STM Topicmodel

Riccardo Ruta

05/2022

Contents

STM Topic model analysis

Preliminary steps

Load the data

```
load("data/dfm.Rda")
load("data/dataset.Rda")
load("data/tw.Rda")
load("data/corpus.Rda")
```

Import the dictionaries

```
# Import dictionaries file
dict <-  read_excel("data/populism_dictionaries.xlsx")

Decadri_Boussalis_Grundl <-
  dictionary(list(people =
                    dict$Decadri_Boussalis_Grundl_People
                  [!is.na(dict$Decadri_Boussalis_Grundl_People)],
                  common_will =
                    dict$`Decadri_Boussalis_Grundl_Common Will`
                  [!is.na(dict$`Decadri_Boussalis_Grundl_Common Will`)],
```

```
                  elite =
                    dict$Decadri_Boussalis_Grundl_Elite
                  [!is.na(dict$Decadri_Boussalis_Grundl_Elite)]))
```

Remove all the account's mentions

```
DFM@Dimnames$features <- gsub("^@", "", DFM@Dimnames$features)
```

Trim the data

```
# Remove text with less than 1 word
DFM <- dfm_subset(DFM, ntoken(DFM) > 1)

# Remove very short words
DFM <- dfm_remove(DFM, min_nchar=2)

# Dfm trimming:only words that occur in the top 20% of the distribution
# and in less than 10% of documents
DFM <- dfm_trim(DFM,min_termfreq = 0.8,
                termfreq_type = "quantile",
                max_docfreq = 0.1, docfreq_type = "prop")
```

Apply dictionary

```
# Apply Dictionary
DFMdict <- dfm_lookup(DFM, dictionary = Decadri_Boussalis_Grundl)

# Convert to a dataframe
DATAdictDFM <- DFMdict %>%
  quanteda::convert(to = "data.frame")
```

Create percentage for each components

```
# RUN ONLY ONCE!
# Add variable with general level of populism & multiply all components by 100
DATAdictDFM <- DATAdictDFM %>%
  mutate(populism = (people + common_will + elite) * 100)

DATAdictDFM <- DATAdictDFM %>%
  mutate(people = people*100,
         common_will = common_will*100,
         elite = elite*100)
```

Add the percentage of populism to the original dfm

```
docvars(DFM) <- cbind(docvars(DFM),DATAdictDFM)
```

Convert DFM to STM format

```
myDFM = DFM
set.seed(123)
DfmStm <- quanteda::convert(myDFM, to = "stm", docvars = docvars(myDFM))
#save(DfmStm, file="data/DfmStm.Rda")
```

Import the original corpus and repeat the same cleanings

This is for search the documents after find a label for the topics

```
# Remove text with less than 1 word
corpus <- corpus_subset(corpus, ntoken(corpus) > 1)

subs_corpus <- corpus_subset(corpus,
                             docnames(corpus) %in% myDFM@Dimnames$docs)

subs_corpus <- as.character(subs_corpus)

subs_corpus <-  as.vector(subs_corpus)

#save(subs_corpus, file="data/subs_corpus.Rda")
```

Find best number of topics k

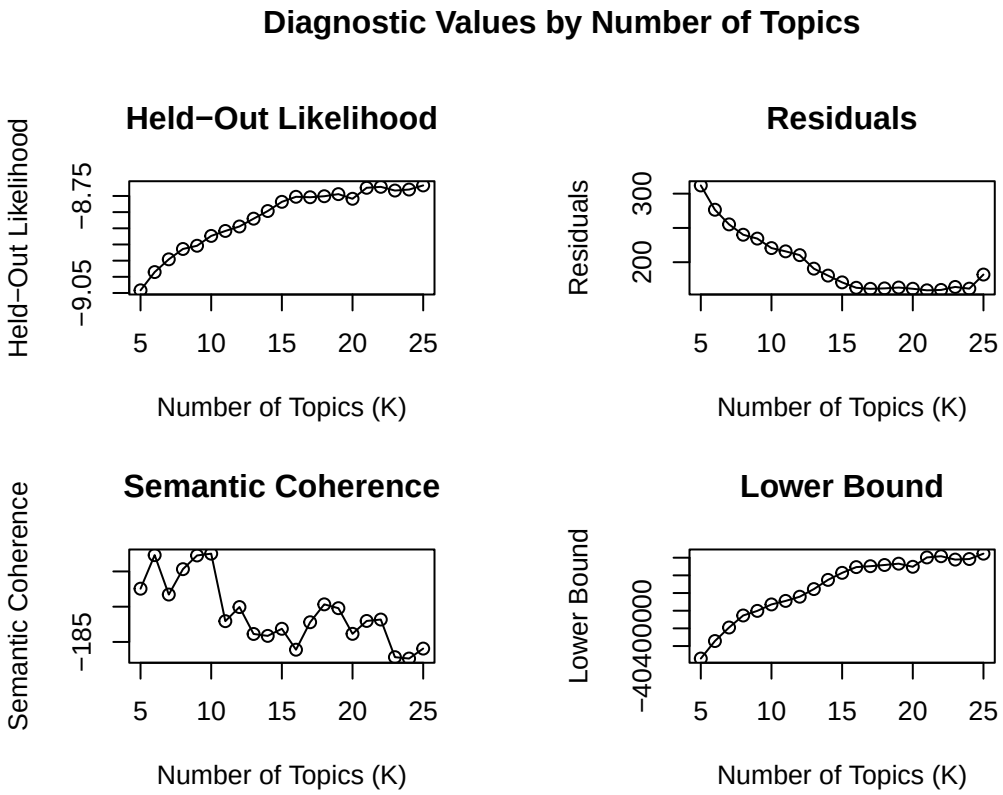
Search the best number of Topics comparing coherence and exclusivity values

K = 5:20

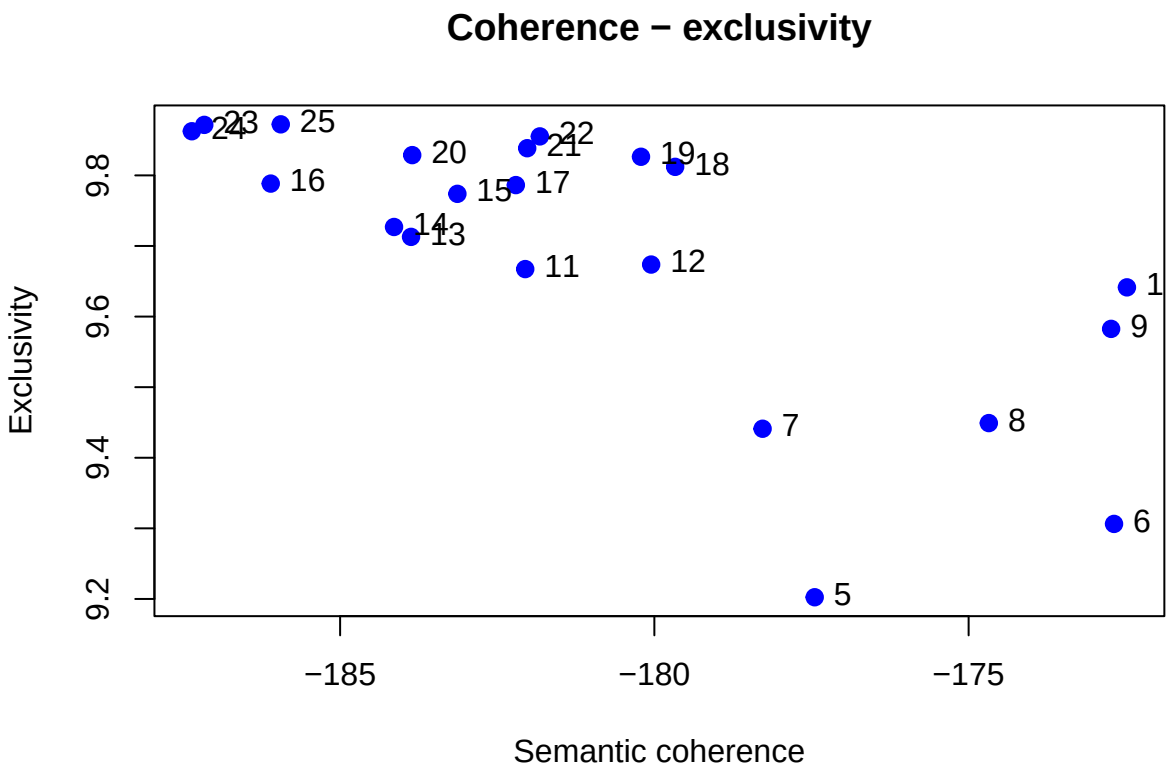
```
k <-c(5:20)
system.time(storageNoG <- searchK(DfmStm$documents,
                                  DfmStm$vocab,
                                  K = k,
                                  prevalence = ~ party_id + populism + s(quarter),
                                  data = DfmStm$meta, init.type = "Spectral"))
#save(storageNoG,file="data/storageNoG.Rda")
```

plot results

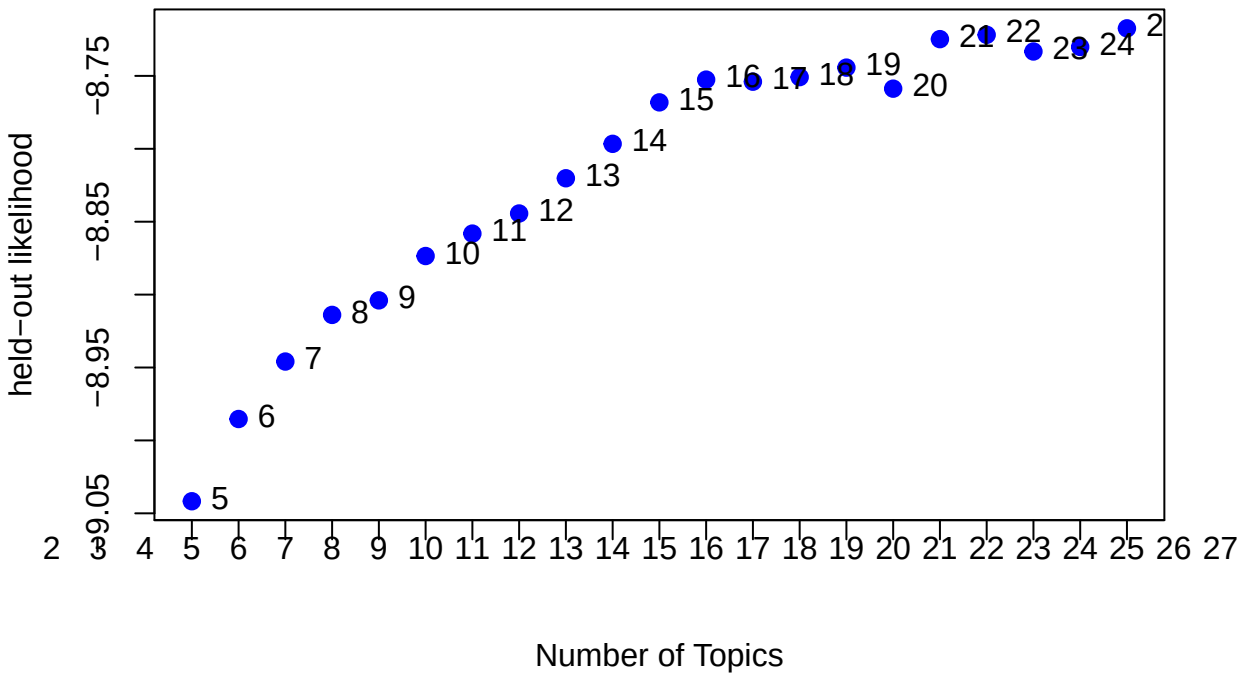
```
plot.searchK(storageNoG)
```




```
plot(storageNoG$results$semcoh, storageNoG$results$exclus,
     xlab= "Semantic coherence",
     ylab= "Exclusivity",
     col= "blue", pch = 19, cex = 1, lty = "solid", lwd = 2, main = "Coherence - exclusivity")
text(storageNoG$results$semcoh, storageNoG$results$exclus,
     labels=storageNoG$results$K, cex= 1, pos=4)
```




```
plot(storageNoG$results$K, storageNoG$results$heldout,
     xlab= "Number of Topics",
     ylab= "held-out likelihood",
     col= "blue", pch = 19, cex = 1, lty = "solid", lwd = 2, xaxt="n")
text(storageNoG$results$K, storageNoG$results$heldout,
     labels=storageNoG$results$K, cex= 1, pos=4)
xtick<-seq(2, 50, by=1)
axis(side=1, at=xtick, labels = FALSE)
text(x=xtick,  par("usr")[3],
     labels = xtick, pos = 1, xpd = TRUE)
```


held−out likelihood
−8.75
−8.85
−8.95
−9.05
Number of Topics
2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

Run the analysis selecting k = 10

```
k = 10

mySTM10NoG <- stm(DfmStm$documents, vocab = DfmStm$vocab,
              K = k,
              prevalence = ~ party_id + populism + s(quarter),
              data = DfmStm$meta,
              init.type = "Spectral",
              verbose = TRUE)

 #save(mySTM10NoG,file="data/mySTM10NoG.Rda")
```

Label topics

The frequency/exclusivity (FREX) scoring summarizes words according to their probability of appearance under a topic and the exclusivity to that topic. These words provide more semantically intuitive representations of each topic.

```
labeledtpic <- labelTopics(mySTM10NoG, n=10)

#FREX
FREXmySTM10NoG <- t(as.matrix(labeledtpic[["frex"]]))
kable(FREXmySTM10NoG[,1:5], col.names = c("Topic1","Topic2","Topic3",
                                          "Topic4","Topic5"))
```

Topic1	Topic2	Topic3	Topic4	Topic5
regione	pass	legasalvini	libertà	domani
covid	ragazzi	vittime	diritti	città
vaccini	green	famiglia	democrazia	buona
personale	viva	pensiero	violenza	auguri
virus	#greenpass	foto	minacce	sindaco
medici	vittoria	ricordo	inaccettabile	mattina
sanitario	sport	memoria	costituzione	aspetto
numero	vince	abbraccio	#ddlzan	milano
lombardia	finale	pagina	umani	vediamo
coronavirus	vinto	legacamera	esteri	stasera

```
kable(FREXmySTM10NoG[,6:10], col.names = c("Topic6","Topic7","Topic8",
                                           "Topic9","Topic10"))
```

