README for
No Surprises, Please: Voting Costs and Electoral Turnout
by Jean-Victor Alipour and Valentin Lindlacher
Journal of Political Economy Microeconomics, MS 230259

1. Overview

This readme file contains instructions on how to replicate the results of the article *No Surprises, Please: Voting Costs and Electoral Turnout*. The `_master.do` file runs all of the code to construct the data from scratch, perform the analysis, and produce the output using Stata. The replicator should expect the code to run for about 4 hours.

The replication folder is organized as follows:

- The `_master.do` file defines directories and calls do-files in the correct order
- `01_code/` includes all codes necessary to clean the data and produce the results.
  The codes that read and geolocate addresses are stored in subdirectory `1_prep_and_geolocate_addresses/`;
  the codes that read, clean, and build the datasets for estimation are stored in subdirectory `2_build_datasets/`;
  the codes that conduct the analysis and produce the output are stored in subdirectory `3_analysis/`
- `02_rawdata/` stores the raw data, organized in subdirectories according to the data source
- `03_geocoded_addresses/` stores pre-processed geo-coordinates of residential addresses and polling locations; you can perform the geo-location yourself by providing your opencagegeo API key in the `_master.do` file (see details below)
- `04_newdata/` saves datasets compiled from the raw data, which are used for the final analysis
- `05_packages/` contains ado files, the authors' Stata programs, and other packages needed to make the codes run
- `06_arc_analyze/` stores preprocessed output from GIS calculations
- `07_output/` is where the log files (`07_output/logfiles/`), figures (`07_output/figures/`) and tables (`07_output/tables/`) are saved
- `tmp/` is where auxiliary files produced during the data cleaning process are saved

2. Instructions to replicators

- Replicators need first to open the `_master.do` file and update the global path to the replication folder

- Running the _master.do replicates the project from scratch by calling do-files from the subdirectories of 01_code/ in the correct order
- We provide geocoded voter and polling place addresses in 03_geocoded_addresses/. If replicators want to perform their own geocoding, they need to enter a valid OpenCage Geocoding API key (https://opencagedata.com/) and change the globals RERUN_GEOCODE_VOTER and RERUN_GEOCODE_PP to 1 in _master.do

3. Data

All data used in this project are publicly available and provided in the replication package. The following list details the name, content, and source of the raw data files.

- 02_rawdata/election_office/electoral_rolls/LTW2018_Straßenverzeichnis_einzeln_mit Filter.xlsx
  - Contains all addresses of eligible voters in the 2018 State Election
  - Source: Munich Electoral Office, Heiko Flehmig (heiko.flehmig@muenchen.de)
  - Provided: Yes
- 02_rawdata/election_office/electoral_rolls/X_Straßenverzeichnis.xlsx
  - Contains the correspondence of eligible voters' addresses to their precinct in each election X ∈ {BTW2009, BTW2017, EuW2014, EuW2019, KommW2014, KommW2020, LTW2013, LTW2018}
  - Source: Munich Electoral Office, Heiko Flehmig (heiko.flehmig@muenchen.de)
  - Provided: Yes
- 02_rawdata/election_office/electoral_rolls/BTW2009-Strvz-num.csv
  - Contains the address and the assigned precinct of polling places used in the 2009 Federal Election
  - Source: Munich Electoral Office, Heiko Flehmig (heiko.flehmig@muenchen.de)
  - Provided: Yes
- 02_rawdata/election_office/polling_places/X_Wahllokalverzeichnis.xlsx
  - Contains the address and the assigned precinct of polling places used in election X ∈ {BTW_2017, BTW_2013, EuW_2019, KommW_2014, KommW_2020, LTW_2018}
  - Source: Munich Electoral Office, Heiko Flehmig (heiko.flehmig@muenchen.de)
  - Provided: Yes
- 02_rawdata/election_office/X/X_strukturindikatoren.xlsx
  - Contains structural indicators at precinct level for each election X ∈ {btw_2013, btw_2017, euw_2014, euw_2019, kommw_2014, kommw_2020, ltw_2013, ltw_2018}
  - Source: Munich Electoral Office, Heiko Flehmig (heiko.flehmig@muenchen.de)
  - Provided: Yes
- 02_rawdata/election_office/X/X_amtlich_ergebnisdownload.xlsx
  - Contains election results for each election X ∈ {btw_2013, btw_2017, euw_2014, euw_2019, kommw_2014, kommw_2020, ltw_2013, ltw_2018}
  - Source: Munich Electoral Office, Heiko Flehmig (heiko.flehmig@muenchen.de)

