Social Media and Policy Responses to theCOVID-19 Pandemic in Switzerland

Description of Replication Material, Datasets, and R Scripts

Fabrizio Gilardi Theresa Gessler Maël Kubli Stefan Müller

In this file, we describe all datasets and R scripts required to rerun and reproduce the analyses reported in the paper and Supporting Information. In order to have the data, scripts, and outputs saved in the correct folders, which is required to run the scripts successfully, we recommend to download all files as a single ZIP file (click the "Access Dataset" button in the right-hand corner and select "Original Format ZIP"). If you only need selected files, make sure to download data in the original file format. Please do not hesitate contacting the authors if you have any questions.

Datasets

The repository for replicating the empirical analyses contains the following datasets. The following files (in alphabetic order) are used in the R scripts listed below. All other files in the repository, which are not listed here are intermediate products of the analysis and are there to save time if only some steps of the analysis should be reproduced.

- *auto_dictionaries_lsd.RData*: Simple dictionaries for sentiment calculation
- *Facebook_data_minified.RDS*: Minified Corpus of Facebook data (no text)
- *lsde_frenche_germane.RData*: Extended dictionaries for sentiment calculation
- *onetime-structural-shock-irfs-results_all_small_de_only_no_fb.RDS*: Onetime structural shocks of item response function for the VAR.
- *SMD_CDT_data_minified.RDS*: Minified Corpus of Newspaper Articles data (no text)
- *Twitter_data_minified.RDS*: Minified Corpus of Tweets (no text)
- *Twitter_app.*RDS: All tweets of topic app
- *Twitter_mask.RDS*: All tweets of topic maks
- *Twitter_covid.RDS*: All tweets of topic covid
- *var_irfs_main_all_small_de_only_no_fb.RDS*: Main VARs IRFs
- *var_model_all_small_de_only_no_fb.RDS*: Main VAR model

R Scripts

The repository contains the following R scripts. If you want to reproduce the entire analysis, please load and execute the scripts in the following order:

- *01-01-twitter_data_preparation.R*: script to prepare and classify all tweets. This script cannot run since the input data set cannot be added to the replication files. We are unable to provide the tweets for replication, due to the Developer Agreement we have

with Twitter. To run this script if you are in possession of the Tweets you will need the following files:
    - *Twitter_data.RDS (missing dataset)*
    - *auto_dictionaries_lsd.RData*
    - *lsde_frenche_germane.RData*
- *01-02-smd_data_preparation.R*: script to classify all articles from newspapers. This script cannot run since the input data set cannot be added to the replication files. We are unable to provide the newspaper articles for replication, due to a binding data protection agreement with the Swiss Media Database. To run this script (if you are in possession of the Newspaper Articles Dataset) you will need the following files:
    - *SMD_data.RDS (missing dataset)*
    - *auto_dictionaries_lsd.RData*
    - *lsde_frenche_germane.RData*
- *01-03-CdT_data_preparation.R:* script to classify all articles from the newspaper Correire del Ticino and adding these articles to the Swiss Media Database dataset. This script cannot run since the input data set cannot be added to the replication files. We are unable to provide the newspaper articles for replication, due to a binding data protection agreement with the Correire del Ticino. To run this script (if you are in possession of the Correire del Ticino data) you will need the following files:
    - *CdT_data.RDS (missing dataset)*
    - *auto_dictionaries_lsd.RData*
    - *lsde_frenche_germane.RData*
- *01-04-facebook_data_preparation.R*: script to classify all posts from Facebook pages. This script cannot run since the input data set cannot be added to the replication files. We are unable to provide the tweets for replication, due to the Developer Agreement we have with Facebook. To run this script if you are in possession of the Facebook Posts you will need the following files:
    - *Facebook_data.RDS (missing dataset)*
    - *auto_dictionaries_lsd.RData*
    - *lsde_frenche_germane.RData*
- *02-01-data_preparation_var.R:* This script produces the data for the VAR analysis. To run the script, you need the following files:
    - *Twitter_data_minified.RDS*
    - *SMD_CDT_data_minified.RDS*
    - *Facebook_data_minified.RDS*
- *03-01-var_model.R*: This script contains the vector autoregression model used in the paper. To run this script, you need the following files:
    - *main_data_for_paper_de_only.RDS*
- *04-01-figure_1.R*: This script produces Figure 1. To run this script, you need the following files:
    - *Twitter_data_minified.RDS*
    - *SMD_CDT_data_minified.RDS*
    - *Facebook_data_minified.RDS*
    - *ggplot_theme_ddl.R*
- *04-01-figure_2.R*: This script produces Figure 2. To run this script, you need the following files:
    - *Twitter_data_minified.RDS*
    - *SMD_CDT_data_minified.RDS*
    - *Facebook_data_minified.RDS*
    - *ggplot_theme_ddl.R*

