)
library(stm)

DATA_PATH <- "data/"
DIST_PATH <- "dist/"

NUM_TOPICS <- 18
```

```
climate_congressional_record <-  
  read_csv(paste0(DATA_PATH, "climate_congressional_record.csv"))  
climate_hansard <-  
  read_csv(paste0(DATA_PATH, "climate_hansard.csv"))
```

## Data Preprocessing

### Cleaning the Climate Congressional Record dataset

```
partitioned_climate_congressional_record <- climate_congressional_record |>  
  rename(  
    first_name = first_name_x,  
    family_name = last_name_x,  
    speech_date = date,  
    speech_party = party,  
    text = speech  
  ) |>  
  select(  
    speech_id,  
    first_name,  
    family_name,  
    speech_date,  
    speech_party,  
    text,  
    cleaned_stems  
  ) |>  
  filter(speech_party == "R" | speech_party == "D") |>  
  mutate(  
    name = paste(first_name, family_name, sep = " ") |>  
      str_to_title(),  
    first_name = str_to_title(first_name),  
    family_name = str_to_title(family_name),  
    speech_party = if_else(speech_party == "R", "Republican", "Democrat"),  
    speech_id = as.character(speech_id),  
    country = "USA"  
  )
```

## Cleaning the Climate Hansard dataset

```
climate_hansard_party_filter <-  
  c("Conservative", "Labour", "Liberal Democrat", "Labour/Co-operative")  
partitioned_climate_hansard <- climate_hansard |>  
  select(  
    speech_id,  
    name,  
    first_name,  
    family_name,  
    speech_date,  
    speech_party,  
    text,  
    cleaned_stems  
  ) |>  
  filter(speech_party %in% climate_hansard_party_filter) |>  
  mutate(  
    speech_party =  
      if_else(  
        speech_party == "Labour/Co-operative", "Labour", speech_party  
      ),  
    country = "UK"  
  )
```

## Combining the Climate Congressional Record and Climate Hansard datasets

```
complete_climate_speeches <- bind_rows(  
  partitioned_climate_congressional_record,  
  partitioned_climate_hansard  
) |>  
  arrange(speech_date) |>  
  mutate(  
    date_as_numeric = as.numeric(speech_date)  
  )
```

## Creating a Climate Speeches DFM

```

climate_speeches_corpus <-
  corpus(
    complete_climate_speeches,
    text_field = "cleaned_stems"
  )
climate_speeches_dfm <-
  climate_speeches_corpus |>
  tokens() |>
  dfm()

```

## Testing Hypothesis 2a

Fitting a structural topic model with an interaction between the `speech_party` covariate and `date_as_numeric`

```

# party_interaction_stm <- stm(
#   documents = climate_speeches_dfm,
#   prevalence = ~ speech_party * s(date_as_numeric),
#   K = NUM_TOPICS,
#   seed = 42
# )
# save(party_interaction_stm, file = paste0(
#   DATA_PATH,
#   "party_interaction_stm.RData"
# ))
load(paste0(DATA_PATH, "party_interaction_stm.RData"))

```

## Investigating the model's topics

```
labelTopics(party_interaction_stm)
```

```

## Topic 1 Top Words:
##   Highest Prob: paper, green, shall, conserv, white, remark, answer
##   FREX: chingford, woodford, duncan, hour, paper, ashford, damian
##   Lift: aberconwi, aesthet, amendmentsinterrupt, areassuch, bagot, barbado, bebb
##   Score: paper, shall, green, hour, white, sit, chingford
## Topic 2 Top Words:
##   Highest Prob: green, plan, invest, deal, job, protect, new

```

