Instructions for Reproducing the Empirical Analyses

Leader of the Pack?

Changes in 'Wolf Warrior Diplomacy' After A Politburo Collective Study Session

Samuel Brazys

Alexander Dukalskis

Stefan Müller

The China Quarterly

All data and code scripts are stored on Harvard Dataverse at https://doi.org/10.7910/DVN/FR2ZPL

This file contains instructions on how to reproduce the figures and tables reported in the paper and appendix. Note that we provide the full code used to run the LSS models. Due to Twitter's data sharing policy we are not allowed to share the text of each tweet. While the datasets provided in the replication material allow for reproducing our results, they do not allow for re-running the text scaling model. The file data_tweets_en.csv contains a unique identifier (called status_id) for each tweet, which allows users to retrieve the tweets again if they have not been deleted since we downloaded the tweets through the API in 2021. The get_collections() function from the *rtweet* R package should allow you to retrieve the raw data.[1]

Below, we describe the R scripts and do files.

- `01_run_lss.R` runs the LSS model. The main output dataset of this file is data_tweets_en.csv which we use for descriptive plots and as the input for the regression models. The script lists the CRAN versions of each package used when running the models reported in the paper.
- `02_run_regresssions.do` runs all regression models in Stata and prepares datasets to assess the parallel trends assumptions.
- `03_plot_results.R` imports files created in `02_run_regressions.do` and creates all plots. The script lists the CRAN versions of each package used when running the models reported in the paper.

If you have any questions, please do not hesitate to contact the authors of the paper.

[1] `https://cran.r-project.org/web/packages/rtweet/rtweet.pdf`