Topic6	Topic7	Topic8	Topic9	Topic10
fratellilditalia	presidente	conte	italiaviva	imprese
stampasgarbi	forza_italia	salvini	davvero	euro
giorgiameloni	pdnetwork	pd	giusto	lavoratori
#fratellilditalia	giuseppeconteit	#salvini	cambiare	decreto
meloni	deputatipd	#conte	matteorenzi	misure
vocedelpatriota	berlusconi	m5s	guardare	piano
fratelli	#mattarella	vergogna	maestro	risorse
#meloni	enricoletta	#iostoconsalvini	l'intervista	miliardi
ilgiornale	mattarella	vogliono	parità	bilancio
adnkronos	gruppoficamera	referendum	elenabonetti	servono

Looking at the FREX words, it is complicated to give a substantive interpretation of the content of the topics. We therefore made a second attempt using k = 18

Run the analysis selecting k = 18

```
k = 18

mySTM18NoG <- stm(DfmStm$documents, vocab = DfmStm$vocab,
             K = k,
             prevalence = ~ party_id + populism + s(quarter),
             data = DfmStm$meta,
             init.type = "Spectral",
             verbose = TRUE)

# save(mySTM18NoG,file="data/mySTM18NoG.Rda")
```

Label topics

The frequency/exclusivity (FREX) scoring summarizes words according to their probability of appearance under a topic and the exclusivity to that topic. These words provide more semantically intuitive representations of each topic.

```
labeledtpic <- labelTopics(mySTM18NoG, n=30)

#FREX
FREXmySTM18NoG <- t(as.matrix(labeledtpic[["frex"]]))
kable(FREXmySTM18NoG[,1:5], col.names = c("Topic1","Topic2","Topic3",
                                          "Topic4","Topic5")) %>%
  kable_styling(latex_options="scale_down")
```

```
kable(FREXmySTM18NoG[,6:10], col.names = c("Topic6","Topic7","Topic8",
                                           "Topic9","Topic10")) %>%
  kable_styling(latex_options="scale_down")
```

```
kable(FREXmySTM18NoG[,11:15], col.names = c("Topic11","Topic12","Topic13",
                                            "Topic14","Topic15")) %>%
  kable_styling(latex_options="scale_down")
```

```
kable(FREXmySTM18NoG[,16:18], col.names = c("Topic16","Topic17","Topic18"))
```

Topic1	Topic2	Topic3	Topic4	Topic5
great	dosi	pubblicata	maestro	anzaldi
muto	dose	foto	mancherà	lastampa
will	molecolari	#venetodaamare	addio	intervista
red_marxist	tamponi	vicenza	mancherai	adnkronos
government	vaccinale	italiachiama	viaggio	pierosansonetti
hope	terapie	#abruzzo	ciao	libero_official
together	vaccinati	laguna	#enniomorricone	adginforma
democracy	vaccinazione	pubblicato	musica	edicola
people	registrati	palazzo	sergio	ilmessaggeroit
good	gialla	chigi	artista	corriere
always	prossima	d'alfonso	#gigiproietti	lapresse_news
us	vaccino	#italianinelmondo	natale	radioradicale
even	intensive	treviso	#davidsassoli	repubblica
right	intensiva	ddl	#battiato	avvenire_nei
can	terapia	all'unanimità	curini	corriere
must	all'aperto	senato	#carlafracci	fuortes
today	decessi	dolomiti	icona	askanews_ita
really	#vaccino	ricevo	amato	ilprimaton
one	ricoveri	museo	onomastico	agenzia_dire
now	test	villa	poeta	agenzia_italia
rights	somministrazioni	canova	#monicavitti	ilriformista
years	rapidi	ig	#festadellarepubblica	messaggero
colinphoenix	settimana	neve	carla	il_piccolo
get	zona	belluno	lucianoghelfi	l'intervista
friend	prenotazioni	zan	simonamalpezzi	cda
newwaveandpunk	contagi	falsa	martina_carone	italpress
welikeduel	arancione	camera	sassoli	formichenews
make	guariti	mostra	appassionato	scanzi
work	ricoverati	deputati	#epifani	mattinodinapoli
ever	l'obbligo	pdabruzzo	eletta	ansaromalazio

Topic6	Topic7	Topic8	Topic9	Topic10
rapite	volontari	risposte	pour	strage
dall'oglio	vigili	economica	iva	scorta
#padredalloglio	infermieri	scelte	bollette	#foibe
paoladelusa	#forzearmate	responsabilità	avec	uccisi
tornino	poliziadistato	decisioni	fiscali	attanasio
sostenibilità	armate	sanitaria	autonomi	ucciso
ecologica	gdf	gestire	du	#giornodelricordo
transizione	svolto	concrete	cartelle	tragedia
ambientale	emergenzavvf	sociali	au	#giornatadellamemoria
sostenibile	ringraziare	tavolo	esattoriali	innocenti
climatici	penitenziaria	opposizioni	et	uccise
pianeta	ringraziamento	#lockdown	perduto	iacovacci
innovazione	dell'ordine	ripresa	dans	attentato
dell'ambiente	#vigilidelfuoco	affrontare	cassa	ferita
cambiamenti	divisa	#fase2	fiscale	l'orrore
angelazoppo	sm_difesa	l'emergenza	prestiti	vittime
loops40994697	_carabinieri_	necessarie	fatturato	#aldomoro
alatigiulio	italiannavy	ripartire	contributi	foibe
scureggione	ministerodifesa	collaborare	scadenze	#congo
l'ambiente	mamme	chiare	credito	#giornodellamemoria
gfi65	soccorso	giuste	pagamenti	ucraino
stretto	svolgono	soluzioni	tax	#falcone
massionline	compleanno	esecutivo	mutui	odio
renzo_pisu	#festadellamamma	uscire	tasse	#paoloborsellino
aledeniz	plauso	#smartworking	nous	violenza
gianni_dragoni	prezioso	servono	liquidità	uccisa
task	mille	messe	bonus	l'odio
sviluppo	#carabinieri	serve	sur	falcone
climatico	esercito	fronteggiare	prezzi	vittima
ugoarrigo	nonni	superare	#bollette	persero

Topic11	Topic12	Topic13	Topic14	Topic15
l'aggiornamento		diretta	#sanità	para
bollettino	5s	seguitemi	#quirinale	por
docenti	imbarazzante	aspetto	regione_sicilia	el
pittoni	grillini	ospite	#parlamento	giusvapulejo
scuola	franferrante	aggiornamenti	#quirinale2022	los
orizzontescuola	imbecille	streaming	#presidentedellarepubblica	las
ordinario	vabbè	interverrò	#governomusumeci	#amala
studenti	brutta	rete4	#verità	paolabottelli
orizzonte	capito	domattina	#rassegnastampa	tommasolabate
didattica	ridicolo	parleremo	#buongiorno	dispiace
#scuola	a_lisacorrado	seguiteci	#libero	italianos
scolastica	dica	rassegna	#salute	esta
scolastico	robdellaseta	seguite	#primapagina	hoy
concorsi	poltrone	facebook	#edicola	mariolavia
paritarie	l'unico	collegatevi	#giornali	nomfup
insegnanti	letta	sull'emergenza	#tempo	juve
asili	neanche	canale	#giornale	cottarellicpi
scolastici	legalemeglio	parteciperò	#consultazioni	exterior
azzolina	teatrino	mancate	#edicolalucidi	gobierno
campania	zingaretti	fb	coraggio_italia	adalucde
alunni	virologo	vediamo	#fiducia	gianlookingfor
azzolinalucia	povero	organizzato	#lavoro	appunto
scuole	talmente	seguire	eleggere	gracias
nido	leu	link	#maggioranza	#inter
precari	smentisce	l'audizione	#leu	ardigiorgio
classe	#arcuri	#danielasantanche	#europa	todos
#azzolina	giasilvestrini	raitre	#covid19italia	mancava
#dad	doveva	parlerò	#infrastrutture	lauracesaretti1
dell'istruzione	pur	conoscitiva	terni	grotondi
dad	monicafrassoni	stasera	#emergenza	il_cappellini

Topic16	Topic17	Topic18
#tokyo2020	#iostoconsalvini	youtube
#italiateam	molinaririk	speriamo
medaglie	maxromeomb	arrivato
tokyo	angelociocca	all'estero
atleti	#processateancheme	finito
olimpiadi	#salvinipremier	confermato
azzurri	patriziarametta	ricorso
alex	votalega	fatti
medaglia	lega_senato	deciso
argento	sbarchi	tratta
oro	a_gusmeroli	dato
bronzo	noiconsalvini	diversi
podio	ponytaele	germania
#borgonzonipresidente	lacavandoli	rapporti
#olimpiadi	lntoscana	riconosce
#euro2020	legacamera	bruxelles
#paralimpiadi	clandestina	penale
#paralympics	massimobitonci	l'ora
#jacobs	lucabattanta	sostengo
#berrettini	clandestini	lanciato
#tamberi	legasalvini	corruzione
vince	matteosalvinimi	sospeso
federica	#blocconavale	l'abbiamo
forza	#primagliitaliani	ilpost
azzurro	massimogara	tocca
#giochiolimpici	alex63roy	strada
fi_ultimissime	giuliocentemero	mediazione
paralimpiadi	lampedusa	passi
#pechino2022	albertobagnai	vedono
l'oro	alessia_smile6	appello

```
# PROB
PROBmySTM18NoG <- t(as.matrix(labeledtpic[["prob"]]))
kable(PROBmySTM18NoG[,1:5], col.names = c("Topic1","Topic2","Topic3",
                                          "Topic4","Topic5")) %>%
  kable_styling(latex_options="scale_down")
```

```
kable(PROBmySTM18NoG[,6:10], col.names = c("Topic6","Topic7","Topic8",
                                           "Topic9","Topic10")) %>%
  kable_styling(latex_options="scale_down")
```

```
kable(PROBmySTM18NoG[,11:15], col.names = c("Topic11","Topic12","Topic13",
                                            "Topic14","Topic15")) %>%
  kable_styling(latex_options="scale_down")
```

```
kable(PROBmySTM18NoG[,16:18], col.names = c("Topic16","Topic17","Topic18"))
```

Topic1	Topic2	Topic3	Topic4	Topic5
mov5stelle	covid	legge	presidente	via
leggi	dati	commissione	grande	stampasgarbi
parlamentari	vaccini	senato	anni	intervista
giorgiameloni	vaccino	camera	buona	rai
giuseppeconteit	#covid19	appena	pdnetwork	parlato
scritto	settimana	foto	mondo	repubblica
post	virus	aula	storia	vocedelpatriota
fratelliditalia	vaccinale	video	buon	fattoquotidiano
m5s_senato	#covid	palazzo	repubblica	pubblico
social	regioni	voto	italiana	politica
italymfa	#greenpass	lavori	donna	matteorenzi
amp	campagna	proposta	famiglia	sud
cina	numero	città	città	leggere
tweet	regione	deputati	cultura	corriere
usa	vaccinati	gruppo	uomo	draghi
l'ho	lombardia	approvato	festa	ilgiornale
#m5s	contagi	pubblicata	deputatipd	recovery
referendum	tamponi	testo	comunità	anzaldi
luigidimaio	zona	#ddlzan	politica	adnkronos
fdi_parlamento	casi	sede	enricoletta	direttore
corte	italia	venezia	mattarella	libero_official
costituzionale	mascherine	milano	italiano	lettera
italy	settimane	#venezia	tanti	quotidiano
news	positivi	ddl	auguri	giornalisti
letto	#vaccini	veneto	giornata	parlo
#referendum	dosi	firma	natale	lastampa
scrivere	test	centro	paese	riflessione
fake	#liguria	giunta	amico	nord
trump	primi	roma	persona	sole
#iovotono	numeri	approvata	simbolo	articolo

Topic6	Topic7	Topic8	Topic9	Topic10
presto	grazie	governo	imprese	solidarietà
sviluppo	lavoro	paese	milioni	libertà
territorio	donne	cittadini	euro	anni
settore	buon	crisi	decreto	parole
progetto	auguri	serve	miliardi	diritti
qualità	forze	l'italia	famiglie	vittime
ricerca	impegno	momento	lavoratori	democrazia
mondo	uomini	dare	fondo	violenza
tutela	servizio	responsabilità	tasse	guerra
importante	giornata	#coronavirus	riforma	giustizia
nuove	linea	sicurezza	bilancio	popolo
opportunità	medici	pandemia	fiscale	memoria
crescita	personale	salute	governo	vicinanza
investimenti	generale	misure	aziende	morte
comuni	sicurezza	bisogna	sostegno	minacce
futuro	cuore	italiani	misura	pensiero
transizione	civile	piano	emendamento	verità
turismo	tanti	politiche	stop	pace
infrastrutture	ringrazio	parlamento	soldi	dovere
sostenibile	dell'ordine	emergenza	dl	rispetto
progetti	anni	lavorare	aiuti	diritto
territori	operatori	dobbiamo	difficoltà	donne
italiana	protezione	ripartire	risorse	ricordo
modello	polizia	economica	aiutare	inaccettabile
centro	compleanno	bene	de	tragedia
sociale	paese	servono	approvato	dolore
digitale	nazionale	priorità	pagare	ricordare
produzione	sanitario	mettere	bonus	atto
città	difesa	devono	reddito	valori
fondamentale	l'impegno	lavoro	mld	mafia

