Replication Data for: Talking to the Populist Radical Right: A Comparative Analysis of Parliamentary Speech (Schwalbach, 2022)

Description

This readme script explains the files that are provided in order to replicate the analyses of the paper. This include four different types of data as well as two separate analyses that are provided.

Data Types

There are four different types of data in the replication folder: R scripts, which contain the code for the respective analyses. Four large .RData files, which contain the speech corpora of the four different countries. Many small .RData files, which store intermediate results of the respective analyses, in order to be able to control them. As well as dictionaries in .csv format for the topics immigration and education and for the respective parties, which are considered in the party mentions analysis.

All R script were produced with R version 4.1.1

Analyses

There are two analyses that are carried out in the main part of the paper: a correspondence analysis and a dictionary analysis of party mentions. All .R and RData files that are only used for one the analyses either have a “CA” or a “party_mentions” in their name. The four speech corpora are used for both analyses. Furthermore, the dictionaries for the detection of speech topics were used prior to the analyses and the hits are already included in the corpora. All files need to be stored in the same location and the R directory needs to be set to this location in order to run the analyses. The respective .R files explain in detail what happens during each step of the analysis.

Correspondence Analysis

For the correspondence analysis, there are three different types of R scripts, which are labeled with the initial numbers 1-3 to indicate the order of execution: The scripts 1a-d produce the raw position scores of the parties that are included in the analysis for each country. The scripts 2a-d merge these scores and annotate them with necessary metadata. Additionally, these are divided into _a and _b, with the _b variant taking the normalised scores for the later analysis. In order to get the scores without preprocessing for the regression analysis, the respective steps need to be left out in the scripts 1a-d. The script 3 then runs the regression analysis for all models and reproduces figure 1 on the basis of the main model.

Party Mentions Analysis

For the party mentions analysis, there are two different types of R scripts. The scripts 1a-d take the dictionaries and the corpora as an input. They check if the parties were mentioned in the respective corpus and annotate the result with metadata from each country. The script 2 then merges the four datasets and runs the logit analysis to estimate the likelihood that a PRRP party is mentioned in a given speech and plots figure 2.