README for The Political Fallout of Air Pollution

Overview

The code in this replication package replicates all the findings from the paper *The Political Fallout of Air Pollution* by Luna Bellani, Stefano Ceolotto, Benjamin Elsner and Nico Pestel, submitted to the *Proceedings of the National Academy of Sciences*. The replication package constructs the main dataset (regional data at the country level) from various data sources as well as the survey dataset from the weekly opinion poll, and produces all the tables and figures shown in the paper. We also provide code for the analysis based on the Socio-Economic Panel, although we cannot make the data available due to confidentiality. All codes are Stata do-files (Version 15), except for one R script (R version 4.2.2, 2022-10-31). The run time of the code depends on the operating system. On a Macbook Pro with a M2 Pro chip and 16GB RAM, the run time is about 25 minutes.

Data Availability and Provenance Statements

Statement about Rights

We certify that the authors of the manuscript have legitimate access to and permission to use the data used in this manuscript. Furthermore, we certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package.

Summary of Availability

The regional data used in this manuscript are publicly available. The survey data from the Socio-Economic Panel and Politbarometer are confidential and cannot be made available. We provide instructions how to obtain the data from the providers.

Details on each Data Source

Geospatial Data

- Information on changes to county definitions following administrative reforms. The administrative reform in Saxony-Anhalt in 2007 led to a reduction in the number of counties through mergers and the re-assignment of several municipalities. The data on the changes of municipalities and counties are taken from Destatis (2008).

  Datafile: `/rawdata/refromed_counties/2007.xlsx`

- Data on the geo-location of weather stations is provided by the German Metererological Service (DWD). Datafile: `/rawdata/pollution/meta_station_DWD.dta`

- Data on the geo-location of pollution measuring stations is provided by the German Environment Agency (Umweltbundesamt).

  Datafile: `/rawdata/pollution/meta_station_UBA.dta`

- The geocodes of all counties in the definition post-2011 is provided in the file `/rawdata/maps/kreise_geocodes.dta`

Pollution Data

Pollution data for PM10, Ozone and NO_2 were compiled from daily published air quality readings by the German Federal Environment Agency (*Umweltbundesamt*). On its website https://www.umweltbundesamt.de/daten/luft/luftdaten (last accessed November 14, 2023), users can download spreadsheets with daily averages of various pollutants for all measuring stations in Germany. The data were compiled from thousands of spreadsheets — one per day and pollutant. We provide here the .dta file for PM10, Ozone and NO_2.

Datafile: `/rawdata/pollution/pollution_UBA_d_19902019.dta`

Weather Data

Data on weather and climate (temperature (°C), relative humidity (%), wind (m/s), precipitation (mm/m^2), solar radiation (h), air pressure (hpa) and dew point (°C)) were compiled from data from the German Metereological Service (Deutscher Wetterdienst, DWD). On its website, the DWD offers .txt files with daily averages or sums of the aforementioned weather variables. URL: https://opendata.dwd.de/climate_environment/CDC/observations_germany /climate/daily/kl/historical/. We provide here the .dta file, which contains the weather data used in the manuscript.

Datafile: `/rawdata/pollution/weather_d_19802019.dta`

Population Data

- Annual population data at the county level are taken from the Statistische Ämter des Bundes und der Länder (2021d). Separate files for each year can be found in the folder `/rawdata/demoecon/`.
- The population in Mecklenburg-Vorpommern from 2000 to 2011 at the municipality level is taken from the LAIV (2023). The folder `/refromed_counties/population` contains 12 files with the population numbers for each year.

Economic Data

- Annual GDP data at the county level are taken from the Statistische Ämter des Bundes und der Länder (2021b). Separate files for each year can be found in the folder `/rawdata/demoecon/`.

- Annual employment data at the county level are taken from the Statistische Ämter des Bundes und der Länder (2021c). Separate files for each year can be found in the folder `/rawdata/demoecon/`.