Topic11	Topic12	Topic13	Topic14	Topic15
scuola	conte	diretta	paese	de
giovani	pd	domani	futuro	vero
scuole	governo	consiglio	#governo	ragione
ragazzi	vuole	punto	#draghi	credo
bambini	m5s	sera	sanità	vedere
settembre	sinistra	parlare	vogliamo	giusto
regione	pass	mattina	#italia	bene
figli	maggioranza	aspetto	#senato	penso
studenti	partito	stampa	pubblica	porta
#scuola	andare	amici	politica	carlocalenda
presenza	green	intervento	presidente	differenza
ministra	cose	vediamo	bisogno	bravo
#covid19	draghi	conferenza	#m5s	magari
notizie	#conte	ospite	costruire	posso
ministero	dovrebbe	pagina	#lavoro	dico
formazione	italiani	gt	dobbiamo	serie
ministro	voto	stasera	#leu	ovviamente
campania	parlamento	vista	#mes	peggio
classe	parla	incontro	#pnrr	guidocrosetto
distanza	problema	link	l'italia	gran
famiglie	destra	confronto	riforme	assolutamente
diritto	votare	libro	sistema	en
studio	sa	ministri	#sicilia	partita
dell'unità	capire	facebook	#recoveryfund	resto
bollettino	casa	presentazione	riforma	el
l'aggiornamento	male	pomeriggio	forza_italia	vedo
genitori	cittadinanza	coronavirus	italiaviva	spero
personale	pensa	insieme	giuseppeconteit	persona
scolastico	bene	seguire	#europa	sbagliato
riapertura	davvero	presidente	visione	azione_it

Topic16	Topic17	Topic18
italia	lega	bene
forza	matteosalvinimi	politica
forza_italia	salvini	fatti
sindaco	#lega	anni
insieme	italiani	dato
roma	fratelliditalia	anno
l'italia	#salvini	ministro
battaglia	legasalvini	europa
viva	matteo	strada
notizia	#fratelliditalia	casa
grande	ministro	realtà
bella	meloni	parlamento
elettorale	d'italia	risposta
complimenti	processo	deciso
coraggio	governo	tratta
vittoria	#iostoconsalvini	italiaviva
elezioni	legacamera	possibile
sport	fratelli	pochi
#roma	vergogna	movimento
squadra	borghi_claudio	all'estero
berlusconi	#meloni	scelta
candidato	testa	tante
risultato	lega_senato	unico
bocca	migranti	speriamo
piazza	piazza	arriva
centrodestra	confini	arrivato
campagna	casa	nuova
vincere	giorgia	continua
uniti	clandestini	roma
regionali	lamorgese	rispetto

Meaningfull labels with the first 10 FREX word associated

```
labeledtpic <- labelTopics(mySTM18NoG, n=10)

FREXmySTM18NoG <- t(as.matrix(labeledtpic[["frex"]]))
```

6) Sustainable energy	7) Categories involved in covid emergency	8) Economic relaunch
rapite	volontari	risposte
dall'oglio	vigili	economica
#padredalloglio	infermieri	scelte
paoladelusa	#forzearmate	responsabilità
tornino	poliziadistato	decisioni
sostenibilità	armate	sanitaria
ecologica	gdf	gestire
transizione	svolto	concrete
ambientale	emergenzavvf	sociali
sostenibile	ringraziare	tavolo

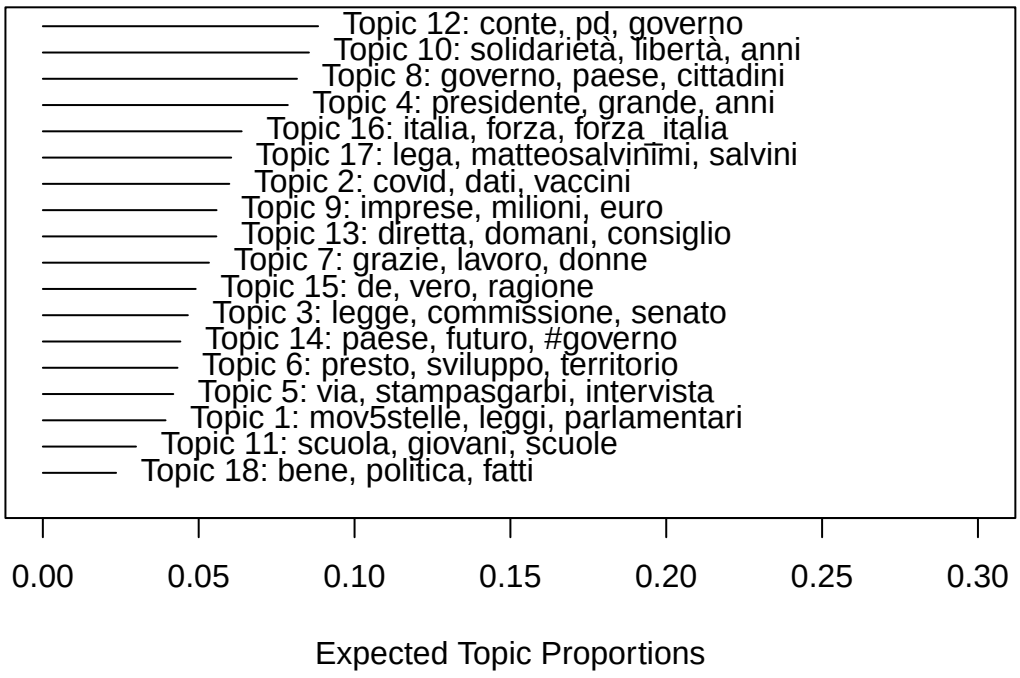
9) Economic hardship and taxes	10) Victims of violent deaths	11) Public education
pour	strage	l'aggiornamento
iva	scorta	bollettino
bollette	#foibe	docenti
avec	uccisi	pittoni
fiscali	attanasio	scuola
autonomi	ucciso	orizzontescuola
du	#giornodelricordo	ordinario
cartelle	tragedia	studenti
au	#giornatadellamemoria	orizzonte
esattoriali	innocenti	didattica

2) Covid-19	4) Epitaphs	5) Journals and media
dosi	maestro	anzaldi
dose	mancherà	lastampa
molecolari	addio	intervista
tamponi	mancherai	adnkronos
vaccinale	viaggio	pierosansonetti
terapie	ciao	libero_official
vaccinati	#enniomorricone	adginforma
vaccinazione	musica	edicola
registrati	sergio	ilmessaggeroit
gialla	artista	corriere

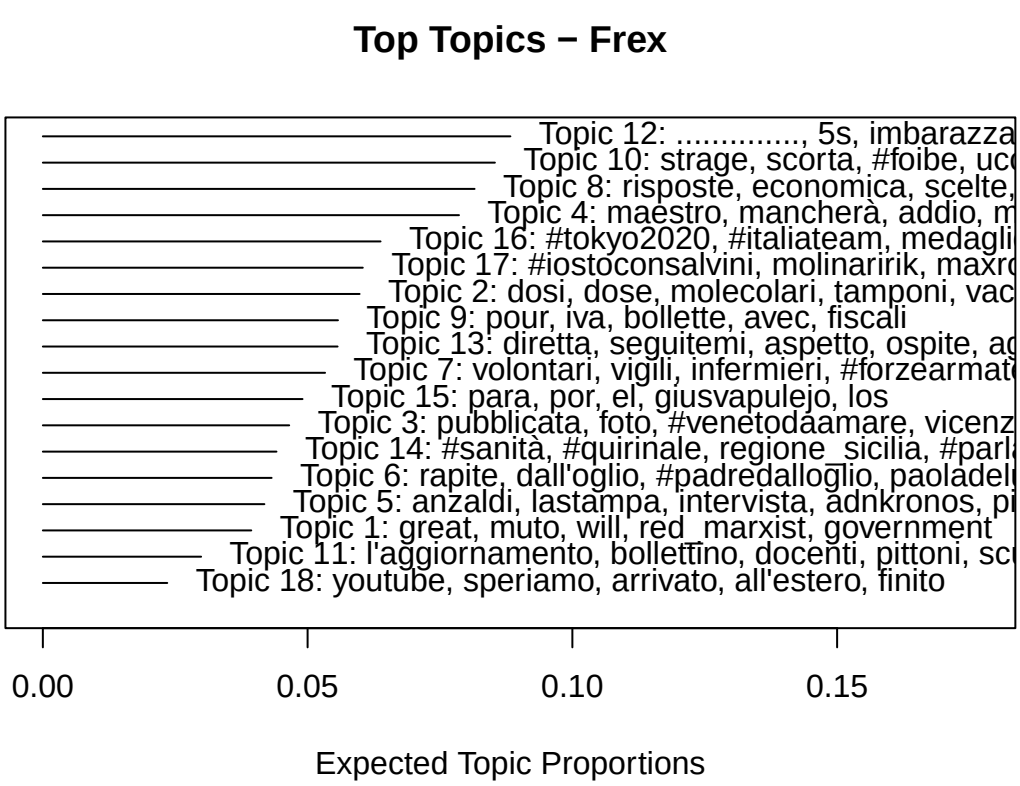
12) Anti-elitism	13) Social and TV live broadcasts	16) Olympics game	17) Right-wing party topics
	diretta	#tokyo2020	#iostoconsalvini
5s	seguitemi	#italiateam	molinaririk
imbarazzante	aspetto	medaglie	maxromeomb
grillini	ospite	tokyo	angelociocca
franferrante	aggiornamenti	atleti	#processateancheme
imbecille	streaming	olimpiadi	#salvinipremier
vabbè	interverrò	azzurri	patriziarametta
brutta	rete4	alex	votalega
capito	domattina	medaglia	lega_senato
ridicolo	parleremo	argento	sbarchi

Most frequent topic

```
plot(mySTM18NoG, type = "summary", xlim = c(0, .3),
     main = "Top Topics - Prob")
```




```
# plot just frex words for each topic
plot(mySTM18NoG, type = "summary", labeltype = c("frex"), n=5,
     main = "Top Topics - Frex")
```



Which are the the most likely topics across our documents?

```
tab <- table(apply(mySTM18NoG$theta,1,which.max))
topics_label <- c("01) Junk topic",
                  "02) Covid-19",
                  "03) Junk topic",
                  "04) Epitaphs",
                  "05) Journals and media",
                  "06) Sustainable energy",
                  "07) Categories involved in the covid emergency",
                  "08) Economic relaunch",
                  "09) Economic hardship and taxes",
                  "10) Victims of violent deaths",
                  "11) Public education",
                  "12) Anti-elitism",
                  "13) Social and TV live broadcasts",
                  "14) Junk topic",
                  "15) Junk topic",
                  "16) Olympics game",
                  "17) Right-wing parties topic",
                  "18) Junk topic")
tab <- as.matrix(tab)
tab2 <- cbind(topics_label,tab)
tab2 <- as.data.frame(tab2)
```

```
colnames(tab2) <- c("Topic label","Freq")
tab2$Freq <- as.numeric(tab2$Freq)
kable(tab2 %>% arrange(desc(Freq)))
```

	Topic label	Freq
12	12) Anti-elitism	41212
10	10) Victims of violent deaths	38252
8	08) Economic relaunch	31610
4	04) Epitaphs	30860
17	17) Right-wing parties topic	26510
16	16) Olympics game	23121
9	09) Economic hardship and taxes	22513
2	02) Covid-19	22233
13	13) Social and TV live broadcasts	20170
15	15) Junk topic	19391
7	07) Categories involved in the covid emergency	17417
6	06) Sustainable energy	16688
3	03) Junk topic	15645
5	05) Journals and media	14788
14	14) Junk topic	14516
1	01) Junk topic	13234
11	11) Public education	9454
18	18) Junk topic	627

Save them back in the original corpus

```
subs_corpus$topic <- apply(mySTM18NoG$theta,1,which.max)
```

Find the most associated document for each topics

This list of 18 items represent the respective document with highest theta for each topic ordered from 1 to 18.

```
apply(mySTM18NoG$theta,2,which.max)
```

```
##  [1]  12710   1080  26346  52361  41589 198020 234701   8705  12415 248340
## [11]  80644 353132 200651 342504 267537 162724 199419  22068
```

```
Tweet_number <- apply(mySTM18NoG$theta,2,which.max)

kable(cbind(topics_label,Tweet_number))
```

topics_label	Tweet_number
01) Junk topic	12710
02) Covid-19	1080
03) Junk topic	26346
04) Epitaphs	52361
05) Journals and media	41589
06) Sustainable energy	198020
07) Categories involved in the covid emergency	234701
08) Economic relaunch	8705
09) Economic hardship and taxes	12415
10) Victims of violent deaths	248340
11) Public education	80644
12) Anti-elitism	353132
13) Social and TV live broadcasts	200651
14) Junk topic	342504
15) Junk topic	267537
16) Olympics game	162724
17) Right-wing parties topic	199419
18) Junk topic	22068

Coefficients

```
# set PD as reference
DfmStm$meta$party_id <- relevel(as.factor(DfmStm$meta$party_id), ref = "PD")

prep_K18NoG <- estimateEffect(1:18 ~ party_id + populism + s(quarter),
                      mySTM18NoG,metadata = DfmStm$meta,
                      uncertainty = "Global")

#save(prep_K18NoG,file="data/prep_K18NoG.Rda")
```

Regression coefficients for all topics are shown here

01) Junk topic

```
summary(prep_K18NoG, topics = 1)
```

```
##
## Call:
## estimateEffect(formula = 1:18 ~ party_id + populism + s(quarter),
##     stmobj = mySTM18NoG, metadata = DfmStm$meta, uncertainty = "Global")
##
##
## Topic 1:
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           3.977e-02  7.255e-04  54.823  < 2e-16 ***
## party_idCI           -2.107e-02  1.386e-03 -15.205  < 2e-16 ***
## party_idFDI           4.563e-03  7.832e-04   5.827 5.65e-09 ***
## party_idFI           -4.996e-03  5.739e-04  -8.706  < 2e-16 ***
## party_idINDIPENDENTE -2.394e-02  2.498e-03  -9.581  < 2e-16 ***
## party_idIV           -1.828e-02  2.122e-03  -8.612  < 2e-16 ***
## party_idLEGA         -1.196e-02  6.593e-04 -18.135  < 2e-16 ***
## party_idLEU          -1.537e-02  1.453e-03 -10.579  < 2e-16 ***
## party_idM5S           2.441e-02  7.314e-04  33.368  < 2e-16 ***
## party_idMISTO        -3.981e-03  7.312e-04  -5.444 5.21e-08 ***
## party_idREG_LEAGUES  -2.675e-02  2.929e-03  -9.134  < 2e-16 ***
## populism             -7.219e-05  6.354e-06 -11.363  < 2e-16 ***
## s(quarter)1           2.837e-02  1.337e-02   2.121 0.033910 *
## s(quarter)2          -1.526e-02  5.661e-03  -2.695 0.007032 **
## s(quarter)3           2.573e-02  2.882e-03   8.928  < 2e-16 ***
## s(quarter)4           1.152e-02  1.725e-03   6.679 2.41e-11 ***
## s(quarter)5          -1.998e-03  1.285e-03  -1.556 0.119811
## s(quarter)6           4.476e-03  1.169e-03   3.829 0.000129 ***
## s(quarter)7           5.941e-03  1.822e-03   3.262 0.001107 **
## s(quarter)8          -6.245e-03  1.782e-03  -3.503 0.000459 ***
## s(quarter)9          -7.494e-04  1.582e-03  -0.474 0.635768
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

02) Covid-19

