COVID-19 containment measures and the public's response

R. Atkinson, M. Gonzalez, J. Interrante, K. Lee, L.Tian

12-05-2021

Motivation and project objective

The COVID-19 pandemic has elicited a wide array of measures from policymakers and political authorities. Among similar types measures, there is heterogeneity in their stringency or their degree of enforcement.

Also, the public responds differently to the different types of measures.

Objectives of this project

- Group US states according to different strategies of COVID containment.
- Identify variables that are correlated with each group.
- Understand the public's response to these measures.

Analyzing COVID-19 US states' measures

Data and Methodology

COVID measures dataset: Oxford Policy Tracker

- Follows daily levels of 10 types of COVID measures for the 50 US states.
- Each variable is an integer taking values from zero to three or four.
- The higher the value, the higher the restrictions/stringency of the implemented policy.

Our methodology

- Select the period of 30 days before and after the peak number of new deaths,
- Average the level of response, for each type of variable.
- Compare this average across states

Categories of measures and measures included

Table 1: Categories and Variable types

Containment	Health	Economic
School Closings	Public info campaigns	Income Support
Workplace closing	Testing Policy	Debt relief
Cancel public events	Contact Tracing	Fiscal Measures
Restrictions on gatherings	Emergency investment	International Support
Close public transport	Investment in vaccines	
Stay at home requirements	Facial coverings	
Restrictions on internal movement	Vaccination policy	
International travel controls	Protection of elderly people	

Understanding types of responses : PCA results

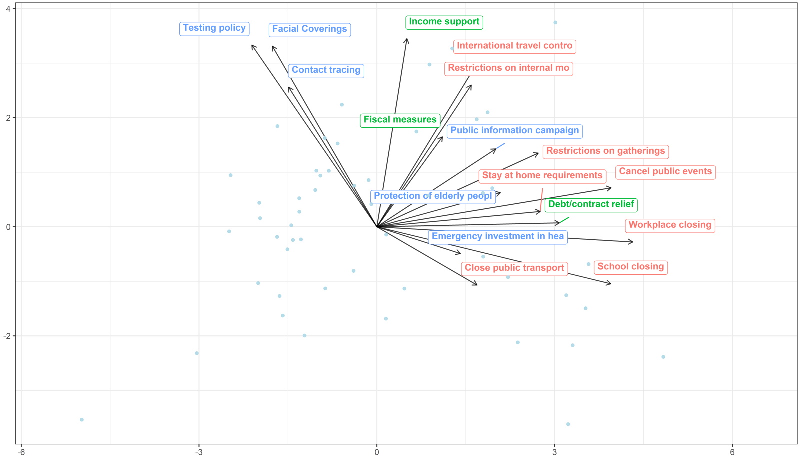


Figure 1: PCA components and original features

Where are US states positioned?

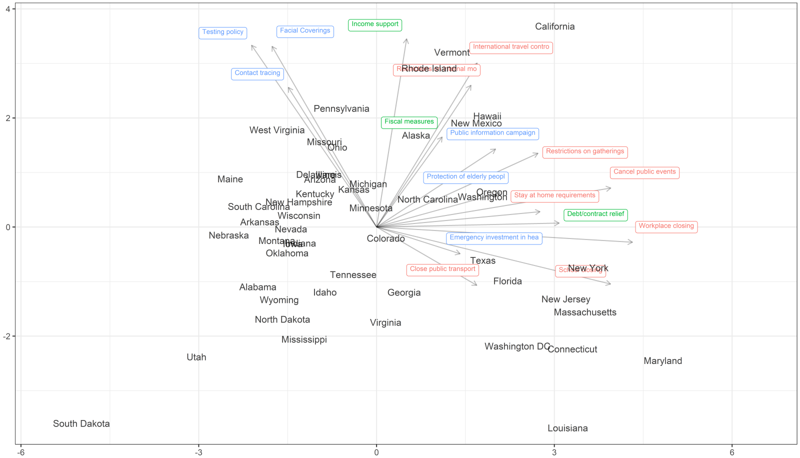


Figure 2: PCA components for USA states

What groups can be identified?