```

##      FREX: belt, brownfield, centr, protect, petit, greenbelt, stretford
##      Lift: bungalow, cirl, communityl, greentech, highgrad, hounslow, humanist
##      Score: green, invest, belt, job, protect, plan, bank
## Topic 3 Top Words:
##      Highest Prob: chang, framework, report, set, respons, consult, public
##      FREX: regul, publish, document, consult, draft, detail, bodi
##      Lift: accountsnarr, agencyand, alliancegeorg, alludedentir, andif, appropriatel
##      Score: report, publish, chang, framework, regul, set, consult
## Topic 4 Top Words:
##      Highest Prob: global, climat, warm, us, nation, make, issu
##      FREX: leader, dont, that, said, good, health, agenda
##      Lift: ab, aboutthi, absolv, achiei, adida, admini, affelt
##      Score: nation, make, know, way, import, polici, epa
## Topic 5 Top Words:
##      Highest Prob: climat, atmospher, ocean, water, warm, temperatur, increas
##      FREX: speci, fish, forest, sea, river, wildlif, habitat
##      Lift: abalon, accordg, acidcould, actionsth, actlaw, adaptationto, aeon
##      Score: ocean, speci, atmospher, forest, water, ice, temperatur
## Topic 6 Top Words:
##      Highest Prob: carbon, emiss, target, per, reduc, cent, tell
##      FREX: transport, tonn, reduct, cent, tell, cc, reduc
##      Lift: tyndal, zealot, achil, agreean, airsourc, airyfairi, alstom
##      Score: carbon, emiss, reduc, cent, per, target, reduct
## Topic 7 Top Words:
##      Highest Prob: climat, chang, countri, issu, world, need, intern
##      FREX: adapt, pari, g, confer, countri, agreement, leadership
##      Lift: actionaid, africainclud, agow, aidthat, alarmiststhey, allpeacekeep, angl
##      Score: climat, chang, countri, agreement, world, unit, intern
## Topic 8 Top Words:
##      Highest Prob: net, author, billion, hear, figur, million, meet
##      FREX: net, sd, figur, room, pm, borrow, migrat
##      Lift: outturn, abc, abeut, aboutexpenditur, aboveaverag, accel, accent
##      Score: net, figur, billion, consent, hear, author, migrat
## Topic 9 Top Words:
##      Highest Prob: zero, polic, peopl, warn, money, women, true
##      FREX: crime, violenc, teacher, prison, women, victim, zero
##      Lift: aberfan, acrylamid, annumth, antibulli, arab, arei, ariel
##      Score: zero, polic, true, women, crime, prison, young
## Topic 10 Top Words:
##      Highest Prob: chang, time, matter, propos, whether, legisl, case
##      FREX: vote, law, elector, elect, legisl, rule, mind
##      Lift: absolutelybecaus, actit, adeptli, ado, adumbr, agreeand, alfr
##      Score: chang, vote, law, legisl, matter, time, elect
## Topic 11 Top Words:
##      Highest Prob: tax, use, chang, power, industri, cost, levi

```