  - Provided: Yes
- 02_rawdata/statistical_office/Umrechnung_Wahlen_2013-2020_auf_LTW2018.xlsx
  - Contains conversion keys to harmonize precinct-level variables (outcomes and controls) from all elections between 2013 to 2020 to the precincts' boundaries of the 2018 State Election.
  - Source: Munich Statistical Office, Ingrid Kreuzmair (ingrid.kreuzmair@muenchen.de)
  - Provided: Yes
- 02_rawdata/statistical_office/migration_precincts.xlsx
  - Contains migration counts within and across precinct boundaries (already with boundaries as of 2018)
  - Source: Munich Statistical Office, Ingrid Kreuzmair (ingrid.kreuzmair@muenchen.de)
  - Provided: Yes
- 02_rawdata/shp/
  - Subdirectory contains shapefiles of precinct boundaries (by election) and district boundaries (time-invariant).
  - Source: Munich Statistical Office, Ingrid Kreuzmair (ingrid.kreuzmair@muenchen.de)
  - Provided: Yes
- 02_rawdata/RWI/GEO_RED_v3/Stata/WM_SUF_ohneText/WM_SUF_ohneTextX.dta
  - Contains data on quoted rents for flats (X between 1 and 8 for data chunks) by listings with an identifier of the location in a 1 km^2 grid (INSPIRE ID).
  - We cannot provide the SUF but only the data aggregated to the precinct level
  - Source: RWI Leibniz Institute for Economic Research (RWI-GEO-RED): Scientific Use File (SUF) is available free of charge, but a data application is required. Data access: https://www.rwi-essen.de/en/research-advice/further/research-data-center-ruhr-fdz/data-access
  - Provided: No
- 02_rawdata/RWI/02_data/final/election_rwi_red_sb_final.dta
  - Contains data on the quoted rents for flats aggregated to the precinct level for each election year. This is the data used in the estimations. We provide this data as the disaggregated data has to be requested from RWI. This file can be produced using the code provided with the SUF (see point above) from RWI.
  - Source: RWI (and own calculations, see section Section 5 for details)
  - Provided: Yes
- 02_rawdata/RWI/02_data/tmp/precinct_grid_inter_area.shp
  - Contains a shapefile of the intersection of the INSPIRE grid ID with the precinct shapefile of the 2018 State Election. This can also be generated with the provided shapefile and the raster shapefile from RWI.
  - Source: Own calculations (code provided as well)
  - Provided: Yes
- 02_rawdata/CoronaDatenplattform/corplat_kreise_yrl.dta
  - Contains characteristics of German counties

  - Source: Federal Statistical Office (Destatis) and Federal Employment Agency, accessed via BMWK-Corona Datenplattform (https://www.healthcare-datenplattform.de/pages/projekt, accessed 11/2021)
  - Provided: Yes
- 02_rawdata/IDEA/
  - Folder contains statistics on the prevalence of early voting and postal voting arrangements
  - Source: International Institute for Democracy and Electoral Assistance (International IDEA), available for download at: https://www.idea.int/data-tools/tools/special-voting-arrangements/early-voting [accessed 11/2023]
  - Note: for better machine-readability, the .csv files were cleaned by hand and saved with the prefix cleaned_
  - Provided: Yes
- 02_rawdata/schuloffensive/Wahllokale_merge.xlsx
  - A spreadsheet containing information about whether a school that served as a polling place in at least one election underwent renovation due to the city's school construction program (*Schulbauoffensive*).
  - Source: Own research (information gathered from https://risi.muenchen.de/risi/)
  - Provided: Yes
- 06_arc_anlyze/arc_output/ltw18_sb_intersected.shp
  - Contains a shapefile to identify neighboring precincts for the local matching approach.
  - Source: Own calculations (code provided as well)
  - Provided: Yes
- 03_geocoded_addresses/voter_addresses_ltw18_geocoded.dta
  - Contains geo-referenced addresses of eligible voters in the 2018 State Election (can also be generated).
  - Source: Own calculations using opencagegeo module (https://opencagedata.com/) with data provided by Munich Election Office
  - Provided: Yes
- 03_geocoded_addresses/wl_addresses_geocoded.dta
  - Contains geo-referenced addresses of all polling places (can also be generated).
  - Source: Own calculations using opencagegeo module (https://opencagedata.com/) with data provided by Munich Election Office
  - Provided: Yes