Figure 3: US States clustered by k-means cluster

Is governors' partisanship a good segmenting variable?

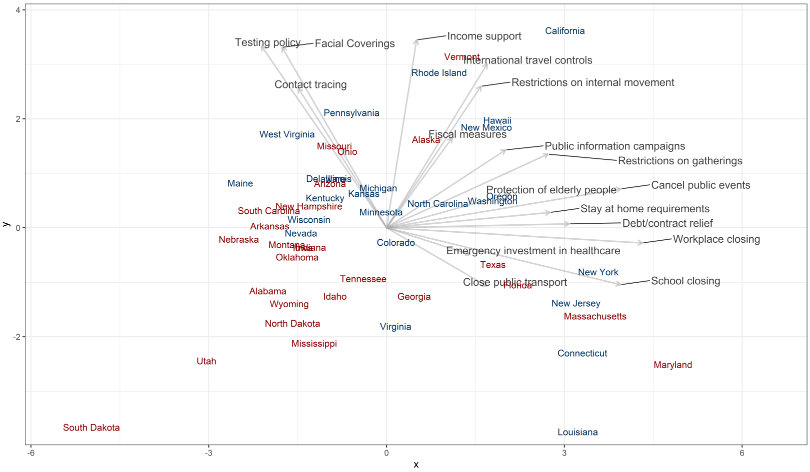


Figure 4: PCA components for USA states. Color is political partisanship of current governor.

Is governors' partisanship a good segmenting variable?

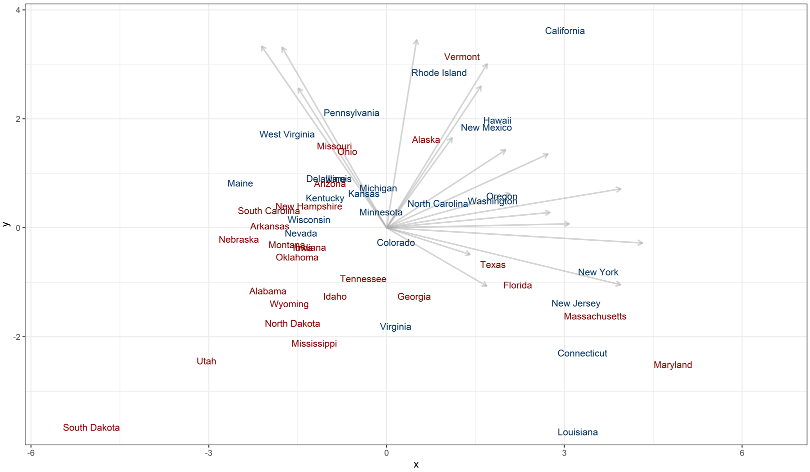


Figure 5: PCA components for USA states. Color is political partisanship of current governor.

Are measures better separated by people's affinities or authorities' partisanship?

- We are interested in investigating wether people's preferences are a better segmenting variable than governor's partizanship.
- We leverage the cases where the governor is of the opposite political side of the majority of the population (e.g. A. Schwarzenegger)
- This leads to the following set of states:

State	Party	Republican Advantage in Population	Leaning
Florida	republican	-5.38	democrat
Kansas	democrat	9.36	republican
Maine	democrat	9.14	republican
Maryland	republican	-22.24	democrat
Massachusetts	republican	-10.87	democrat
Oklahoma	republican	-5.25	democrat
Vermont	republican	-10.29	democrat