Voting Data

- **Federal Elections**
  - Voting data for the federal election at the county level is taken from Statistische Ämter des Bundes und der Länder (2021). The spreadsheets with the raw data can be found in the folder `/rawdata/voting/Bundestagswahlen_Kreis-level`.
  - Saxony, Federal elections 2002 at the municipality level; datafile: `/rawdata/refromed_counties/votes/Sachsen_federal_2002.xlsx`. Source: Statistisches Landesamt Sachsen (2002).
  - Saxony, Federal elections 2005 at the municipality level; datafile: `/rawdata/refromed_counties/votes/Sachsen_federal_2005.xlsx`. Source: Statistisches Landesamt Sachsen (2005).
  - Saxony-Anhalt, Federal elections 2002 at the municipality level; datafile: `/rawdata/refromed_counties/votes/Sachsen-Anhalt_federal_2002.csv`. Source: Statistisches Landesamt Sachsen-Anhalt (2002).
  - Saxony-Anhalt, Federal elections 2005 at the municipality level; datafile: `/rawdata/refromed_counties/votes/Sachsen-Anhalt_federal_2005.csv`. Source: Statistisches Landesamt Sachsen-Anhalt (2005).
- **State Elections**
  - County-level data for all state elections in the sample period are taken from Regionaldatenbank Deutschland, Statistische Ämter des Bundes und der Länder (2021a). The datafiles can be found in the folder `/rawdata/voting/Landtagswahlen - Kreis Ebene`.
  - Mecklenburg-Vorpommern, State elections 2002 at the municipality level; datafile: `/rawdata/refromed_counties/votes/Mecklenburg-Vorpommern_state_2002.xls`. Source: LAIV (2002).
  - Mecklenburg-Vorpommern, State elections 2006 at the municipality level; datafile: `/rawdata/refromed_counties/votes/Mecklenburg-Vorpommern_state_2006-2.xls`. Source: LAIV (2006).
  - Mecklenburg-Vorpommern, State elections 2011 at the municipality level; datafile: `/rawdata/refromed_counties/votes/Mecklenburg-Vorpommern_state_2011.xls`. Source: LAIV (2011).
  - Saxony, State elections 2004 at the municipality level; datafile: `/rawdata/refromed_counties/votes/Sachsen_state_2004.xlsx`. Source: Statistisches Landesamt Sachsen (2004).
  - Saxony-Anhalt, State elections 2002 at the municipality level; datafile: `/rawdata/refromed_counties/votes/Sachsen-Anhalt_State_2002.csv`. Source: Statistisches Landesamt Sachsen-Anhalt (2002).

  - Saxony-Anhalt, State elections 2006 at the municipality level; datafile: `/rawdata/refromed_counties/votes/Sachsen-Anhalt_State_2006.xlsx`. Source: Statistisches Landesamt Sachsen-Anhalt (2006).

Survey data from the Socio-Economic Panel

We use version 35 of the SOEP-core dataset (SOEP, 2019). The SOEP data are confidential and, thus, cannot be distributed in a replication package. We used the remote access through SOEPremote, which provides identifiers for respondents' current county of residence. Researchers who want to access the SOEP data need to sign an agreement with DIW Berlin. Instructions for general access and remote access are provided here: https://www.diw.de/en/diw_01.c.601584.en/data_access.html#c_741353 (last accessed 14 November 2023). The file below is the datafile provided by SOEPremote.

Datafile: `V:/soep/datasets/2018/V35/STATA_DEEN_v35/pequiv.dta` (not provided)

Survey data from Politbarometer

We use survey data from Politbarometer, which is offered by Forschungsgruppe Wahlen (2020). The data are confidential but researchers can access them free of charge from Gesis. Instructions are provided here: https://www.gesis.org/wahlen/politbarometer.