4. Computational requirements

The replication can be performed using Stata only. The code was tested using Stata 17. Necessary ado files are provided in 05_packages/ado/ and listed here:

- osrmtime.ado (Version 1.3.2)
- missings.ado (Version 1.4.0)

- gduplicates.ado (Version 1.0.0)
- gcollapse.ado (Version 1.3.1)
- opencagegeo.ado (Version 1.2.0)
- distinct.ado (Version 1.2.1)
- labmask.ado (Version 1.0.0)
- heatplot.ado (Version 1.1.1)
- spmap.ado (Version 1.3.2)
- grc1leg.ado (Version 1.0.5)
- encoder.ado
- fre.ado (Version 1.2.5)
- eststo.ado (Version 1.1.0)
- esttab.ado (Version 2.0.6)
- estpost.ado (Version 1.1.6)
- outreg.ado (Version 4.32)
- reghdfe.ado (Version 6.12.3)
- coefplot.ado (Version 1.8.6)
- event_plot.ado
- zscore.ado (Version 2.0.1)
- mkcorr.ado (Version 2.0.1)
- title.ado (Version 2.0)
- did_imputation.ado
- did_multiplegt.ado
- eventstudyinteract.ado (Version 0.1)
- csdid.ado (Version 1.72)
- texsave.ado (Version 1.6.1)
- honestdid.ado (Version 1.1.2)
- grc1leg2.ado (Version 2.26)
- savesome.ado (Version 1.1.0)
- ebalance.ado (Version 1.5.4)
- grconst.ado (Version 1.0)
- psmatch2.ado (Version 4.0.12)
- did_multiplegt_dyn.ado

To calculate walking distances, we used the module `osrmtime`, available for download on Github (https://github.com/christophrust/osrmtime). We use the package version 1.3.3, which is stored in `05_packages/osrmtime_release1.3.3/`. The module draws on OSRM basemaps, which were downloaded in 2020 from http://download.geofabrik.de/europe/germany/bayern.html (download map "Oberbayern", link last tested in Feb 2024), and stored in `05_packages/osrm_basemaps/`.

Two shapefiles included in this replication package were generated using ArcGIS (ArcMap (Version 10.4.1) and IDLE Python 2.7.10 Shell with `arcpy` and the extension `spatial`) and R (R-4.2 with the

libraries sf (1.0-12) and dplyr (1.1.2) (stored in 05_packages/library)). The R library folder is only needed if you want to replicate the provided shapefiles.

We used a Microsoft Corporation Virtual Machine with the following features:

- Operating system version: Microsoft Windows Server 2019 Standard
- Processors: Intel(R) Xeon(R) CPU E5-2697 v4 @ 2.30GHz, Intel(R) Xeon(R) CPU E5-2697 v4 @ 2.30GHz
- Installed memory (RAM): 128 GB
- Total disk space: 994.91 GB

With our setting, it takes about 4 hours to run the code (without geocoding). If the replicator runs their own geocoding, this takes at least an additional hour (depending on the purchased plan).