- *04-01-figure_3.R*: This script produces Figure 3. To run this script, you need the following files:
    - *onetime-structural-shock-irfs-results_all_small_de_only_no_fb.RDSggplot_theme_ddl.R*
    - *ggplot_theme_ddl.R*
- *0A-01-false_negative_false_positive_sample.R*: This script produces all small data frames used in the appendix and for evaluations. This script will not run since it requires the full data sets from Twitter and the SMD, which are not available. To run this script, you need the following files:
    - *SMD_CDT_data.RDS (missing dataset)*
    - *Twitter_data.RDS (missing dataset)*
    - *Facebook_data.RDS (missing dataset)*
- *0A-02-figure_A1_A2_A3.R*: This script produces Figure A1, A2 and A3. To run this script, you need the following files:
    - *Twitter_app.RDS*
    - *Twitter_mask.RDS*
    - *Twitter_covid.RDS*
    - *ggplot_theme_ddl.R*
- *0A-03-figure_A4.R* This script produces Figure A4. To run this script, you need the following files:
    - *onetime-structural-shock-irfs-results_all_small_de_only_no_fb.RDS*
    - *ggplot_theme_ddl.R*

The beginning of each R script lists the packages that need to be loaded and the package versions that have been used when running the scripts.
The entire analysis has been successfully executed using the following platform and package versions:

```
R version 4.0.2 (2020-06-22)
Platform: x86_64-apple-darwin17.0 (64-bit)
Running under: macOS  10.16

Matrix products: default
LAPACK: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRlapack.dylib

locale:
[1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8

attached base packages:
[1] compiler  parallel  stats     graphics  grDevices utils     datasets  methods
[9] base

other attached packages:
 [1] stargazer_5.2.2    cowplot_1.1.1      magrittr_2.0.1     scales_1.1.1
 [5] graphlayouts_0.7.1 lubridate_1.7.9.2  forcats_0.5.1      purrr_0.3.4
 [9] tibble_3.0.5       tidyverse_1.3.0    stringr_1.4.0      pbmcapply_1.5.0
[13] doParallel_1.0.16  iterators_1.0.13   foreach_1.5.1      data.table_1.14.0
[17] readr_1.4.0        jsonlite_1.7.2     rjson_0.2.20       showtext_0.9-2
[21] showtextdb_3.0     sysfonts_0.8.3     forecast_8.13      ggrepel_0.9.1
[25] ggforce_0.3.3      ggplotify_0.0.5    ggraph_2.0.5       ggplot2_3.3.3
[29] rio_0.5.16         boot_1.3-27        vars_1.5-3         lmtest_0.9-38
[33] urca_1.3-0         strucchange_1.5-2  sandwich_3.0-0     zoo_1.8-8
[37] MASS_7.3-53        dplyr_1.0.5        tidyr_1.1.2

loaded via a namespace (and not attached):
 [1] readxl_1.3.1         backports_1.2.1      plyr_1.8.6           igraph_1.2.6
 [5] splines_4.0.2        digest_0.6.27        htmltools_0.5.1.1    viridis_0.5.1
 [9] checkmate_2.0.0      cluster_2.1.1        openxlsx_4.2.3       recipes_0.1.15
```

```
 [13] modelr_0.1.8          gower_0.2.2           xts_0.12.1            tseries_0.10-48
 [17] jpeg_0.1-8.1          colorspace_2.0-0      rvest_0.3.6           haven_2.3.1
 [21] xfun_0.20             crayon_1.3.4          RCurl_1.98-1.2        glue_1.4.2
 [25] polyclip_1.10-0       gtable_0.3.0          ipred_0.9-9           quantmod_0.4.18
 [29] DBI_1.1.1             Rcpp_1.0.6            viridisLite_0.3.0     htmlTable_2.1.0
 [33] gridGraphics_0.5-1    foreign_0.8-81        stats4_4.0.2          lava_1.6.8.1
 [37] prodlim_2019.11.13    htmlwidgets_1.5.3     httr_1.4.2
RColorBrewer_1.1-2
 [41] ellipsis_0.3.1        pkgconfig_2.0.3       farver_2.0.3          nnet_7.3-14
 [45] dbplyr_2.1.0          tidyselect_1.1.0      rlang_0.4.10          reshape2_1.4.4
 [49] munsell_0.5.0         cellranger_1.1.0      tools_4.0.2           cli_2.3.1
 [53] generics_0.1.0        broom_0.7.3           fs_1.5.0
ModelMetrics_1.2.2.2
 [57] knitr_1.30            tidygraph_1.2.0       zip_2.1.1             nlme_3.1-151
 [61] slam_0.1-48           xml2_1.3.2            rstudioapi_0.13       curl_4.3
 [65] png_0.1-7             reprex_0.3.0          tweenr_1.0.1          stringi_1.5.3
 [69] Matrix_1.3-2          vctrs_0.3.6           pillar_1.4.7          lifecycle_1.0.0
 [73] BiocManager_1.30.10   bitops_1.0-6          R6_2.5.0
latticeExtra_0.6-29
 [77] gridExtra_2.3         codetools_0.2-18      assertthat_0.2.1      withr_2.4.0
 [81] fracdiff_1.5-1        hms_1.0.0             quadprog_1.5-8        grid_4.0.2
 [85] rpart_4.1-15          timeDate_3043.102     class_7.3-17          rvcheck_0.1.8
 [89] TTR_0.24.2            pROC_1.17.0.1         base64enc_0.1-3
```