```
summary(prep_K18NoG, topics = 2)
```

```
##
## Call:
## estimateEffect(formula = 1:18 ~ party_id + populism + s(quarter),
##     stmobj = mySTM18NoG, metadata = DfmStm$meta, uncertainty = "Global")
##
##
## Topic 2:
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           4.149e-02  8.826e-04  47.008  < 2e-16 ***
## party_idCI            1.787e-03  1.807e-03   0.989 0.322489
## party_idFDI          -6.548e-03  8.987e-04  -7.286 3.19e-13 ***
## party_idFI            1.045e-02  7.350e-04  14.212  < 2e-16 ***
## party_idINDIPENDENTE  4.498e-02  3.140e-03  14.324  < 2e-16 ***
## party_idIV           -1.084e-02  2.375e-03  -4.563 5.05e-06 ***
## party_idLEGA          1.983e-02  7.299e-04  27.172  < 2e-16 ***
## party_idLEU          -1.510e-02  1.504e-03 -10.046  < 2e-16 ***
## party_idM5S           4.135e-04  7.560e-04   0.547 0.584369
## party_idMISTO         1.978e-02  1.039e-03  19.046  < 2e-16 ***
## party_idREG_LEAGUES   2.335e-02  3.945e-03   5.919 3.24e-09 ***
## populism              7.151e-05  7.857e-06   9.102  < 2e-16 ***
## s(quarter)1           1.822e-01  1.279e-02  14.247  < 2e-16 ***
## s(quarter)2          -7.049e-02  5.704e-03 -12.359  < 2e-16 ***
## s(quarter)3           2.316e-02  3.221e-03   7.188 6.59e-13 ***
## s(quarter)4           7.325e-03  2.138e-03   3.426 0.000614 ***
## s(quarter)5           3.165e-02  1.504e-03  21.046  < 2e-16 ***
## s(quarter)6           1.348e-02  1.395e-03   9.660  < 2e-16 ***
## s(quarter)7           1.727e-02  1.862e-03   9.275  < 2e-16 ***
## s(quarter)8          -7.363e-04  2.165e-03  -0.340 0.733784
## s(quarter)9          -1.259e-02  1.743e-03  -7.222 5.14e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

03) Junk topic

```
summary(prep_K18NoG, topics = 3)
```

```
##
## Call:
## estimateEffect(formula = 1:18 ~ party_id + populism + s(quarter),
##     stmobj = mySTM18NoG, metadata = DfmStm$meta, uncertainty = "Global")
##
##
## Topic 3:
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           3.148e-02  7.378e-04  42.668  < 2e-16 ***
## party_idCI            7.255e-02  2.137e-03  33.946  < 2e-16 ***
## party_idFDI          -4.523e-03  8.608e-04  -5.255 1.48e-07 ***
## party_idFI           -2.879e-03  6.226e-04  -4.624 3.76e-06 ***
## party_idINDIPENDENTE -4.258e-03  2.611e-03  -1.631  0.10289
## party_idIV            1.607e-03  2.234e-03   0.720  0.47178
## party_idLEGA          2.023e-02  6.240e-04  32.415  < 2e-16 ***
## party_idLEU          -7.666e-03  1.429e-03  -5.366 8.07e-08 ***
## party_idM5S           6.549e-03  7.164e-04   9.142  < 2e-16 ***
## party_idMISTO         1.273e-02  8.882e-04  14.328  < 2e-16 ***
## party_idREG_LEAGUES   6.194e-03  3.536e-03   1.752  0.07979 .
## populism             -7.608e-05  6.662e-06 -11.419  < 2e-16 ***
## s(quarter)1           1.297e-02  1.183e-02   1.096  0.27287
## s(quarter)2          -1.335e-02  5.086e-03  -2.626  0.00864 **
## s(quarter)3           3.366e-02  2.783e-03  12.094  < 2e-16 ***
## s(quarter)4          -9.382e-03  1.671e-03  -5.615 1.97e-08 ***
## s(quarter)5           2.015e-02  1.293e-03  15.588  < 2e-16 ***
## s(quarter)6           1.472e-02  1.325e-03  11.105  < 2e-16 ***
## s(quarter)7           3.393e-02  1.623e-03  20.913  < 2e-16 ***
## s(quarter)8           2.929e-03  2.014e-03   1.454  0.14595
## s(quarter)9           2.202e-02  1.838e-03  11.981  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

04) Epitaphs

```
summary(prep_K18NoG, topics = 4)
```

```
##
## Call:
## estimateEffect(formula = 1:18 ~ party_id + populism + s(quarter),
##     stmobj = mySTM18NoG, metadata = DfmStm$meta, uncertainty = "Global")
##
##
## Topic 4:
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           8.350e-02  9.136e-04  91.397  < 2e-16 ***
## party_idCI            2.128e-02  2.273e-03   9.363  < 2e-16 ***
## party_idFDI          -4.149e-02  9.496e-04 -43.695  < 2e-16 ***
## party_idFI           -2.065e-02  8.667e-04 -23.823  < 2e-16 ***
## party_idINDIPENDENTE -1.625e-02  3.258e-03  -4.987 6.12e-07 ***
## party_idIV           -7.140e-03  2.608e-03  -2.737   0.0062 **
## party_idLEGA         -4.061e-02  7.243e-04 -56.063  < 2e-16 ***
## party_idLEU          -3.341e-02  1.835e-03 -18.204  < 2e-16 ***
## party_idM5S          -3.668e-02  7.696e-04 -47.661  < 2e-16 ***
## party_idMISTO        -3.676e-02  9.488e-04 -38.745  < 2e-16 ***
## party_idREG_LEAGUES  -3.184e-02  4.077e-03  -7.809 5.80e-15 ***
## populism             -2.239e-04  7.844e-06 -28.547  < 2e-16 ***
## s(quarter)1           2.277e-03  1.433e-02   0.159   0.8737
## s(quarter)2          -3.787e-03  6.134e-03  -0.617   0.5370
## s(quarter)3           2.202e-02  3.396e-03   6.483 8.98e-11 ***
## s(quarter)4           1.381e-02  2.059e-03   6.708 1.98e-11 ***
## s(quarter)5           2.382e-02  1.653e-03  14.415  < 2e-16 ***
## s(quarter)6           1.609e-02  1.571e-03  10.239  < 2e-16 ***
## s(quarter)7           1.803e-02  2.107e-03   8.557  < 2e-16 ***
## s(quarter)8           4.123e-02  2.293e-03  17.980  < 2e-16 ***
## s(quarter)9           2.745e-02  2.220e-03  12.360  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

05) Journals and media

```
summary(prep_K18NoG, topics = 5)
```

```
##
## Call:
## estimateEffect(formula = 1:18 ~ party_id + populism + s(quarter),
##     stmobj = mySTM18NoG, metadata = DfmStm$meta, uncertainty = "Global")
##
##
## Topic 5:
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           4.227e-02  7.201e-04  58.700  < 2e-16 ***
## party_idCI           -2.197e-02  1.372e-03 -16.015  < 2e-16 ***
## party_idFDI          -5.427e-03  8.169e-04  -6.644 3.06e-11 ***
## party_idFI           -1.715e-02  6.198e-04 -27.672  < 2e-16 ***
## party_idINDIPENDENTE -3.003e-02  2.282e-03 -13.161  < 2e-16 ***
## party_idIV           -1.348e-02  2.480e-03  -5.438 5.39e-08 ***
## party_idLEGA         -2.786e-02  5.339e-04 -52.182  < 2e-16 ***
## party_idLEU          -1.237e-02  1.533e-03  -8.068 7.17e-16 ***
## party_idM5S          -4.992e-03  6.552e-04  -7.618 2.58e-14 ***
## party_idMISTO         2.931e-02  8.739e-04  33.533  < 2e-16 ***
## party_idREG_LEAGUES  -1.081e-02  3.070e-03  -3.521  0.00043 ***
## populism             -8.642e-05  6.430e-06 -13.441  < 2e-16 ***
## s(quarter)1          -1.625e-02  1.048e-02  -1.551  0.12100
## s(quarter)2           1.926e-02  4.694e-03   4.104 4.07e-05 ***
## s(quarter)3           4.983e-03  2.497e-03   1.996  0.04596 *
## s(quarter)4           1.806e-02  1.594e-03  11.325  < 2e-16 ***
## s(quarter)5           1.020e-02  1.256e-03   8.120 4.68e-16 ***
## s(quarter)6           1.101e-02  1.310e-03   8.403  < 2e-16 ***
## s(quarter)7           3.340e-02  1.652e-03  20.213  < 2e-16 ***
## s(quarter)8          -3.805e-03  2.030e-03  -1.874  0.06087 .
## s(quarter)9           8.658e-03  1.504e-03   5.755 8.65e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

06) Sustainable energy

```
summary(prep_K18NoG, topics = 6)
```

```
##
## Call:
## estimateEffect(formula = 1:18 ~ party_id + populism + s(quarter),
##     stmobj = mySTM18NoG, metadata = DfmStm$meta, uncertainty = "Global")
##
##
## Topic 6:
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           5.310e-02  8.458e-04  62.777  < 2e-16 ***
## party_idCI            1.071e-02  1.966e-03   5.447 5.12e-08 ***
## party_idFDI          -2.795e-02  8.131e-04 -34.370  < 2e-16 ***
## party_idFI           -1.048e-02  7.930e-04 -13.214  < 2e-16 ***
## party_idINDIPENDENTE  4.737e-02  3.780e-03  12.531  < 2e-16 ***
## party_idIV            7.711e-03  2.873e-03   2.684  0.00728 **
## party_idLEGA         -1.863e-02  6.461e-04 -28.831  < 2e-16 ***
## party_idLEU           3.996e-02  1.945e-03  20.539  < 2e-16 ***
## party_idM5S           1.563e-02  7.509e-04  20.815  < 2e-16 ***
## party_idMISTO        -1.036e-02  9.151e-04 -11.324  < 2e-16 ***
## party_idREG_LEAGUES   3.475e-02  5.257e-03   6.610 3.86e-11 ***
## populism             -1.442e-04  6.797e-06 -21.208  < 2e-16 ***
## s(quarter)1          -3.314e-02  1.361e-02  -2.436  0.01487 *
## s(quarter)2           1.495e-02  6.218e-03   2.405  0.01617 *
## s(quarter)3           2.272e-03  3.199e-03   0.710  0.47769
## s(quarter)4          -9.656e-03  2.089e-03  -4.623 3.79e-06 ***
## s(quarter)5           2.146e-02  1.471e-03  14.589  < 2e-16 ***
## s(quarter)6          -1.118e-03  1.517e-03  -0.737  0.46111
## s(quarter)7           1.346e-02  2.067e-03   6.514 7.31e-11 ***
## s(quarter)8          -5.396e-03  2.151e-03  -2.509  0.01211 *
## s(quarter)9           1.388e-02  2.025e-03   6.856 7.11e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

07) Categories involved in the covid emergency

```
summary(prep_K18NoG, topics = 7)
```

```
##
## Call:
## estimateEffect(formula = 1:18 ~ party_id + populism + s(quarter),
##     stmobj = mySTM18NoG, metadata = DfmStm$meta, uncertainty = "Global")
##
##
## Topic 7:
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           6.785e-02  9.100e-04  74.564  < 2e-16 ***
## party_idCI            9.541e-03  1.776e-03   5.373 7.76e-08 ***
## party_idFDI          -1.509e-02  8.099e-04 -18.633  < 2e-16 ***
## party_idFI            1.436e-03  6.469e-04   2.220 0.026393 *
## party_idINDIPENDENTE  2.078e-02  3.228e-03   6.437 1.22e-10 ***
## party_idIV            3.284e-02  2.874e-03  11.427  < 2e-16 ***
## party_idLEGA         -6.553e-03  6.245e-04 -10.493  < 2e-16 ***
## party_idLEU           4.650e-03  1.723e-03   2.699 0.006962 **
## party_idM5S           9.567e-03  8.122e-04  11.780  < 2e-16 ***
## party_idMISTO        -2.006e-02  8.145e-04 -24.623  < 2e-16 ***
## party_idREG_LEAGUES   4.369e-04  4.296e-03   0.102 0.919007
## populism             -7.371e-05  6.833e-06 -10.787  < 2e-16 ***
## s(quarter)1           4.331e-02  1.224e-02   3.538 0.000403 ***
## s(quarter)2          -2.664e-02  5.375e-03  -4.955 7.23e-07 ***
## s(quarter)3          -1.207e-02  2.875e-03  -4.198 2.69e-05 ***
## s(quarter)4          -1.429e-02  1.919e-03  -7.445 9.69e-14 ***
## s(quarter)5          -5.243e-04  1.525e-03  -0.344 0.731031
## s(quarter)6          -1.986e-02  1.346e-03 -14.750  < 2e-16 ***
## s(quarter)7          -9.536e-03  1.803e-03  -5.288 1.24e-07 ***
## s(quarter)8          -1.558e-02  2.046e-03  -7.615 2.63e-14 ***
## s(quarter)9          -3.892e-03  1.764e-03  -2.206 0.027351 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

08) Economic relaunch