5. Details

- **RWI data.** The precinct-level data on quoted rents provided in this replication package (02_rawdata/RWI/02_data/final/election_rwi_red_sb_final.dta) has been aggregated from grid-level data obtained by RWI (RWI-GEO-RED). If you want to perform the aggregation yourself, you need to request the Scientific Use File (see above) and follow the following steps:
  - Store the SUF in 02_rawdata/RWI/02_data/raw/RWI/
  - Then run the do file 02_rawdata/RWI/01_code/1_prep_red.do to perform the aggregation
  - 1_prep_red.do loads a shapefile that contains the intersection of precincts and the identifier from RWI (INSPIRE grid ID). If you want to generate this yourself, you have to adjust the variable project in 02_rawdata/RWI/01_code/01_prep_ltw18_raster_area.py, and then run the Python script using arcpy.
- **Local matching.** The shapefile 06_arc_analyze/arc_output/ltw18_sb_intersected.shp contains the identifiers of matched neighboring precincts. The matching was performed manually in ArcGIS ArcMap (Version 10.4.1) and can be replicated in R following these steps:
  - copy the libraries from 05_packages/library to the respective folder on your machine and adjust setwd in the file 06_arc_analyze/prep_local_matching.R
  - run prep_local_matching.R

6. Output (figures and tables)

With a few exceptions, all figures and tables are produced by the do-files provided in the replication package. (Generated tables are typically further formatted by hand for the manuscript). The exceptions concern the following exhibits:

- Table E.7 is created as a .log file and converted to a .tex file by hand.
- Figures 2 and 3 were generated using ArcGIS ArcMap (Version 10.4.1). Geo-located addresses are exported as .txt files in 01_code/1_prep_and_geolocate_addresses/09_export_for_arcgis.do. The shapefiles for the maps are located in the respective subfolders of 02_rawdata/shp/. For

Figure 3, the single precinct shapefiles (for precinct numbers 806 in FE 2017 and ME 2014, 612 in FE 2017, and 612 and 613 in ME 2014) are a subset of the respective precinct shapefile for each election. You can replicate the maps following these steps:

- **Figure 2**:
  - Add the "Light Gray Canvas" basemap and drag all files from subfolder `shp/shapes_fig2` to the basemap.
  - Arrange the files as follows (from top to bottom): "stadtbezirke.shp", "landtagwahl2018_stimmbezirke.shp", "wl18.shp", "voter18_line.shp".
  - Go to "Layout View". Insert a north arrow, a scale bar, and a legend and change the names of the shapefiles to the following: "stadtbezirke" to "District borders", "landtagwahl2018_stimmbezirke" to "Precinct borders", "wl18" to "Polling places", and "voter18_line" to "Assignments".
- **Figure 3 (left)**:
  - Add the "Light Gray Canvas" basemap and drag all files from subfolder `shp/shapes_fig3a` to the basemap.
  - Arrange the files as follows (from top to bottom): "sb17_806.shp", "sb14_806.shp", "wl17_806.shp", "wl14_806.shp", "voter17_806_line.shp", "voter14_806_line.shp".
  - Go to "Layout View". Insert a scale bar and a legend and change the names of the shapefiles to the following: "sb17_806" to "Precinct borders 2017", "sb14_806" to "Precinct borders 2014", "wl17_806" to "Polling places 2017", "wl14_806" to "Polling places 2014", "voter17_806_line" to "Assignments 2017", and "voter14_806_line" to "Assignments 2014".
- **Figure 3 (right)**:
  - Add the "Light Gray Canvas" basemap and drag all files from subfolder `shp/shapes_fig3b` to the basemap.
  - Arrange the files as follows (from top to bottom): "sb17_612.shp", "sb14_612_613_diss1.shp", "wl17_612.shp", "wl14_612.shp", "voter17_612_line.shp", "voter14_612_line.shp".
  - Go to "Layout View" and insert a north arrow.

- Figure B.1 was created in PowerPoint.