```

##      FREX: power, exempt, taxat, levi, use, revenu, wast
##      Lift: neta, sledgehamm, wda, ableeven, agricultureand, agrochem, answerwhil
##      Score: tax, power, use, levi, chang, cost, industri
## Topic 12 Top Words:
##      Highest Prob: year, work, last, next, week, done, past
##      FREX: aid, past, year, five, week, six, tribut
##      Lift: adolesc, agn, agoni, aheadand, alreadybut, antiretrovir, archeri
##      Score: year, work, week, next, past, last, month
## Topic 13 Top Words:
##      Highest Prob: european, kyoto, protocol, treati, implement, provis, negoti
##      FREX: protocol, treati, ratifi, ratif, lisbon, articl, kyoto
##      Lift: abm, absente, acknowledgedpresum, advocatedthat, afar, afarin, affidavit
##      Score: treati, protocol, kyoto, european, negoti, languag, air
## Topic 14 Top Words:
##      Highest Prob: get, want, come, peopl, us, talk, back
##      FREX: bit, might, bencher, get, got, back, everybodi
##      Lift: collud, reliv, snap, sweater, velcro, aaah, abellio
##      Score: get, back, might, want, come, talk, tri
## Topic 15 Top Words:
##      Highest Prob: fund, provid, move, research, develop, program, purpos
##      FREX: ireland, purpos, northern, research, program, corpor, weapon
##      Lift: abhorr, accuweath, angelo, anni, antiflash, appalachian, asbestosthey
##      Score: program, research, fund, nation, develop, ireland, move
## Topic 16 Top Words:
##      Highest Prob: energi, fuel, renew, effici, technolog, carbon, fossil
##      FREX: coal, biomass, biofuel, electr, sourc, effici, renew
##      Lift: chevron, aboard, aclean, advancemehit, amencan, andutil, andw
##      Score: energi, fuel, renew, fossil, coal, effici, clean
## Topic 17 Top Words:
##      Highest Prob: fuel, peopl, pension, benefit, home, poverti, help
##      FREX: pension, payment, p, liter, poverti, vat, escal
##      Lift: aboutw, achievepens, advisersi, afteri, agesubstanti, allowanceit, allowa
##      Score: fuel, pension, poverti, payment, winter, price, household
## Topic 18 Top Words:
##      Highest Prob: global, econom, achiev, market, much, economi, competit
##      FREX: resourc, pressur, context, market, bear, achiev, scale
##      Lift: spillag, stearn, birdswhich, bornsuch, businesssit, flourishunlik, imd
##      Score: econom, achiev, global, market, much, resourc, competit

```

```

findThoughts(party_interaction_stm,
  texts = climate_speeches_dfm$text,
  n = 2
)

```

```
##
```

## Topic 1:  
 ## I rise to speak primarily in support of my amendment No. 46, but also in support  
 ## I rise to speak primarily in support of my amendment No. 46, but also in support  
 ## Topic 2:  
 ## Ormiston Wire in Isleworth has previously won the Queen's award for sustainable  
 ## One of the monstrosities that have afflicted our green and pleasant land is the  
 ## Topic 3:  
 ## On 31 March, we published "e-government"-a strategic framework for public servi  
 ##  
 ## out their e-business strategies by October 2000, with the e-envoy reporting to my rig  
 ## Does the Minister agree with the recent report by the all-party group on cancer,  
 ## Topic 4:  
 ## Mr. Speaker. many of my colleagues on the other side of the aisle have continua  
 ## Mr. President. as chairman of the Committee on Environment and Public Works. I h  
 ## Topic 5:  
 ## mentioned earlier. the temperature of the planet was about 5 degrees centigrade  
 ## Mr. Speaker. I rise today to bring attention to a recent University of Californi  
 ## Topic 6:  
 ## If he will make a statement on the Government's strategy for meeting the target  
 ## What recent assessment he has made of progress towards meeting the Government's  
 ## Topic 7:  
 ## I regularly discuss the sustainable development goals with my international cou  
 ## We know that the greatest risk that the UK faces from climate change is flooding  
 ## Topic 8:  
 ## Mr. President. this is the Statement of Budgetary Effects of PAYGO Legislation  
 ## Mr. President. this is the Statement of Budgetary Effects of PAYGO Legislation f  
 ## Topic 9:  
 ## I am sure my hon. Friend would agree that there should be zero tolerance of abu  
 ## Yes. and you can have a 10 for definitely true. 7 for mostly true. 5 for undecid  
 ## Topic 10:  
 ## I note that the right hon. Gentleman has changed his tack. The question is whet  
 ## As the Minister said, these are technical changes relating to arrangements and r  
 ## Topic 11:  
 ## Does the hon. Gentleman accept that that damaging trend is likely to be acceler  
 ## If the Minister believes-I am sure that he is right-that the decommissioning on  
 ## Topic 12:  
 ## My right hon. Friend is aware of the work of Derian House children's hospice in  
 ## My hon. Friend might like to know that, over the 21 years that I worked at Cawsa  
 ## Topic 13:  
 ## Mr. President. I ask unanimous consent that Kurt Volker. a legislative fellow i  
 ## If the Senator will yield for just a second. those demarcation protocols to the  
 ## Topic 14:  
 ## I thank this friend of West Virginia. this friend of mine. I shall always remem  
 ## It might surprise my hon. Friend to know that I do talk to many lesbians, and qu  
 ## Topic 15:

```

##      Mr. Speaker. I move to suspend the rules and pass the Senate bill for the relie
##      Mr. President. I ask the Chair lay before the Senate a message from the House of
## Topic 16:
##      Madam Speaker. today we will consider legislation that invests in the future of
##      Mr. Chaitman. I yil .mydelf 2. minutes. Mr. Chairman. this amendment would clari
## Topic 17:
##      The truth is that millions of low-income families are still living in poorly in
##      Does the right hon. Gentleman agree that the difference between this and earlier
## Topic 18:
##      My hon. Friend is absolutely right, but the context for the problem is the new
##      I am slightly baffled by this. The argument for why, in 2005, on these figures,

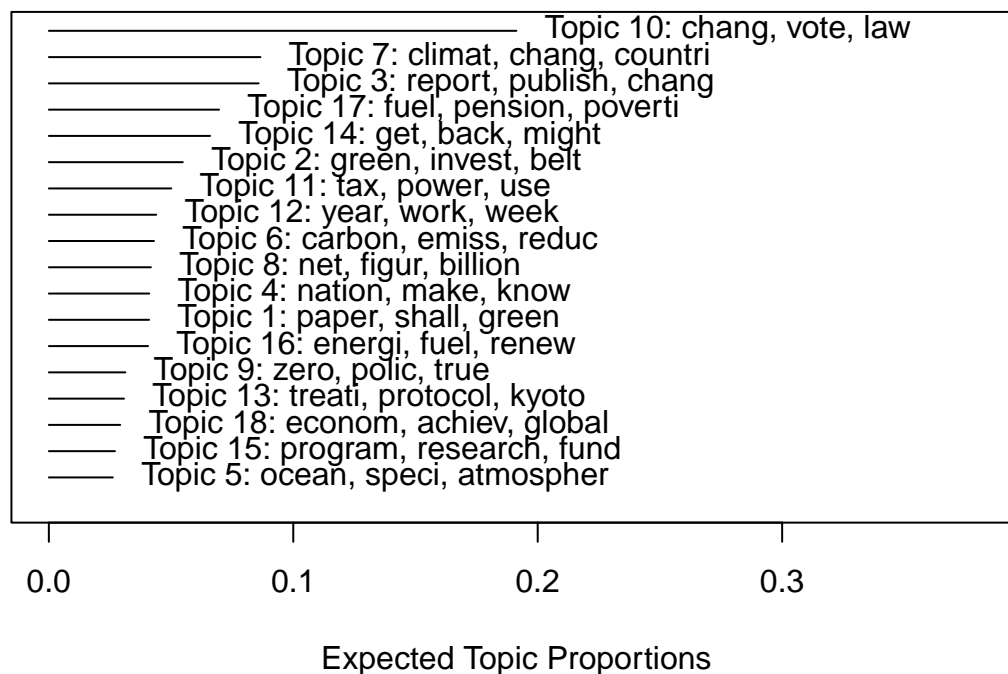
```

```

plot.STM(party_interaction_stm,
  type = "summary",
  labeltype = "score"
)

```

## Top Topics



Visualising the model's topic, covariate relationships



```

party_prevalence_effects <- estimateEffect(
  formula =
    c(seq_len(NUM_TOPICS)) ~ speech_party * s(date_as_numeric),
  stmobj = party_interaction_stm,
  metadata = docvars(climate_speeches_dfm),
  uncertainty = "None"
)
summary(party_prevalence_effects, topic = 5)

```