Datafile: `ZA2391_v12-0-0.dta` (not provided)

Software Requirements

- Stata (code was last run with version 15.1)
  - `estout`
  - `carryforward`
  - `distinct`
  - `reghdfe`
  - `ftools`
  - `coefplot`
  - `ivreg2`
  - `ranktest`
  - `ivreghdfe`
  - `egenmore`
  - `palettes`
  - `colrspace`
  - `labutil`
  - `scheme-burd`
  - `grc1leg`
  - the program "`setup.do`" will install all dependencies locally, and should be run once.
- R version 4.2.2 (2022-10-31)

  - the packages `dplyr`, `ggplot2` and `haven` need to be installed with `install.packages()`

Instructions for Replicators

Replicators need to copy the replication package to their machine and specify the main directory via `global path` in line 28 of the `master.do` file. They also need to specify the main directory in line 6 of the R script `codes/4_figure_S14.R`. When replicators run `master.do`, all the results based on the regional data are getting replicated (i.e. the main results). The master file gives several options, namely that only the dataset gets created or only the analysis is run. The default is that the code creates the dataset and produces all the results. After running the do-file `master.do`, replicators should run the R script `codes/4_figure_S14.R`, which produces Figure S14 in the Appendix.

In the folder `/codes`, we provide codes that run the analysis that is based on the confidential datasets *SOEP* and *Politbarometer*, respectively.

In `soep.do`, researchers need to modify the local `path` in line 3 such that it points to the main directory on the replicator's computer. This path is for saving outputs. Replicators should not change the local `soep` in line 2, as this is the directory on the remote server of SOEPremote.

Description of programs/code

The file `master.do` in the main directory runs all the necessary codes to create the dataset and reproduce all the tables, figures and numbers contained in the manuscript that are based on the regional data. The master file will call all the codes in the `/codes` folder in sequence. These contain the following files:

- data cleaning
  - The program `distanceprog.do` writes a function that computes the distance between two coordinates. This program is subsequently used to calculate the distance between country centroids and weather stations.
  - The programs `1_reformed_counties_votes.do` and `1_reformed_counties_pop.do` produce voting and population data for three states in which the administrative county boundaries changed during our sample period: Saxony-Anhalt in 2007, Saxony in 2008, and Mecklenburg-Vorpommern in 2011. In all three states, counties were merged. The programs create population and voting data for the *post-reform* definition for the entire sample period, i.e. 2000-2019 based on voting and population data at the municipality level.
  - The program `1_data.do` cleans the voting, pollution, weather and demographic data. It produces a county-level panel dataset of voting results (county-election date panel), county-level panel datasets of

daily air pollution readings and daily weather data, and a county-year panel dataset of demographic and economic variables.
  - The program `2_maps.do` takes the geo-coordinates of all counties in Germany and computes the centroids of each county.
  - The program `2_merge.do` first merges the geo-spatial data (pollution, weather) and calculates distance from each country centroid to the nearest measuring station. Second, it merges the geo-spatial data with the voting and demographic/economic data via the county centroids.
- Figures
  - The program `3_figure1.do` merges the pollution data with the survey data from Politbarometer and produces Figure 1 from the manuscript. Replicators need to copy the Politbarometer dataset into the folder `/rawdata`. In the file `master.do`, these files are commented out because the Politbarometer data cannot be provided within this replication package.
  - The program `1_SOEP_merge.do` merges the pollution data with the SOEP data on the SOEPremote server.
  - The program `1_SOEP_reg.do` performs the regressions underlying Figure 2 from the manuscript. Figure 2 gets produced with the program `3_figure2.do`. In the file `master.do` it is commented out because the SOEP data cannot be provided within this replication package.
  - The following programs produce the respective figures in the supplementary material: `4_figure_S1a.do`, `4_figure_S1b.do`, `4_figure_S2.do`, `4_figure_S3.do`, `4_figure_S4.do`, `4_figure_S5.do`, `4_figure_S6.do`, `4_figure_S7.do`, `4_figure_S8.do`, `4_figure_S9.do`, `4_figure_S10.do`, `4_figure_S11a.do`, `4_figure_S11b.do`, `4_figure_S12a.do`, `4_figure_S12b.do` and `4_figure_S13.do`.
  - The file `4_figure_S14S15data.do` prepares the data for Figures S14 and S15.
  - The file `4_figure_S15.do` produces the dumbbell graph, Figure S15 in the Appendix.
  - The R script `4_figure_S14.R` produces the dumbbell graph of Figure S14. To run the script, replicators need to specify the path to the main directory in line 6.
- Tables
  - The program `table1.do` produces Table 1 from the manuscript.
  - The programs `table_S1.do`, `table_S4.do`, `table_S5.do`, `table_S6.do`, `table_S7.do` and `table_S8.do` produce the respective tables in the supplementary material. Tables S2 and S3 do not contain any analysis and, thus, there is no code to generate them.
- Numbers that are mentioned in the text get generated in the various programmes and entered into the file `in_text_numbers_generated_in_stata.txt`.