Table #	Program	Output files
Table 1	03b_main_Tables_1_C1_C2_C3_C4	Table_1_ES_ctr_ln_street_dist.tex
Table C.1	03b_main_Tables_1_C1_C2_C3_C4	Table_C1_ES_bsl.tex
Table C.2	03b_main_Tables_1_C1_C2_C3_C4	Table_C2_ES_rob_clustering.tex
Table C.3	03b_main_Tables_1_C1_C2_C3_C4	Table_C3_ES_het_by_distance.tex
Table C.4	03b_main_Tables_1_C1_C2_C3_C4	Table_C4_ES_het_by_distance_3clust.tex
Table C.5	04h_rob_covariates_Tables_C5_C6	Table_C5_ES_rob_covariates_bsl.tex
Table C.6	04h_rob_covariates_Tables_C5_C6	Table_C6_ES_rob_covariates_dist2.tex
Table C.7	04i_rob_alt_treatm_Figures_C8_C9_C10_C11_Table_C7	Table_C7_Smpl_comp_by_treat_def.tex
Table E.1	01e_descriptives_representativeness_Munich_Table_E1	Table_E1_muc_represent.tex
Table E.2	02_sumstats_and_balance_Figure_6_Tables_E2_E3_E4	Table_E2_sumstat.tex
Table E.3	02_sumstats_and_balance_Figure_6_Tables_E2_E3_E4	Table_E3_uncond_balance_test_zscores.tex
Table E.4	02_sumstats_and_balance_Figure_6_Tables_E2_E3_E4	Table_E4_uncond_balance_test_nosz.tex
Table E.5	04e_rob_novelDiD_Figures_C1_D11_D12_Table_E5	Table_E5_ES_tau1_tau0.tex
Table E.6	03e_het_by_precinct_characs_Figures_10_D13_Tables_E6_E7	Table_E6_ES_het_z_charac.tex
Table E.7	03e_het_by_precinct_characs_Figures_10_D13_Tables_E6_E7	Table_E7_corrs_het_dimensions_2013.log (converted to .tex by hand)
Table E.8	04j_rob_balanced_smpl_Tables_E8_E9	Table_E8_ES_bal_samples_bsl.tex
Table E.9	04j_rob_balanced_smpl_Tables_E8_E9	Table_E9_ES_bal_samples_dist2.tex