```
summary(prep_K18NoG, topics = 8)
```

```
##
## Call:
## estimateEffect(formula = 1:18 ~ party_id + populism + s(quarter),
##     stmobj = mySTM18NoG, metadata = DfmStm$meta, uncertainty = "Global")
##
##
## Topic 8:
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.096e-01  9.644e-04 113.656  < 2e-16 ***
## party_idCI           -5.204e-03  1.886e-03  -2.759 0.005807 **
## party_idFDI           1.027e-03  9.348e-04   1.099 0.271919
## party_idFI            7.865e-03  8.453e-04   9.303  < 2e-16 ***
## party_idINDIPENDENTE -8.233e-03  3.114e-03  -2.644 0.008196 **
## party_idIV            1.565e-02  2.807e-03   5.574 2.49e-08 ***
## party_idLEGA         -2.987e-02  6.838e-04 -43.679  < 2e-16 ***
## party_idLEU           1.551e-02  1.888e-03   8.212  < 2e-16 ***
## party_idM5S          -4.271e-03  8.556e-04  -4.991 6.01e-07 ***
## party_idMISTO        -6.882e-03  8.709e-04  -7.902 2.75e-15 ***
## party_idREG_LEAGUES  -5.070e-03  4.325e-03  -1.172 0.241042
## populism              1.782e-04  7.378e-06  24.159  < 2e-16 ***
## s(quarter)1           1.763e-01  1.314e-02  13.419  < 2e-16 ***
## s(quarter)2          -6.863e-02  5.669e-03 -12.106  < 2e-16 ***
## s(quarter)3          -2.261e-02  3.302e-03  -6.849 7.47e-12 ***
## s(quarter)4          -8.234e-03  2.204e-03  -3.735 0.000188 ***
## s(quarter)5          -3.519e-02  1.649e-03 -21.342  < 2e-16 ***
## s(quarter)6          -6.113e-02  1.398e-03 -43.713  < 2e-16 ***
## s(quarter)7          -5.239e-02  1.985e-03 -26.395  < 2e-16 ***
## s(quarter)8          -3.395e-02  2.163e-03 -15.700  < 2e-16 ***
## s(quarter)9          -4.887e-02  1.861e-03 -26.262  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

09) Economic hardship and taxes

```
summary(prep_K18NoG, topics = 9)
```

```
##
## Call:
## estimateEffect(formula = 1:18 ~ party_id + populism + s(quarter),
##     stmobj = mySTM18NoG, metadata = DfmStm$meta, uncertainty = "Global")
##
##
## Topic 9:
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           5.105e-02  9.106e-04  56.063  < 2e-16 ***
## party_idCI           -5.867e-03  1.674e-03  -3.505 0.000457 ***
## party_idFDI           9.391e-03  9.056e-04  10.370  < 2e-16 ***
## party_idFI            1.999e-02  7.349e-04  27.203  < 2e-16 ***
## party_idINDIPENDENTE -1.472e-02  2.817e-03  -5.225 1.74e-07 ***
## party_idIV            1.310e-02  2.848e-03   4.602 4.19e-06 ***
## party_idLEGA          1.096e-02  6.808e-04  16.103  < 2e-16 ***
## party_idLEU           1.343e-02  1.780e-03   7.543 4.59e-14 ***
## party_idM5S           4.030e-02  8.649e-04  46.598  < 2e-16 ***
## party_idMISTO         6.221e-03  9.296e-04   6.692 2.20e-11 ***
## party_idREG_LEAGUES   1.109e-02  3.995e-03   2.775 0.005519 **
## populism              3.898e-05  8.134e-06   4.793 1.65e-06 ***
## s(quarter)1           2.917e-01  1.382e-02  21.107  < 2e-16 ***
## s(quarter)2          -9.129e-02  6.085e-03 -15.003  < 2e-16 ***
## s(quarter)3           4.470e-02  3.201e-03  13.967  < 2e-16 ***
## s(quarter)4          -3.345e-02  1.890e-03 -17.703  < 2e-16 ***
## s(quarter)5          -7.896e-03  1.540e-03  -5.128 2.92e-07 ***
## s(quarter)6          -3.382e-02  1.515e-03 -22.326  < 2e-16 ***
## s(quarter)7           3.322e-03  1.855e-03   1.791 0.073351 .
## s(quarter)8          -1.651e-02  2.209e-03  -7.476 7.70e-14 ***
## s(quarter)9           3.724e-03  1.952e-03   1.908 0.056426 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

10) Victims of violent deaths

```
summary(prep_K18NoG, topics = 10)
```

```
##
## Call:
## estimateEffect(formula = 1:18 ~ party_id + populism + s(quarter),
##     stmobj = mySTM18NoG, metadata = DfmStm$meta, uncertainty = "Global")
##
##
## Topic 10:
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           8.042e-02  9.631e-04  83.498  < 2e-16 ***
## party_idCI           -4.190e-02  2.016e-03 -20.778  < 2e-16 ***
## party_idFDI          -1.393e-02  1.144e-03 -12.179  < 2e-16 ***
## party_idFI           -2.769e-02  9.163e-04 -30.222  < 2e-16 ***
## party_idINDIPENDENTE -5.448e-02  3.421e-03 -15.927  < 2e-16 ***
## party_idIV           -1.406e-03  3.510e-03  -0.401  0.68878
## party_idLEGA         -3.397e-02  8.537e-04 -39.788  < 2e-16 ***
## party_idLEU           2.854e-02  2.567e-03  11.118  < 2e-16 ***
## party_idM5S          -1.284e-02  9.655e-04 -13.298  < 2e-16 ***
## party_idMISTO        -2.552e-02  1.176e-03 -21.710  < 2e-16 ***
## party_idREG_LEAGUES  -3.309e-02  4.486e-03  -7.377 1.63e-13 ***
## populism              3.117e-04  9.397e-06  33.166  < 2e-16 ***
## s(quarter)1           1.285e-02  1.738e-02   0.739  0.45980
## s(quarter)2          -6.415e-03  7.361e-03  -0.871  0.38350
## s(quarter)3           1.621e-02  3.877e-03   4.181 2.91e-05 ***
## s(quarter)4           2.270e-03  2.511e-03   0.904  0.36606
## s(quarter)5           2.587e-02  1.804e-03  14.339  < 2e-16 ***
## s(quarter)6           2.006e-02  1.649e-03  12.164  < 2e-16 ***
## s(quarter)7          -5.844e-03  2.225e-03  -2.627  0.00862 **
## s(quarter)8           6.092e-02  2.458e-03  24.788  < 2e-16 ***
## s(quarter)9           5.473e-02  2.569e-03  21.300  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

11) Public education

```
summary(prep_K18NoG, topics = 11)
```

```
##
## Call:
## estimateEffect(formula = 1:18 ~ party_id + populism + s(quarter),
##     stmobj = mySTM18NoG, metadata = DfmStm$meta, uncertainty = "Global")
##
##
## Topic 11:
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           3.579e-02  6.786e-04  52.737  < 2e-16 ***
## party_idCI           -1.395e-03  1.749e-03  -0.798 0.424920
## party_idFDI          -1.331e-02  7.396e-04 -17.992  < 2e-16 ***
## party_idFI           -6.546e-03  6.731e-04  -9.725  < 2e-16 ***
## party_idINDIPENDENTE  8.763e-02  3.877e-03  22.601  < 2e-16 ***
## party_idIV            2.404e-02  2.380e-03  10.103  < 2e-16 ***
## party_idLEGA         -7.455e-03  5.667e-04 -13.157  < 2e-16 ***
## party_idLEU          -9.738e-03  1.294e-03  -7.523 5.36e-14 ***
## party_idM5S           1.248e-03  6.962e-04   1.792 0.073161 .
## party_idMISTO        -1.953e-03  7.528e-04  -2.595 0.009468 **
## party_idREG_LEAGUES  -1.223e-02  3.121e-03  -3.918 8.94e-05 ***
## populism             -1.265e-04  5.202e-06 -24.327  < 2e-16 ***
## s(quarter)1          -4.469e-02  1.200e-02  -3.725 0.000195 ***
## s(quarter)2           3.847e-02  5.054e-03   7.611 2.72e-14 ***
## s(quarter)3           6.853e-03  2.623e-03   2.613 0.008981 **
## s(quarter)4           7.296e-03  1.746e-03   4.179 2.93e-05 ***
## s(quarter)5           3.322e-03  1.125e-03   2.954 0.003135 **
## s(quarter)6           7.624e-03  1.186e-03   6.427 1.30e-10 ***
## s(quarter)7          -2.072e-03  1.574e-03  -1.317 0.187946
## s(quarter)8           9.452e-03  1.819e-03   5.196 2.03e-07 ***
## s(quarter)9           3.020e-03  1.520e-03   1.987 0.046968 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

12) Anti-elitism

```
summary(prep_K18NoG, topics = 12)
```

```
##
## Call:
## estimateEffect(formula = 1:18 ~ party_id + populism + s(quarter),
##     stmobj = mySTM18NoG, metadata = DfmStm$meta, uncertainty = "Global")
##
##
## Topic 12:
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           5.103e-02  9.087e-04  56.160  < 2e-16 ***
## party_idCI           -2.579e-02  1.707e-03 -15.111  < 2e-16 ***
## party_idFDI           4.789e-02  1.073e-03  44.655  < 2e-16 ***
## party_idFI            2.080e-02  7.785e-04  26.713  < 2e-16 ***
## party_idINDIPENDENTE -5.280e-02  2.744e-03 -19.242  < 2e-16 ***
## party_idIV           -2.936e-02  2.433e-03 -12.070  < 2e-16 ***
## party_idLEGA          3.912e-02  7.901e-04  49.514  < 2e-16 ***
## party_idLEU           1.078e-02  1.750e-03   6.159 7.32e-10 ***
## party_idM5S          -8.019e-03  8.426e-04  -9.517  < 2e-16 ***
## party_idMISTO         2.993e-02  1.042e-03  28.720  < 2e-16 ***
## party_idREG_LEAGUES  -3.947e-02  3.490e-03 -11.311  < 2e-16 ***
## populism              3.870e-04  9.667e-06  40.035  < 2e-16 ***
## s(quarter)1          -1.085e-01  1.347e-02  -8.057 7.86e-16 ***
## s(quarter)2           6.297e-02  5.877e-03  10.715  < 2e-16 ***
## s(quarter)3          -3.462e-02  3.092e-03 -11.194  < 2e-16 ***
## s(quarter)4           4.571e-02  2.031e-03  22.504  < 2e-16 ***
## s(quarter)5          -3.034e-02  1.490e-03 -20.361  < 2e-16 ***
## s(quarter)6           2.374e-02  1.603e-03  14.813  < 2e-16 ***
## s(quarter)7           2.498e-02  2.059e-03  12.134  < 2e-16 ***
## s(quarter)8          -5.490e-03  2.436e-03  -2.254   0.0242 *
## s(quarter)9          -4.097e-04  2.020e-03  -0.203   0.8392
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

13) Social and TV live broadcasts

```
summary(prep_K18NoG, topics = 13)
```

```
##
## Call:
## estimateEffect(formula = 1:18 ~ party_id + populism + s(quarter),
##     stmobj = mySTM18NoG, metadata = DfmStm$meta, uncertainty = "Global")
##
##
## Topic 13:
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           6.452e-02  8.657e-04  74.529  < 2e-16 ***
## party_idCI           -1.140e-02  1.640e-03  -6.953 3.58e-12 ***
## party_idFDI           2.160e-03  8.546e-04   2.528 0.011486 *
## party_idFI           -4.554e-03  7.379e-04  -6.172 6.76e-10 ***
## party_idINDIPENDENTE  3.919e-02  3.413e-03  11.484  < 2e-16 ***
## party_idIV            2.115e-02  2.791e-03   7.578 3.51e-14 ***
## party_idLEGA         -6.325e-03  6.473e-04  -9.773  < 2e-16 ***
## party_idLEU           1.580e-02  1.620e-03   9.755  < 2e-16 ***
## party_idM5S           4.561e-03  8.865e-04   5.145 2.67e-07 ***
## party_idMISTO         3.901e-03  9.512e-04   4.101 4.11e-05 ***
## party_idREG_LEAGUES  -7.741e-03  3.590e-03  -2.156 0.031081 *
## populism             -2.066e-04  6.414e-06 -32.205  < 2e-16 ***
## s(quarter)1          -2.227e-02  1.299e-02  -1.715 0.086400 .
## s(quarter)2           1.195e-03  5.708e-03   0.209 0.834167
## s(quarter)3          -1.596e-02  3.151e-03  -5.064 4.12e-07 ***
## s(quarter)4          -5.065e-03  1.969e-03  -2.572 0.010102 *
## s(quarter)5          -1.465e-03  1.521e-03  -0.963 0.335403
## s(quarter)6          -1.726e-02  1.299e-03 -13.281  < 2e-16 ***
## s(quarter)7           1.628e-03  1.957e-03   0.832 0.405615
## s(quarter)8          -1.578e-02  2.036e-03  -7.749 9.26e-15 ***
## s(quarter)9          -5.617e-03  1.688e-03  -3.328 0.000875 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

14) Junk topic