```

##
## Call:
## estimateEffect(formula = c(seq_len(NUM_TOPICS)) ~ speech_party *
##   s(date_as_numeric), stmobj = party_interaction_stm, metadata = docvars(climate_sp
##   uncertainty = "None")
##
##
## Topic 5:
##
## Coefficients:
##

```

	Estimate	Std. Error	t value
## (Intercept)	0.0090606	0.0119967	0.755
## speech_partyDemocrat	0.1163877	0.0221417	5.257
## speech_partyLabour	0.0066514	0.0150302	0.443
## speech_partyLiberal Democrat	-0.0156786	0.0266348	-0.589
## speech_partyRepublican	0.0900370	0.0182666	4.929
## s(date_as_numeric)1	0.0166432	0.0202575	0.822
## s(date_as_numeric)2	0.0016058	0.0116411	0.138
## s(date_as_numeric)3	0.0216624	0.0157155	1.378
## s(date_as_numeric)4	0.0022621	0.0121722	0.186
## s(date_as_numeric)5	0.0184864	0.0134990	1.369
## s(date_as_numeric)6	0.0004597	0.0128661	0.036
## s(date_as_numeric)7	0.0008419	0.0126981	0.066
## s(date_as_numeric)8	0.0111917	0.0136590	0.819
## s(date_as_numeric)9	-0.0115136	0.0142723	-0.807
## s(date_as_numeric)10	0.0193991	0.0125254	1.549
## speech_partyDemocrat:s(date_as_numeric)1	-0.0838184	0.0417871	-2.006
## speech_partyLabour:s(date_as_numeric)1	-0.0035787	0.0257669	-0.139
## speech_partyLiberal Democrat:s(date_as_numeric)1	0.0291427	0.0469134	0.621
## speech_partyRepublican:s(date_as_numeric)1	-0.0808731	0.0351883	-2.298
## speech_partyDemocrat:s(date_as_numeric)2	0.0624556	0.0232135	2.690
## speech_partyLabour:s(date_as_numeric)2	-0.0166667	0.0140058	-1.190
## speech_partyLiberal Democrat:s(date_as_numeric)2	0.0080236	0.0246587	0.325
## speech_partyRepublican:s(date_as_numeric)2	0.0724713	0.0205862	3.520
## speech_partyDemocrat:s(date_as_numeric)3	-0.0482281	0.0304517	-1.584

## speech_partyLabour:s(date_as_numeric)3	-0.0118909	0.0197128	-0.603
## speech_partyLiberal Democrat:s(date_as_numeric)3	0.0075384	0.0341626	0.221
## speech_partyRepublican:s(date_as_numeric)3	-0.0022474	0.0277756	-0.081
## speech_partyDemocrat:s(date_as_numeric)4	-0.0252780	0.0241564	-1.046
## speech_partyLabour:s(date_as_numeric)4	-0.0040424	0.0149566	-0.270
## speech_partyLiberal Democrat:s(date_as_numeric)4	0.0262999	0.0265373	0.991
## speech_partyRepublican:s(date_as_numeric)4	0.0769598	0.0213477	3.605
## speech_partyDemocrat:s(date_as_numeric)5	