In general, segmentation improves

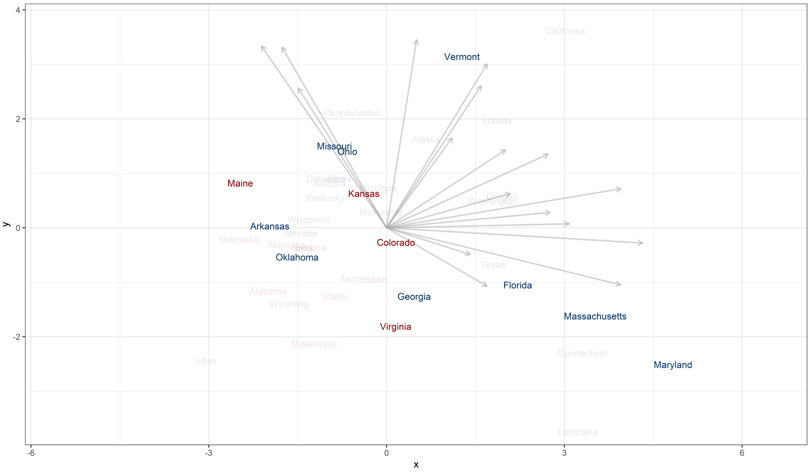


Figure 6: States that change group after considering population's political leaning

Summary

- In general we observe states grouped by the types of measures taken during the peak period of the pandemic.
- Political partizanship of authorities can be a good segmenting variable in the space of measures taken.
- Grouping by the political leaning of the population in general improves the previous segmentation, especially for heavy democrat states like Vermont, Maryland and Massachusets.

Analysis of public's response: tweets

Question of interest and data

Question

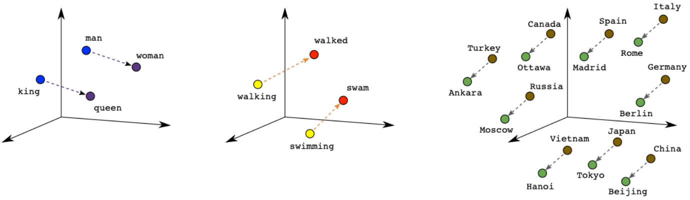
- How different societies, cultures, and nations reacted to the Covid-19 restrictions?

Data

- Preda, Gabriel. “Covid-19 Tweets” Kaggle, 30 Aug. 2020, www.kagglecom/gpreda/covid19-tweets

Methodology

- Word embedding and sentiment analysis
  - `Word2Vec`
  -



Male-Female　　Verb Tense　　Country-Capital

  - sentiment score:

$$\cos(\vec{country}, (\vec{positiveWords} - \vec{negativeWords}))$$

- Text clustering
- Topic modeling
  - Latent Dirichlet Allocation (LDA)
  - Figure out dominant topics for each documents