```
summary(prep_K18NoG, topics = 14)
```

```
##
## Call:
## estimateEffect(formula = 1:18 ~ party_id + populism + s(quarter),
##     stmobj = mySTM18NoG, metadata = DfmStm$meta, uncertainty = "Global")
##
##
## Topic 14:
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           5.725e-02  9.320e-04  61.422  < 2e-16 ***
## party_idCI            1.928e-02  1.825e-03  10.562  < 2e-16 ***
## party_idFDI          -1.129e-02  8.747e-04 -12.907  < 2e-16 ***
## party_idFI            1.403e-02  8.101e-04  17.315  < 2e-16 ***
## party_idINDIPENDENTE  4.833e-03  2.825e-03   1.711  0.08707 .
## party_idIV            1.090e-02  2.448e-03   4.454 8.43e-06 ***
## party_idLEGA         -2.535e-02  6.650e-04 -38.125  < 2e-16 ***
## party_idLEU           6.294e-03  1.794e-03   3.509  0.00045 ***
## party_idM5S           8.751e-03  7.463e-04  11.726  < 2e-16 ***
## party_idMISTO         4.109e-03  8.947e-04   4.593 4.37e-06 ***
## party_idREG_LEAGUES   1.557e-01  4.690e-03  33.204  < 2e-16 ***
## populism             -9.937e-06  6.933e-06  -1.433  0.15178
## s(quarter)1          -2.032e-01  1.184e-02 -17.155  < 2e-16 ***
## s(quarter)2           8.859e-02  5.180e-03  17.104  < 2e-16 ***
## s(quarter)3          -5.690e-02  2.791e-03 -20.388  < 2e-16 ***
## s(quarter)4           3.977e-02  2.074e-03  19.174  < 2e-16 ***
## s(quarter)5          -1.481e-02  1.320e-03 -11.215  < 2e-16 ***
## s(quarter)6          -1.938e-02  1.329e-03 -14.581  < 2e-16 ***
## s(quarter)7          -2.992e-02  1.693e-03 -17.667  < 2e-16 ***
## s(quarter)8           3.377e-02  1.945e-03  17.367  < 2e-16 ***
## s(quarter)9          -1.922e-02  1.680e-03 -11.441  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

15) Junk topic

```
summary(prep_K18NoG, topics = 15)
```

```
##
## Call:
## estimateEffect(formula = 1:18 ~ party_id + populism + s(quarter),
##     stmobj = mySTM18NoG, metadata = DfmStm$meta, uncertainty = "Global")
##
##
## Topic 15:
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           5.740e-02  8.058e-04  71.234  < 2e-16 ***
## party_idCI           -1.567e-02  1.693e-03  -9.258  < 2e-16 ***
## party_idFDI          -1.539e-03  8.377e-04  -1.837   0.0662 .
## party_idFI            7.205e-03  7.449e-04   9.673  < 2e-16 ***
## party_idINDIPENDENTE -3.751e-02  2.470e-03 -15.190  < 2e-16 ***
## party_idIV           -3.306e-02  2.149e-03 -15.388  < 2e-16 ***
## party_idLEGA         -6.443e-03  6.902e-04  -9.335  < 2e-16 ***
## party_idLEU          -2.191e-02  1.429e-03 -15.327  < 2e-16 ***
## party_idM5S          -2.776e-02  7.007e-04 -39.622  < 2e-16 ***
## party_idMISTO         7.877e-03  8.561e-04   9.201  < 2e-16 ***
## party_idREG_LEAGUES  -2.623e-02  3.243e-03  -8.087 6.16e-16 ***
## populism             -1.634e-04  6.581e-06 -24.826  < 2e-16 ***
## s(quarter)1          -2.375e-02  1.192e-02  -1.991   0.0464 *
## s(quarter)2           3.683e-04  5.368e-03   0.069   0.9453
## s(quarter)3          -1.638e-02  2.833e-03  -5.781 7.43e-09 ***
## s(quarter)4          -6.901e-04  1.924e-03  -0.359   0.7198
## s(quarter)5          -1.569e-02  1.312e-03 -11.964  < 2e-16 ***
## s(quarter)6           2.512e-02  1.459e-03  17.221  < 2e-16 ***
## s(quarter)7           1.577e-02  1.965e-03   8.025 1.02e-15 ***
## s(quarter)8           4.275e-03  2.282e-03   1.873   0.0610 .
## s(quarter)9           2.428e-03  1.638e-03   1.482   0.1382
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

16) Olympics game

```
summary(prep_K18NoG, topics = 16)
```

```
##
## Call:
## estimateEffect(formula = 1:18 ~ party_id + populism + s(quarter),
##     stmobj = mySTM18NoG, metadata = DfmStm$meta, uncertainty = "Global")
##
##
## Topic 16:
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           8.277e-02  8.924e-04  92.750  < 2e-16 ***
## party_idCI            2.621e-02  1.943e-03  13.490  < 2e-16 ***
## party_idFDI          -1.523e-02  9.058e-04 -16.816  < 2e-16 ***
## party_idFI            1.092e-02  7.170e-04  15.233  < 2e-16 ***
## party_idINDIPENDENTE  2.252e-02  3.470e-03   6.489 8.67e-11 ***
## party_idIV           -2.581e-03  2.581e-03  -1.000    0.317
## party_idLEGA         -3.744e-03  6.596e-04  -5.676 1.38e-08 ***
## party_idLEU          -2.523e-02  1.609e-03 -15.682  < 2e-16 ***
## party_idM5S          -2.140e-02  7.299e-04 -29.325  < 2e-16 ***
## party_idMISTO        -2.293e-02  8.383e-04 -27.347  < 2e-16 ***
## party_idREG_LEAGUES  -3.988e-02  3.353e-03 -11.892  < 2e-16 ***
## populism              3.234e-05  7.527e-06   4.296 1.74e-05 ***
## s(quarter)1          -2.641e-01  1.328e-02 -19.895  < 2e-16 ***
## s(quarter)2           6.577e-02  5.705e-03  11.529  < 2e-16 ***
## s(quarter)3          -4.196e-02  2.945e-03 -14.247  < 2e-16 ***
## s(quarter)4          -3.348e-02  1.873e-03 -17.871  < 2e-16 ***
## s(quarter)5          -2.865e-02  1.455e-03 -19.694  < 2e-16 ***
## s(quarter)6           3.938e-02  1.518e-03  25.940  < 2e-16 ***
## s(quarter)7          -5.059e-02  1.768e-03 -28.609  < 2e-16 ***
## s(quarter)8          -1.537e-02  2.054e-03  -7.483 7.26e-14 ***
## s(quarter)9          -1.955e-02  1.884e-03 -10.375  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

17) Right-wing parties topic

```
summary(prep_K18NoG, topics = 17)
```

```
##
## Call:
## estimateEffect(formula = 1:18 ~ party_id + populism + s(quarter),
##     stmobj = mySTM18NoG, metadata = DfmStm$meta, uncertainty = "Global")
##
##
## Topic 17:
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           2.705e-02  8.259e-04  32.748  < 2e-16 ***
## party_idCI           -7.863e-03  1.487e-03  -5.288 1.23e-07 ***
## party_idFDI           9.296e-02  9.393e-04  98.971  < 2e-16 ***
## party_idFI            4.379e-03  6.346e-04   6.900 5.22e-12 ***
## party_idINDIPENDENTE -1.772e-02  2.496e-03  -7.097 1.27e-12 ***
## party_idIV           -9.600e-03  2.179e-03  -4.406 1.05e-05 ***
## party_idLEGA          1.312e-01  7.904e-04 165.955  < 2e-16 ***
## party_idLEU           5.843e-03  1.452e-03   4.025 5.69e-05 ***
## party_idM5S           3.318e-03  6.643e-04   4.995 5.89e-07 ***
## party_idMISTO         1.467e-02  8.167e-04  17.963  < 2e-16 ***
## party_idREG_LEAGUES   5.094e-03  3.334e-03   1.528   0.1265
## populism              1.591e-04  7.535e-06  21.114  < 2e-16 ***
## s(quarter)1          -2.752e-02  1.301e-02  -2.115   0.0344 *
## s(quarter)2           4.126e-03  5.740e-03   0.719   0.4723
## s(quarter)3           2.276e-02  3.068e-03   7.417 1.20e-13 ***
## s(quarter)4          -3.034e-02  1.907e-03 -15.911  < 2e-16 ***
## s(quarter)5           1.244e-03  1.461e-03   0.852   0.3944
## s(quarter)6          -2.160e-02  1.556e-03 -13.885  < 2e-16 ***
## s(quarter)7          -1.925e-02  1.803e-03 -10.681  < 2e-16 ***
## s(quarter)8          -3.253e-02  2.062e-03 -15.774  < 2e-16 ***
## s(quarter)9          -2.683e-02  1.638e-03 -16.382  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

18) Junk topic

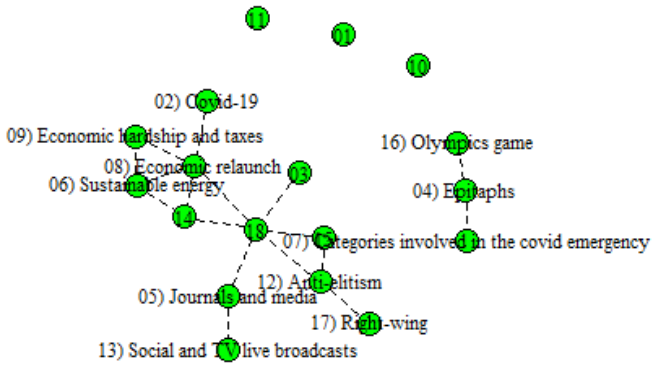
```
summary(prep_K18NoG, topics = 18)
```

```
##
## Call:
## estimateEffect(formula = 1:18 ~ party_id + populism + s(quarter),
##     stmobj = mySTM18NoG, metadata = DfmStm$meta, uncertainty = "Global")
##
##
## Topic 18:
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)           2.362e-02  2.539e-04  93.046  < 2e-16 ***
## party_idCI           -3.229e-03  5.122e-04  -6.304 2.91e-10 ***
## party_idFDI          -1.654e-03  2.592e-04  -6.381 1.76e-10 ***
## party_idFI           -2.110e-03  2.480e-04  -8.511  < 2e-16 ***
## party_idINDIPENDENTE -7.412e-03  8.055e-04  -9.201  < 2e-16 ***
## party_idIV           -1.242e-03  8.335e-04  -1.490 0.136106
## party_idLEGA         -2.574e-03  2.065e-04 -12.464  < 2e-16 ***
## party_idLEU           1.375e-05  5.038e-04   0.027 0.978231
## party_idM5S           1.224e-03  2.341e-04   5.232 1.68e-07 ***
## party_idMISTO        -4.784e-05  2.876e-04  -0.166 0.867882
## party_idREG_LEAGUES  -3.428e-03  1.069e-03  -3.208 0.001338 **
## populism              4.073e-06  2.240e-06   1.818 0.069058 .
## s(quarter)1          -7.050e-03  3.947e-03  -1.786 0.074079 .
## s(quarter)2           4.426e-04  1.658e-03   0.267 0.789521
## s(quarter)3          -1.973e-03  8.858e-04  -2.227 0.025952 *
## s(quarter)4          -1.098e-03  5.652e-04  -1.942 0.052078 .
## s(quarter)5          -1.133e-03  4.854e-04  -2.334 0.019584 *
## s(quarter)6          -1.528e-03  4.293e-04  -3.560 0.000371 ***
## s(quarter)7           1.871e-03  6.132e-04   3.052 0.002275 **
## s(quarter)8          -1.114e-03  6.141e-04  -1.814 0.069718 .
## s(quarter)9           1.857e-03  6.193e-04   2.999 0.002707 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interpretation

Correlation between topics

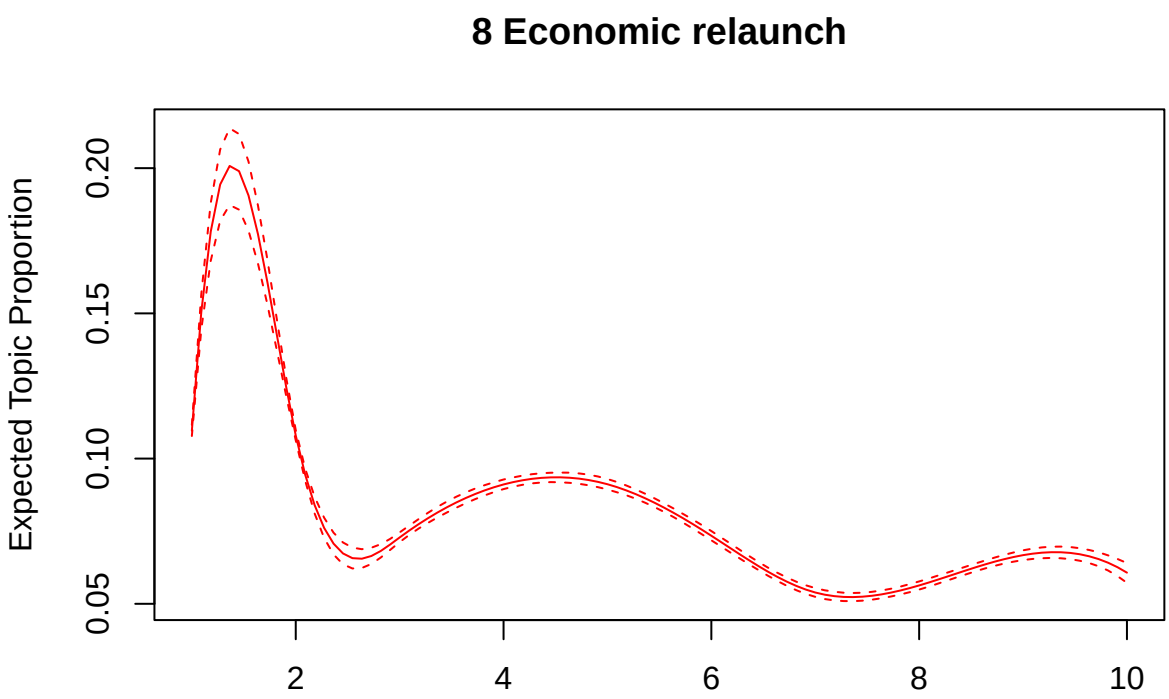
```
mod.out.corr <- topicCorr(mySTM18NoG)

ll <- c("01","02) Covid-19","03",
        "04) Epitaphs",
        "05) Journals and media",
        "06) Sustainable energy",
        "07) Categories involved in the covid emergency",
        "08) Economic relaunch",
        "09) Economic hardship and taxes","10","11",
        "12) Anti-elitism",
        "13) Social and TV live broadcasts","14","15",
        "16) Olympics game",
        "17) Right-wing","18")

set.seed(10)
plot(mod.out.corr, vlabels = ll, layout = layout.auto )
```