-0.0609472	0.0243449	-2.503
## speech_partyLabour:s(date_as_numeric)5	-0.0137125	0.0172432	-0.795
## speech_partyLiberal Democrat:s(date_as_numeric)5	0.0082285	0.0295108	0.279
## speech_partyRepublican:s(date_as_numeric)5	-0.1235304	0.0227458	-5.431
## speech_partyDemocrat:s(date_as_numeric)6	-0.0272951	0.0263687	-1.035
## speech_partyLabour:s(date_as_numeric)6	-0.0022031	0.0161546	-0.136
## speech_partyLiberal Democrat:s(date_as_numeric)6	0.0121628	0.0278077	0.437
## speech_partyRepublican:s(date_as_numeric)6	0.0785587	0.0281655	2.789
## speech_partyDemocrat:s(date_as_numeric)7	0.0034073	0.0266403	0.128
## speech_partyLabour:s(date_as_numeric)7	-0.0150150	0.0163999	-0.916
## speech_partyLiberal Democrat:s(date_as_numeric)7	0.0113055	0.0288973	0.391
## speech_partyRepublican:s(date_as_numeric)7	-0.1522301	0.0299971	-5.075
## speech_partyDemocrat:s(date_as_numeric)8	0.1908436	0.0287698	6.633
## speech_partyLabour:s(date_as_numeric)8	-0.0068959	0.0184295	-0.374
## speech_partyLiberal Democrat:s(date_as_numeric)8	0.0034725	0.0296802	0.117
## speech_partyRepublican:s(date_as_numeric)8	0.0680798	0.0352750	1.930
## speech_partyDemocrat:s(date_as_numeric)9	0.0030167	0.0307140	0.098
## speech_partyLabour:s(date_as_numeric)9	0.0006521	0.0182937	0.036
## speech_partyLiberal Democrat:s(date_as_numeric)9	0.0261305	0.0324368	0.806
## speech_partyRepublican:s(date_as_numeric)9	-0.0351055	0.0304138	-1.154
## speech_partyDemocrat:s(date_as_numeric)10	-0.0347286	0.0252191	-1.377
## speech_partyLabour:s(date_as_numeric)10	-0.0193774	0.0166524	-1.164
## speech_partyLiberal Democrat:s(date_as_numeric)10	-0.0088053	0.0359476	-0.245
## speech_partyRepublican:s(date_as_numeric)10	-0.0033850	0.0255316	-0.133
##	Pr(> t )		
## (Intercept)	0.450100		
## speech_partyDemocrat	1.48e-07	***	
## speech_partyLabour	0.658105		
## speech_partyLiberal Democrat	0.556101		
## speech_partyRepublican	8.31e-07	***	
## s(date_as_numeric)1	0.411321		
## s(date_as_numeric)2	0.890290		
## s(date_as_numeric)3	0.168089		
## s(date_as_numeric)4	0.852569		
## s(date_as_numeric)5	0.170865		
## s(date_as_numeric)6	0.971496		
## s(date_as_numeric)7	0.947135		
## s(date_as_numeric)8	0.412588		

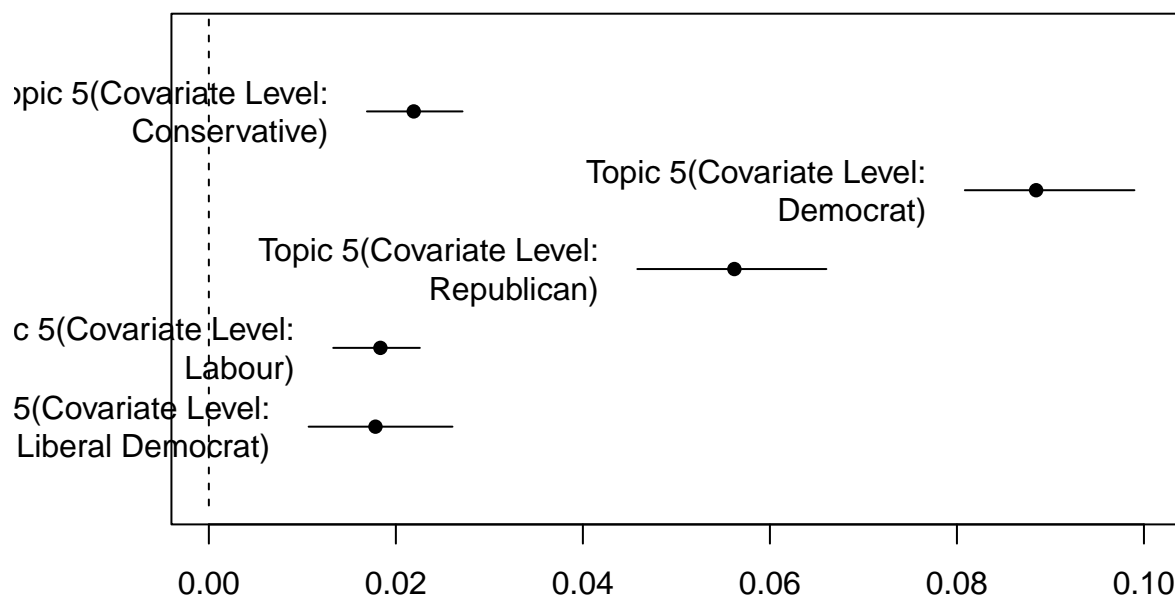
```