The Tweets Corpus: Word frequencies

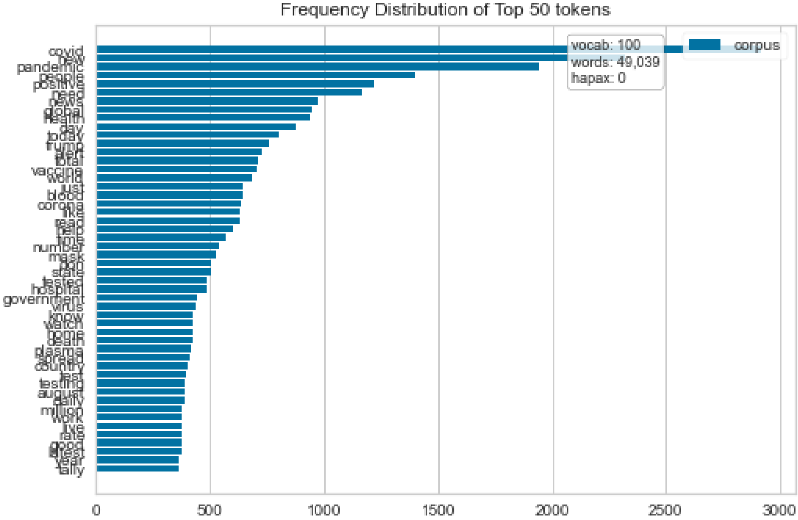


Figure 7: Top words

The Tweets Corpus: lexical dispersion

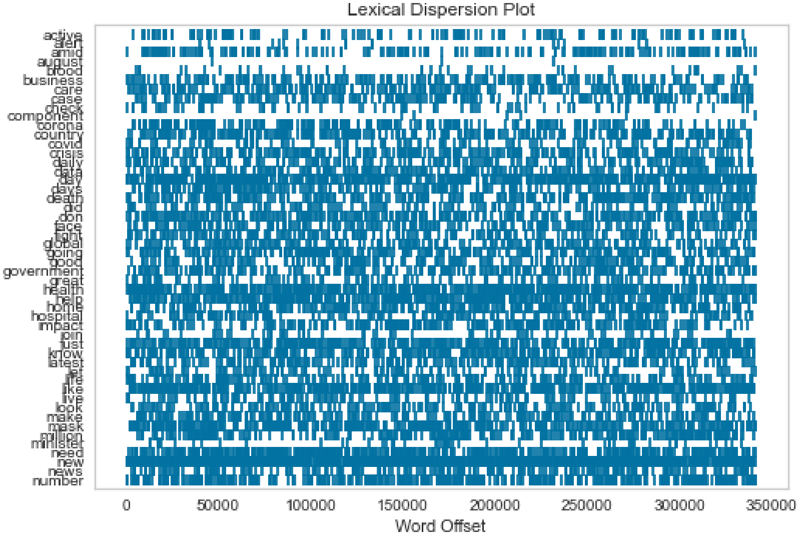


Figure 8: Lexical dispersion

The Tweets Corpus: PCA of TF-IDF

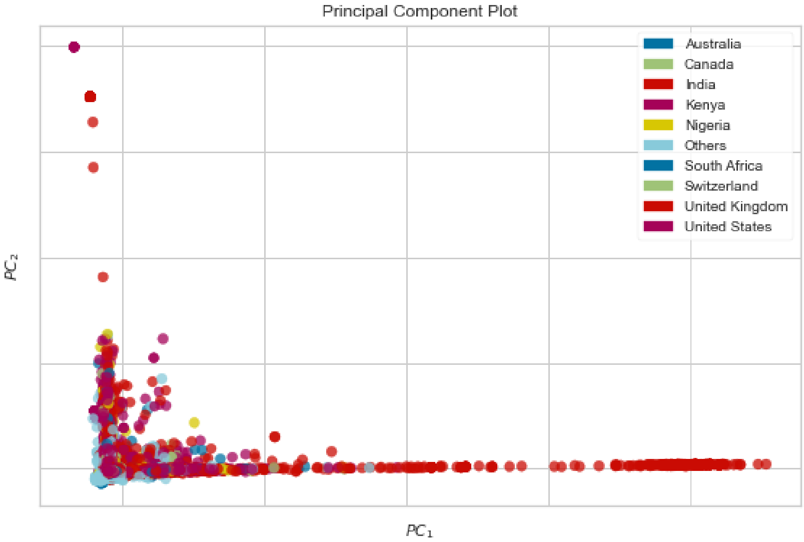


Figure 9: PCA of TF-IDF colored by countries

The Tweets Corpus: TSNE of TF-IDF

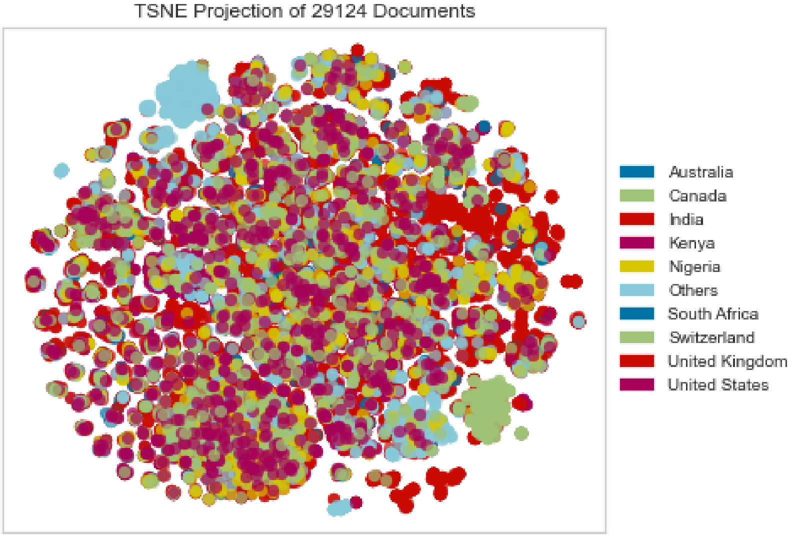


Figure 10: TSNE of TF-IDF colored by countries

Word2Vec

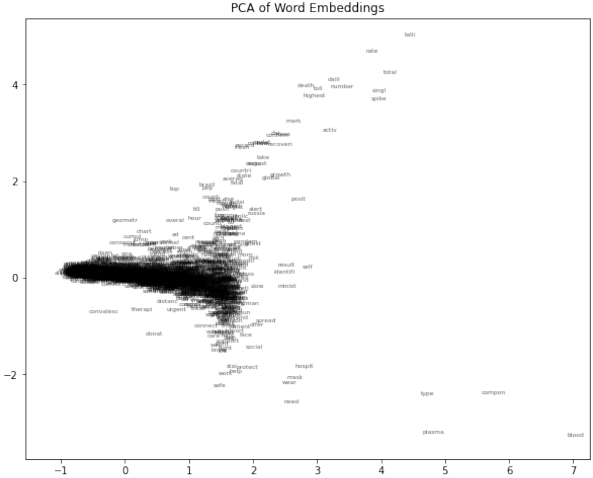


Figure 11: PCA of Word Embeddings

Word2Vec (cont.)

```
>>> w2vmodel.wv.most_similar(
  positive=['wear', 'inject'],
  negative=['vaccine']
)
```

```
'mask'
```

$$\text{wear} - \text{mask} = \text{inject} - \text{vaccine}$$

