Description of Replication Material, Datasets, and R Scripts

The Incumbency Advantage in Lower-Order PR Elections: Evidence from the Irish Context, 1942–2019

Michael Jankowski 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. If you have any questions, please do not hesitate to contact the authors.

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.

- `data_dates_byeelections.csv`: candidates elected in all bye-elections to the Dáil Eireann
- `data_dates_general_elections.csv`: dataset with dates of Irish general elections
- `data_dates_local_elections.csv`: dataset with dates of Irish local elections
- `data_elections_general_raw.rds`: scraped dataset of candidate-level election results in Irish general elections (Source: http://irelandelection.com)
- `data_elections_local_raw.rds`: scraped dataset of candidate-level election results in Irish local elections (Source. http://irelandelection.com)
- `data_general_elections_complete.rds`: merged and harmonised dataset for Irish general elections. This file is created in 01b_prepare_rdd_data_generalelections.R
- `data_local_elections_complete.rds`: merged and harmonised dataset for Irish local elections. This file is created in 01b_prepare_rdd_data_local_elections.R
- `data_local_elections_firstcount.rds`: dataset indicating the availability of results from first counts in Irish local elections. In the analysis, we use the last count in each constituency/local electoral area. This dataset is used to compare the data availability regarding the first count and last count
- `data_redmond_regan_replication.xlsx`: replication dataset from Redmond and Regan (2015). We extend this dataset by adding the results from the most recent general elections
- `data_tds.csv`: spreadsheet with names and party affiliations of all elected TDs in each legislative period
- `data_turnout_ireland.csv`: levels of aggregated turnout in Irish local and general elections

- `iebabynames.rda`: Full baby name data (1964–2019) for the Republic of Ireland from the Central Statistics Office. Due to confidentiality reasons, only names with 3 or more instances in the relevant year are included[1]
- `data_ines_subset.rds`: A subset of variables from the Irish National Election Study (INES), a five-wave panel survey of (initially) 2663 respondents, carried out through the period 2002-2007 and encompassing the Irish general elections of 2002 and 2007, as well as the local and European Parliament elections of 2004. This dataset only contains the year, respondent ID, and response to the question as to whether at least one candidate called the respondent's home.[2]

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. You can use an R project[3], use the **here** package or set the working directory manually:

- *01a_prepare_rdd_data_local_elections.R*: script to clean the results from local elections and code which local candidate was also a TD at the time of the local election. The script produces `data_local_elections_complete.rds`. To run the script, you need the following files (in order of appearance in the script):
  - `data_elections_local_raw.rds`
  - `data_tds.csv`
  - `data_dates_general_elections.csv`
  - `data_dates_byeelections.csv`
  - `data_dates_local_elections.csv`
  - `iebabynames.rda`
- *01b_prepare_rdd_data_general_elections.R*: script to clean the results from general elections and merge them with the replication data from Redmond and Regan (2015). The script produces `data_general_elections_complete.rds`. To run the script, you need the following files (in order of appearance in the script):
  - `data_elections_general_raw.rds`
  - `data_redmond_regan_replication.xlsx`
- *02_rdd_analysis*: This script includes empirical analyses reported in the paper and the appendix that relate to the Regression Discontinuity Design. To run the script, you need the following files (in order of appearance in the script):
  - `data_local_elections_complete.rds`
  - `data_general_elections_complete.rds`
- *03_descriptive_plots.R*: This script produces all descriptive plots reported in the paper (Figure 1) and in SI Section A and SI Section C. To run the script, you need the following files (in order of appearance in the script):
  - `data_local_elections_complete.rds`
  - `data_turnout_ireland.csv`
  - `data_local_elections_firstcount.rds`
  - `data_ines_subset.rds`

[1] https://www.cso.ie/en/interactivezone/visualisationtools/babynamesofireland/. Müller, Stefan (2020). iebabynames: Full baby name data for the Republic of Ireland. R package version 0.1.1. http://github.com/stefan-mueller/iebabynames

[2] Source and description: https://www.ucd.ie/issda/data/irishnationalelectionstudy/

[3] https://support.rstudio.com/hc/en-us/articles/200526207-Using-Projects

- *helper_functions.R*: This script includes R functions for various plots and tables required for carrying out the RDD analysis.

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:

```
> sessionInfo()
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_IE.UTF-8/en_IE.UTF-8/en_IE.UTF-8/C/en_IE.UTF-8/en_IE.UTF-8

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

other attached packages:
 [1] foreign_0.8-80    forcats_0.5.0     scales_1.1.1      Hmisc_4.4-1
lattice_0.20-41
 [6] purrr_0.3.4       kableExtra_1.2.1  rddensity_2.1     gridExtra_2.3
rdrobust_0.99.9
[11] rdd_0.57          Formula_1.2-3     AER_1.2-9         survival_3.1-12
lmtest_0.9-38
[16] zoo_1.8-8         sandwich_3.0-0    ggplot2_3.3.3     tidyr_1.1.2
rio_0.5.16
[21] iebabynames_0.1.1 quanteda_2.1.2    stringr_1.4.0     car_3.0-9
carData_3.0-4
[26] dplyr_1.0.2

loaded via a namespace (and not attached):
 [1] httr_1.4.2          viridisLite_0.3.0   splines_4.0.2       RcppParallel_5.0.2
 [5] latticeExtra_0.6-29 cellranger_1.1.0    yaml_2.2.1          backports_1.2.1
 [9] pillar_1.4.7        glue_1.4.2          digest_0.6.27       checkmate_2.0.0
[13] RColorBrewer_1.1-2  rvest_0.3.6         colorspace_2.0-0    htmltools_0.5.0
[17] Matrix_1.2-18       pkgconfig_2.0.3     haven_2.3.1         webshot_0.5.2
[21] jpeg_0.1-8.1        openxlsx_4.1.5      htmlTable_2.1.0     tibble_3.0.4
[25] generics_0.1.0      usethis_2.0.0       ellipsis_0.3.1      withr_2.3.0
[29] nnet_7.3-14         magrittr_2.0.1      crayon_1.3.4        readxl_1.3.1
[33] evaluate_0.14       stopwords_2.1       fs_1.5.0            MASS_7.3-51.6
[37] xml2_1.3.2          tools_4.0.2         data.table_1.13.6   hms_0.5.3
[41] lifecycle_0.2.0     munsell_0.5.0       cluster_2.1.0       zip_2.1.1
[45] compiler_4.0.2      rlang_0.4.10        grid_4.0.2          rstudioapi_0.13
[49] htmlwidgets_1.5.1   base64enc_0.1-3     rmarkdown_2.6       gtable_0.3.0
[53] abind_1.4-5         curl_4.3            R6_2.5.0            knitr_1.30
[57] fastmatch_1.1-0     lpdensity_2.1       stringi_1.5.3       Rcpp_1.0.5
[61] rpart_4.1-15        png_0.1-7           vctrs_0.3.6         tidyselect_1.1.0
[65] xfun_0.19
```