## s(date_as_numeric)9 0.419841
## s(date_as_numeric)10 0.121447
## speech_partyDemocrat:s(date_as_numeric)1 0.044883 *
## speech_partyLabour:s(date_as_numeric)1 0.889540
## speech_partyLiberal Democrat:s(date_as_numeric)1 0.534472
## speech_partyRepublican:s(date_as_numeric)1 0.021553 *
## speech_partyDemocrat:s(date_as_numeric)2 0.007139 **
## speech_partyLabour:s(date_as_numeric)2 0.234062
## speech_partyLiberal Democrat:s(date_as_numeric)2 0.744892
## speech_partyRepublican:s(date_as_numeric)2 0.000432 ***
## speech_partyDemocrat:s(date_as_numeric)3 0.113261
## speech_partyLabour:s(date_as_numeric)3 0.546377
## speech_partyLiberal Democrat:s(date_as_numeric)3 0.825356
## speech_partyRepublican:s(date_as_numeric)3 0.935511
## speech_partyDemocrat:s(date_as_numeric)4 0.295371
## speech_partyLabour:s(date_as_numeric)4 0.786953
## speech_partyLiberal Democrat:s(date_as_numeric)4 0.321667
## speech_partyRepublican:s(date_as_numeric)4 0.000313 ***
## speech_partyDemocrat:s(date_as_numeric)5 0.012303 *
## speech_partyLabour:s(date_as_numeric)5 0.426480
## speech_partyLiberal Democrat:s(date_as_numeric)5 0.780377
## speech_partyRepublican:s(date_as_numeric)5 5.66e-08 ***
## speech_partyDemocrat:s(date_as_numeric)6 0.300616
## speech_partyLabour:s(date_as_numeric)6 0.891526
## speech_partyLiberal Democrat:s(date_as_numeric)6 0.661831
## speech_partyRepublican:s(date_as_numeric)6 0.005288 **
## speech_partyDemocrat:s(date_as_numeric)7 0.898230
## speech_partyLabour:s(date_as_numeric)7 0.359910
## speech_partyLiberal Democrat:s(date_as_numeric)7 0.695631
## speech_partyRepublican:s(date_as_numeric)7 3.90e-07 ***
## speech_partyDemocrat:s(date_as_numeric)8 3.34e-11 ***
## speech_partyLabour:s(date_as_numeric)8 0.708275
## speech_partyLiberal Democrat:s(date_as_numeric)8 0.906863
## speech_partyRepublican:s(date_as_numeric)8 0.053621 .
## speech_partyDemocrat:s(date_as_numeric)9 0.921760
## speech_partyLabour:s(date_as_numeric)9 0.971563
## speech_partyLiberal Democrat:s(date_as_numeric)9 0.420491
## speech_partyRepublican:s(date_as_numeric)9 0.248404
## speech_partyDemocrat:s(date_as_numeric)10 0.168500
## speech_partyLabour:s(date_as_numeric)10 0.244578
## speech_partyLiberal Democrat:s(date_as_numeric)10 0.806498
## speech_partyRepublican:s(date_as_numeric)10 0.894525
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```
plot.estimateEffect(party_prevalence_effects,
  type = "labels",
  covariate = "speech_party",
  topics = 5,
  method = "pointestimate",
  xlim = c(0, 0.1)
)
```

```
## Warning in plot.xy(xy, type, ...): plot type 'labels' will be truncated to
## first character
```



**Testing Hypothesis 2b**

**Testing Hypothesis 4**