Public opinion revelation (next-step)

- sentiment score for each country:

$$\cos(\vec{country}, (\vec{positiveWords} - \vec{negativeWords}))$$

Topic modeling

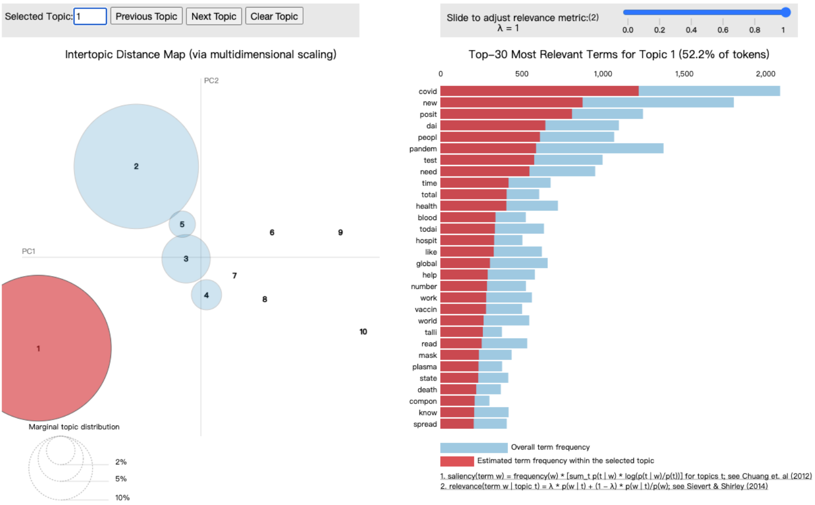


Figure 12: The dominant topic

Topic modeling (cont.)

- Topic are less heterogeneous across countries.
- Topic modeling was not able to reveal meaningful information about public opinion.

Data scarcity

- The corpus only covers a short time span - otherwise dynamic topic modeling can probably reveal shifts in the most concerned topics as the situation evolves, if any.

Next-step plans

- Text clustering: identifying similar countries in terms of public opinion
- Sentiment analysis with Word2Vec
- Word embedding with Autoencoder NN