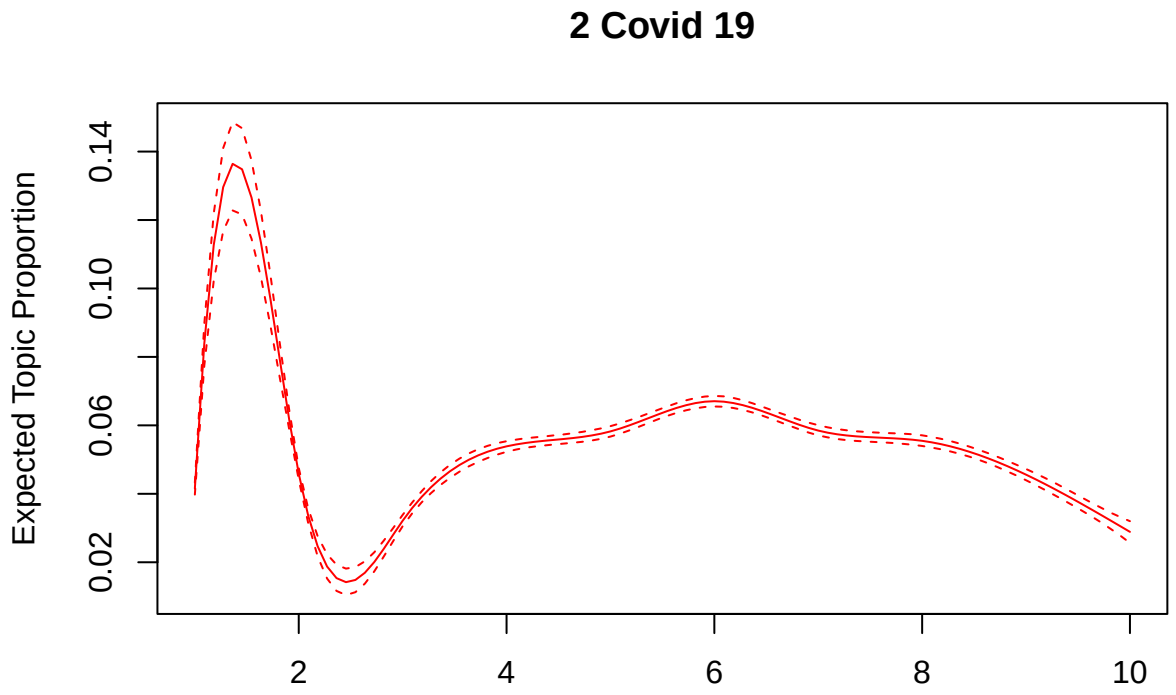
Topic variation over time

Covid cluster

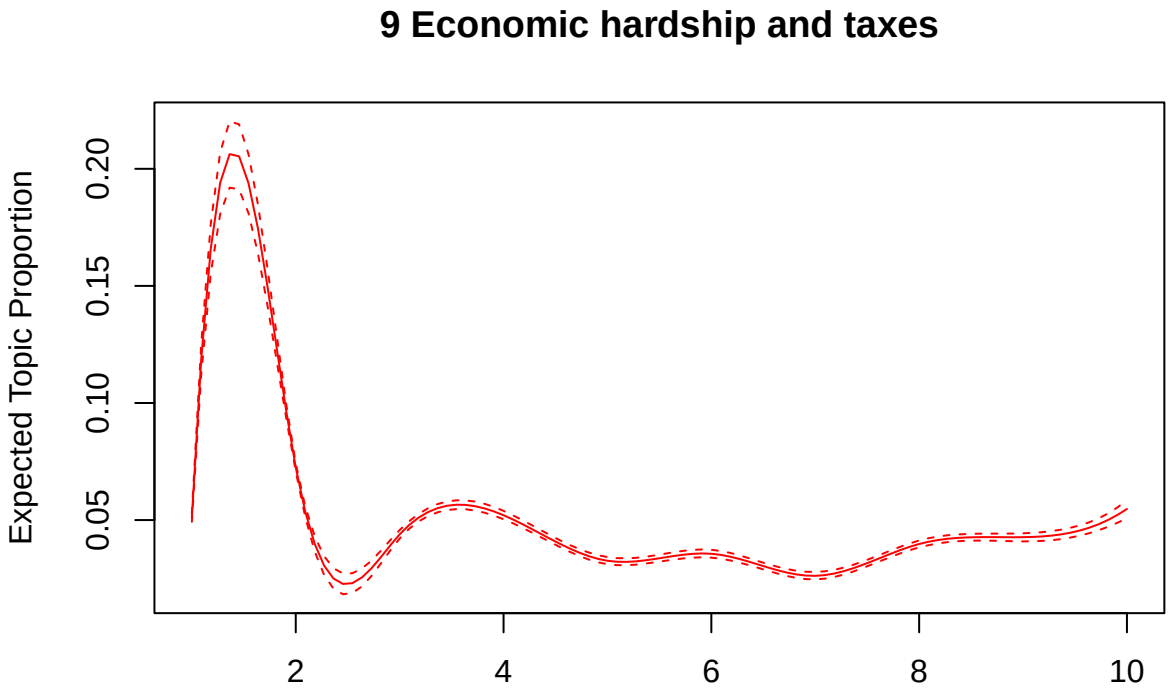
```
# TOPIC 8 Economic relaunch
plot(prep_K18NoG, "quarter", method = "continuous",
     topics = 8, printlegend = F, main = "8 Economic relaunch")
```



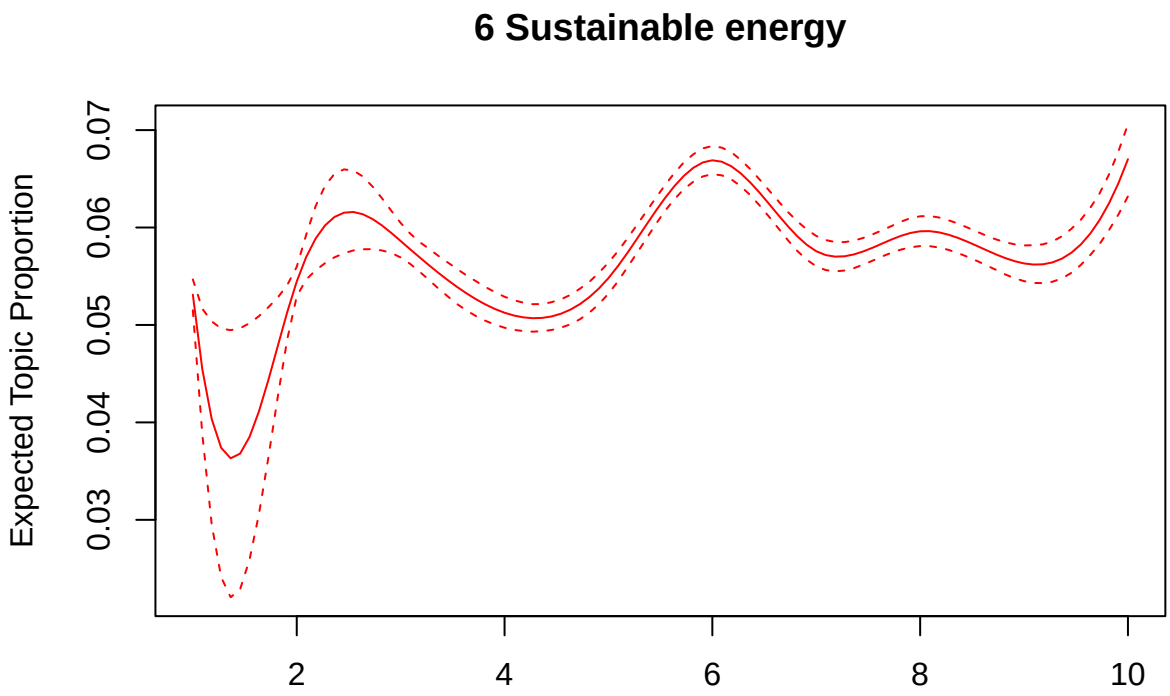
```
# TOPIC 2 Covid 19
plot(prep_K18NoG, "quarter", method = "continuous",
     topics = 2, printlegend = F, main = "2 Covid 19")
```



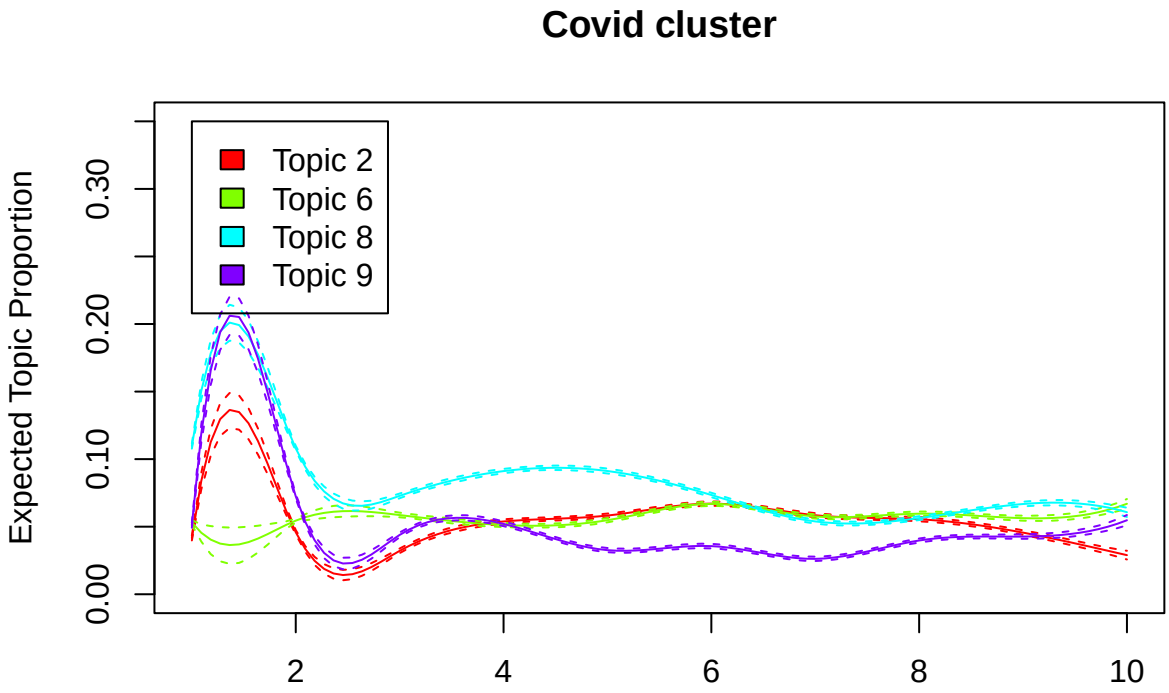

```
# TOPIC 9 Economic hardship and taxes
plot(prep_K18NoG, "quarter", method = "continuous",
     topics = 9, printlegend = F, main = "9 Economic hardship and taxes")
```




```
# TOPIC 6 Sustainable energy
plot(prep_K18NoG, "quarter", method = "continuous",
     topics = 6, printlegend = F, main = "6 Sustainable energy")
```

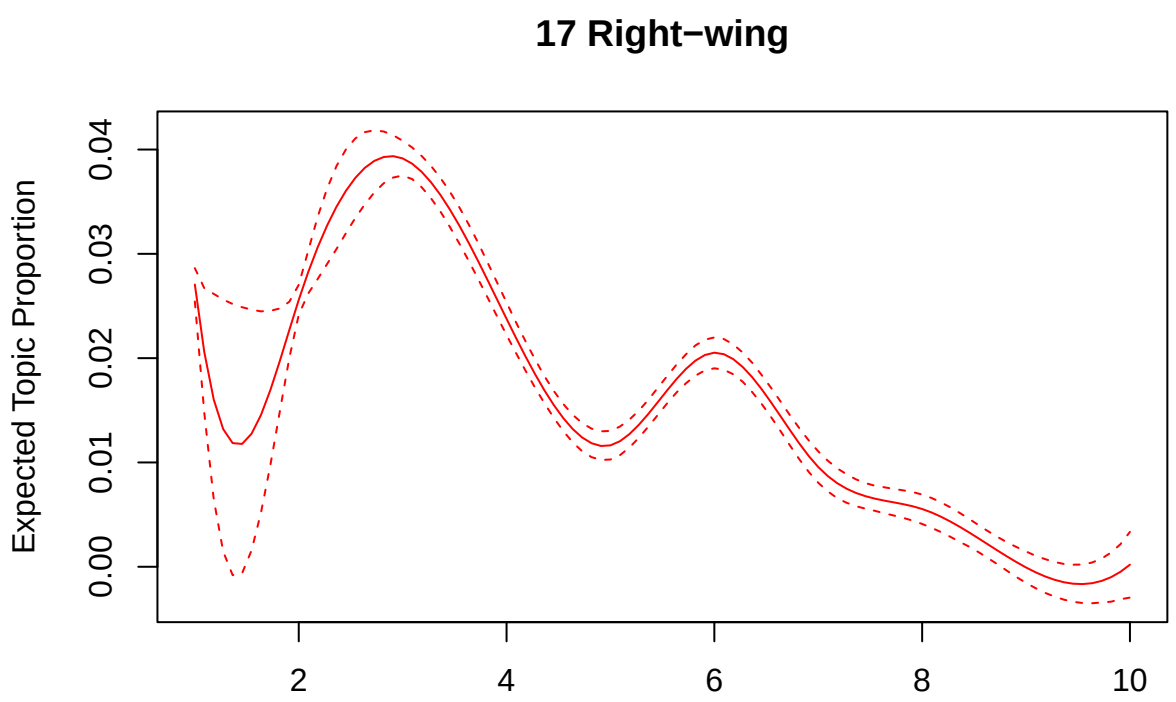



```
# Covid cluster
plot(prep_K18NoG, "quarter", method = "continuous",
     topics = c(2,6,8,9), printlegend = T,
     ylim = c(0,0.35), main = "Covid cluster")
```