Figure #	Program	Output files
Figure 1	01b_descriptives_precinct_level_Figures_1_D3_D5_D6	Figure_1_wahlber_turnout_over_time.pdf
Figure 2	*Map processed in ArcGIS ArcMap (Version 10.4.1)*	Figure_2_map_munich_overview.pdf
Figure 3a	*Map processed in ArcGIS ArcMap (Version 10.4.1)*	Figure_3a_map_euw14_no_consol_806.pdf
Figure 3b	*Map processed in ArcGIS ArcMap (Version 10.4.1)*	Figure_3b_map_euw14_612.pdf
Figure 4	01a_descriptives_address_level_Figures_4_5_D1_D2_D4	Figure_4_bar_share_treated2_HNR_over_time.pdf
Figure 5	01a_descriptives_address_level_Figures_4_5_D1_D2_D4	Figure_5_Density_distances_street
Figure 6	02_sumstats_and_balance_Figure_6_Tables_E2_E3_E4	Figure_6_COEFPLOT_balancetest_z_FEs.pdf
Figure 7	03a_main_Figure_7	Figure_7_ES_baseline.pdf
Figure 8	03d_het_by_distance_Figures_8_9	Figure_8_ES_het_by_distance_v2.pdf
Figure 9	03d_het_by_distance_Figures_8_9	Figure_9_ES_het_by_distance_3clust.pdf
Figure 10	03e_het_by_precinct_characs_Figures_10_D13_Tables_E6_E7	Figure_10_ES_het_z_charac_top5.pdf
Figure 11	03c_party_outcomes_Figures_11_D14	Figure_11_ES_party_outc_two.pdf
Figure B.1a	*Figure created in PowerPoint*	Figure_B1a_theory_diagram_t1_blue.pdf
Figure B.1b	*Figure created in PowerPoint*	Figure_B1b_theory_diagram_t1_blue.pdf
Figure B.2	06_conceptual_framework_Figures_B2_B3	Figure_B2_theory_rel_turnout_noIA.pdf
Figure B.3a	06_conceptual_framework_Figures_B2_B3	Figure_B3a_theory_rel_turnout_IA1.pdf
Figure B.3b	06_conceptual_framework_Figures_B2_B3	Figure_B3b_theory_rel_turnout_IA2.pdf
Figure C.1	04e_rob_novelDiD_Figures_C1_D11_D12_Table_E5	Figure_C1_ES_novel_ctr_wgt.pdf
Figure C.2	04d_rob_het_by_reason_Figure_C2	Figure_C2_ES_het_by_reason.pdf
Figure C.3	04g_rob_matching_Figures_C3_C4_C5	Figure_C3_ES_matching_bsl.pdf
Figure C.4	04g_rob_matching_Figures_C3_C4_C5	Figure_C4_ES_matching_dist2.pdf
Figure C.5	04g_rob_matching_Figures_C3_C4_C5	Figure_C5_map_local_matching_smpl.pdf
Figure C.6	04f_rob_honestDiD_Figures_C6_C7	Figure_C6_ES_RR_2023_bsl.pdf
Figure C.7	04f_rob_honestDiD_Figures_C6_C7	Figure_C7_ES_RR_2023_dist2.pdf
Figure C.8	04i_rob_alt_treatm_Figures_C8_C9_C10_C11_Table_C7	Figure_C8_ES_alternative_treatm_bsl.pdf
Figure C.9	04i_rob_alt_treatm_Figures_C8_C9_C10_C11_Table_C7	Figure_C9_ES_alternative_treat_dist2.pdf
Figure C.10	04i_rob_alt_treatm_Figures_C8_C9_C10_C11_Table_C7	Figure_C10_ES_DDD_RSevent_delta_dist_2slopes.pdf
Figure C.11	04i_rob_alt_treatm_Figures_C8_C9_C10_C11_Table_C7	Figure_C11_ES_dCdH23_dist.pdf
Figure D.1	01a_descriptives_address_level_Figures_4_5_D1_D2_D4	Figure_D1_WL_categories.pdf
Figure D.2	01a_descriptives_address_level_Figures_4_5_D1_D2_D4	Figure_D2_heatmap_wl_active_2009_20.pdf
Figure D.3	01b_descriptives_precinct_level_Figures_1_D3_D5_D6	Figure_D3_nbr_eligible_voters_stimmbez.pdf
Figure D.4	01a_descriptives_address_level_Figures_4_5_D1_D2_D4	Figure_D4_voter_reassignment_freq.pdf
Figure D.5	01b_descriptives_precinct_level_Figures_1_D3_D5_D6	Figure_D5_sb_treat_intensity.pdf
Figure D.6	01b_descriptives_precinct_level_Figures_1_D3_D5_D6	Figure_D6_heatmap_reassigned_bi.pdf
Figure D.7	01c_descriptives_precinct_maps_Figure_D7	Figure_D7_map_treatment_timing.pdf
Figure D.8	04a_rob_het_by_distance_Figures_D8_D10	Figure_D8_ES_het_by_distance_4clust.pdf
Figure D.9	04b_rob_het_by_distance_no_ambiguity_Figure_D9	Figure_D9_ES_het_by_distance_ambi.pdf
Figure D.10	04a_rob_het_by_distance_Figures_D8_D10	Figure_D10_ES_het_by_distance_shrHHdist4.pdf
Figure D.11	04e_rob_novelDiD_Figures_C1_D11_D12_Table_E5	Figure_D11_ES_novel_ctr_wgt_dist_increase.pdf
Figure D.12	04e_rob_novelDiD_Figures_C1_D11_D12_Table_E5	Figure_D12_ES_novel_ctr_wgt_ln_dist
Figure D.13	03e_het_by_precinct_characs_Figures_10_D13_Tables_E6_E7	Figure_D13_ES_het_z_charac_top5_ln_dist.pdf
Figure D.14	03c_party_outcomes_Figures_11_D14	Figure_D14_ES_party_outc_all.pdf
Figure D.15	04c_rob_party_district_level_Figure_D15	Figure_D15_ES_party_outc_two_district_postal.pdf