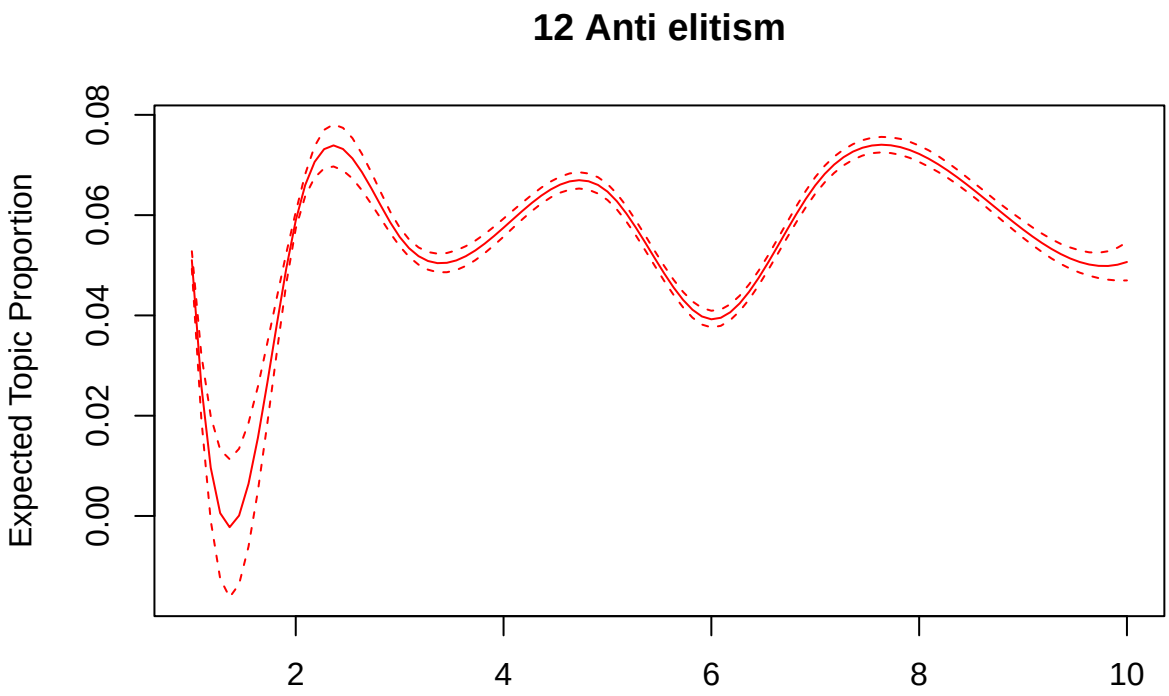


Populism cluster

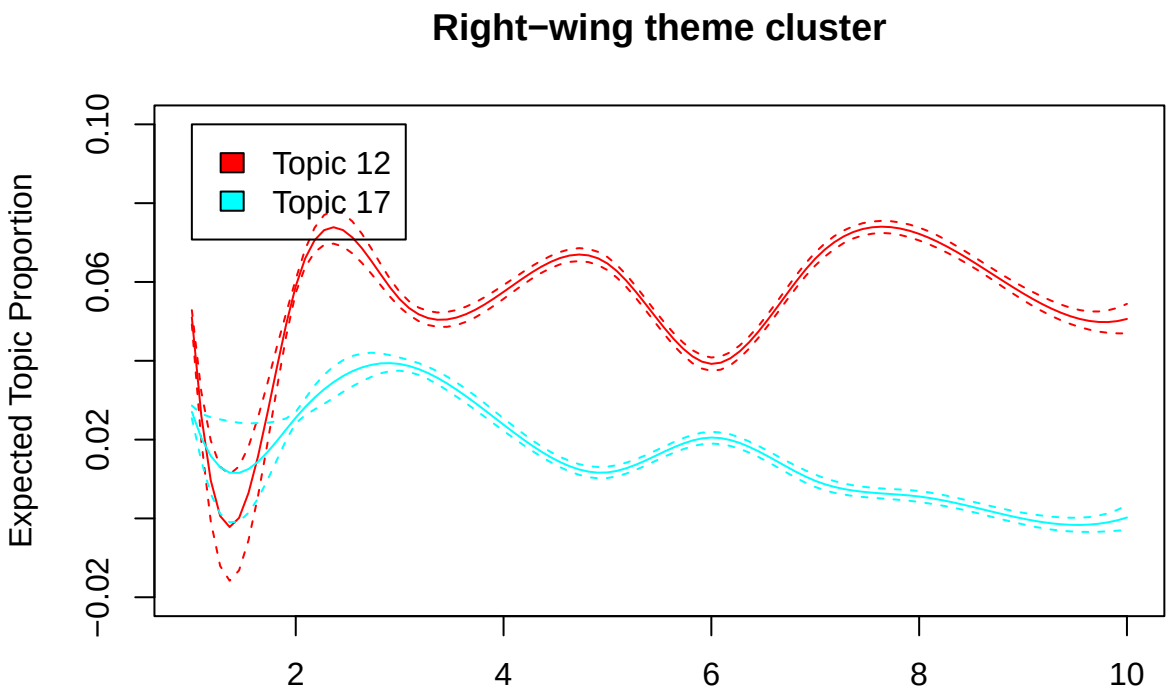
```
# TOPIC 17
plot(prep_K18NoG, "quarter", method = "continuous",
     topics = 17, printlegend = F, main = "17 Right-wing")
```




```
# TOPIC 12
plot(prep_K18NoG, "quarter", method = "continuous",
     topics = 12, printlegend = F, main = "12 Anti elitism")
```

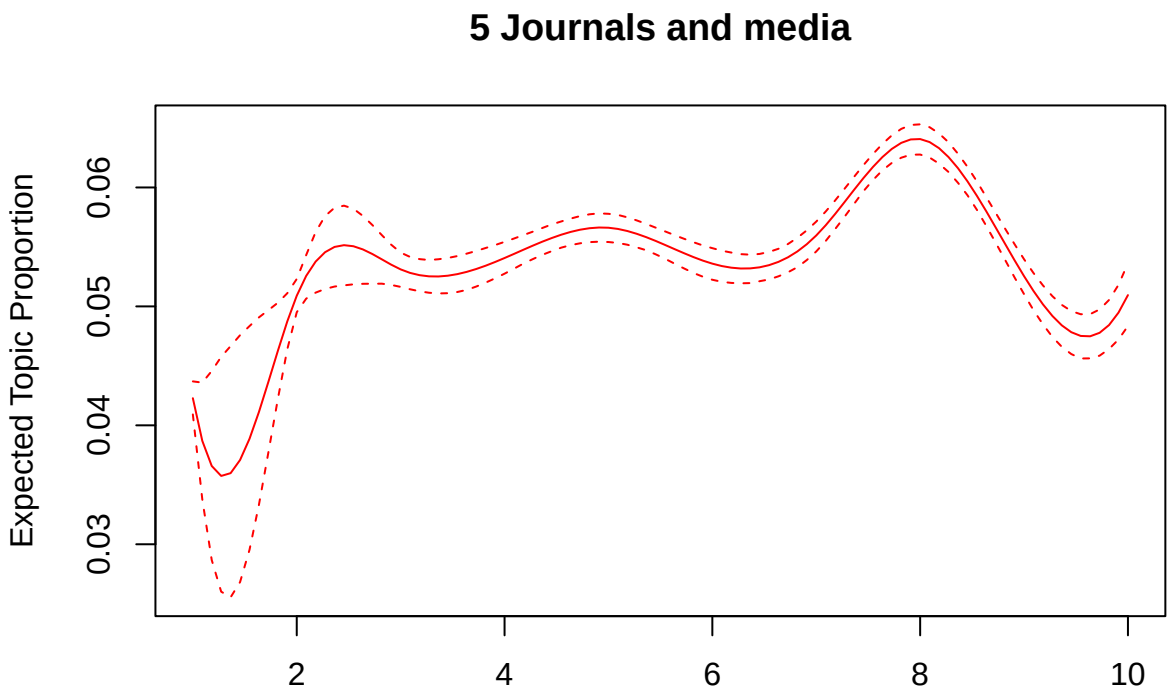



```
# Right-wing theme cluster cluster
plot(prep_K18NoG, "quarter", method = "continuous",
     topics = c(12,17), printlegend = T,
     ylim = c(-0.02,0.1), main = "Right-wing theme cluster")
```

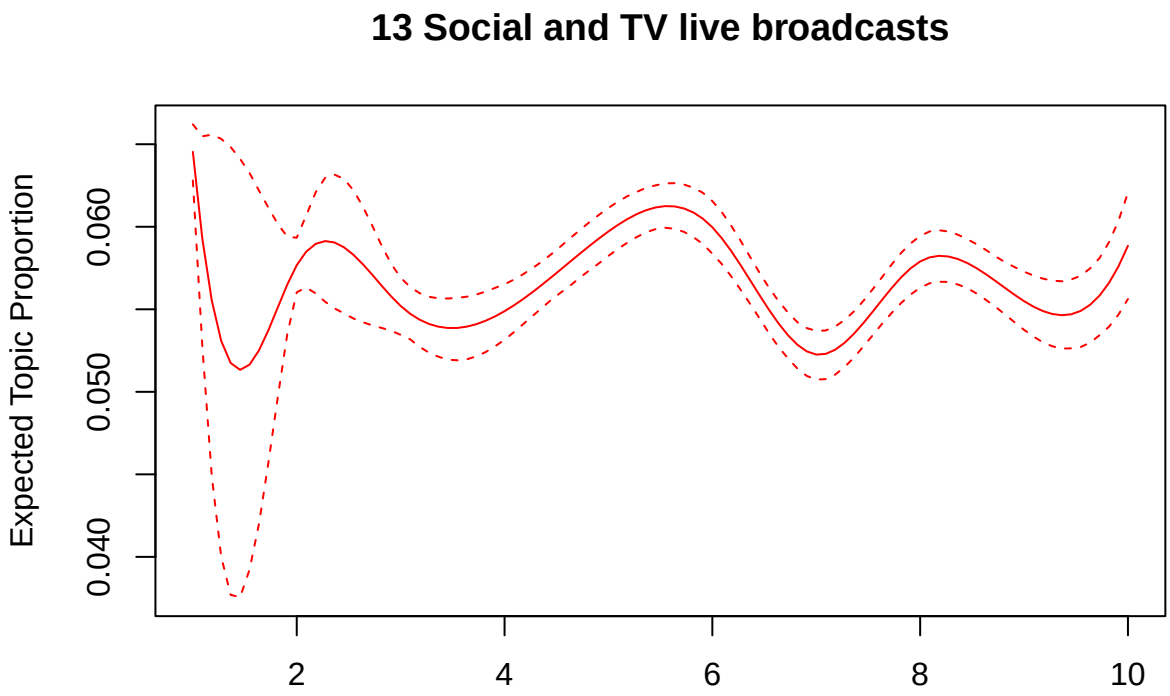


Communication cluster

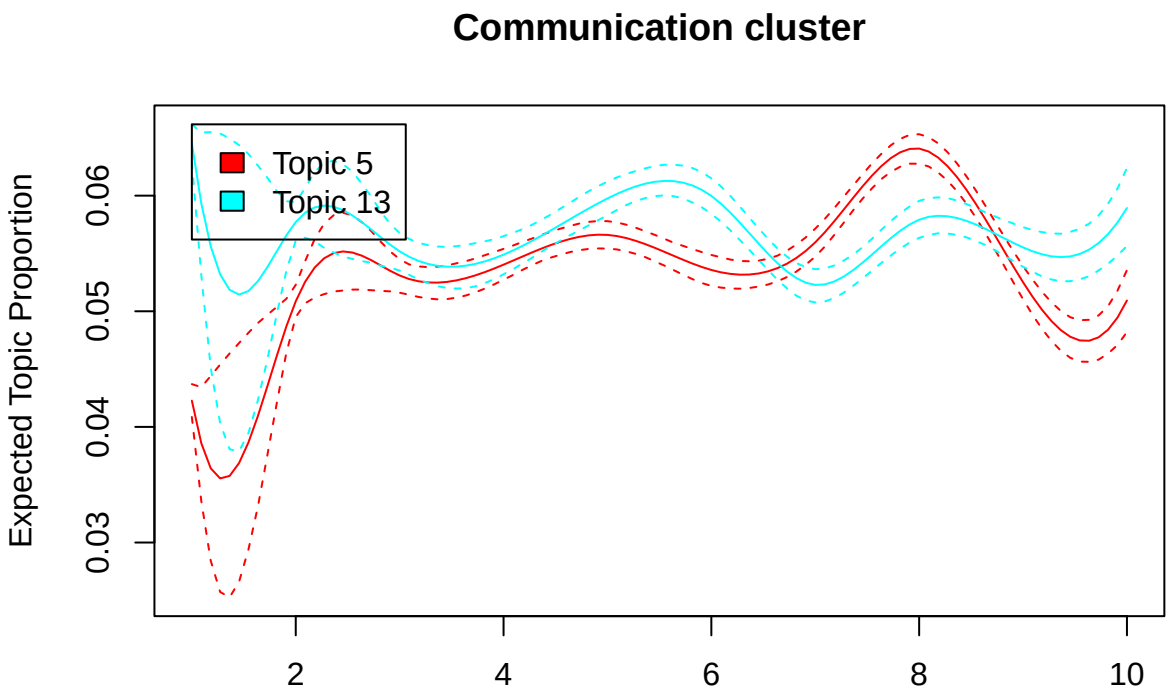
```
# TOPIC 5
plot(prep_K18NoG, "quarter", method = "continuous",
     topics = 5, printlegend = F, main = "5 Journals and media")
```




```
# TOPIC 13
plot(prep_K18NoG, "quarter", method = "continuous",
     topics = 13, printlegend = F, main = "13 Social and TV live broadcasts")
```




```
# Communication cluster
plot(prep_K18NoG, "quarter", method = "continuous",
     topics = c(5,13), printlegend = T, main = "Communication cluster")
```


Communication cluster
Topic 5
Topic 13
Expected Topic Proportion
0.03
0.04
0.05
0.06
2
4
6
8
10