References

Destatis (Statistisches Bundesamt). 2008. Daten aus dem Gemeindeverzeichnis. Gebietsänderungen vom 01.01.2007 bis 31.12.2007. Datafile version: https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Namens-Grenz-Aenderung/2007.xlsx?___blob=publicationFile. Last accessed: 13 November 2023.

Forschungsgruppe Wahlen. 2020. Politbarometer 1977–2019 (Partielle Kumulation). GESIS Datenarchiv, Koeln. ZA2391 Datenfile Version 12.0.0, https://doi.org/10.4232/1.13631, doi:10.4232/1.13631

LAIV (Landesamt für Innere Verwaltung). 2002. Landtagswahl 2002 in Mecklenburg-Vorpommern, Ergebnisse nach Gemeinden. Datafile version: https://www.laiv-mv.de/static/LAIV/Wahlen/Dateien/Publikationen/B721/B724G%202002%2001.xls. Last accessed: 13 November 2023.

LAIV (Landesamt für Innere Verwaltung). 2002. Bundestagswahl 2002 in Mecklenburg-Vorpommern, Ergebnisse nach Gemeinden. Datafile version: https://www.laiv-mv.de/static/LAIV/Wahlen/Dateien/Publikationen/B711/B714G%202002%2001.xls. Last accessed: 13 November 2023.

LAIV (Landesamt für Innere Verwaltung). 2005. Bundestagswahl 2005 in Mecklenburg-Vorpommern, Ergebnisse nach Gemeinden. Datafile version: https://www.laiv-mv.de/static/LAIV/Wahlen/Dateien/Publikationen/B711/B714G%202005%2001.xls. Last accessed: 13 November 2023.

LAIV (Landesamt für Innere Verwaltung). 2006. Landtagswahl 2006 in Mecklenburg-Vorpommern, Ergebnisse nach Gemeinden. Datafile version: https://www.laiv-mv.de/static/LAIV/Wahlen/Dateien/Publikationen/B721/B724G%202006%2001.xls. Last accessed: 13 November 2023.

LAIV (Landesamt für Innere Verwaltung). 2009. Bundestagswahl 2009 in Mecklenburg-Vorpommern, Ergebnisse nach Gemeinden. Datafile version: https://www.laiv-mv.de/static/LAIV/Wahlen/Dateien/Publikationen/B711/B714G%202009%2001.xls. Last accessed: 13 November 2023.

LAIV (Landesamt für Innere Verwaltung). 2011. Landtagswahl 2011 in Mecklenburg-Vorpommern, Ergebnisse nach Gemeinden. Datafile version: https://www.laiv-mv.de/static/LAIV/Wahlen/Dateien/Publikationen/B721/B724G%202011%2001.xls. Last accessed: 13 November 2023.

LAIV (Landesamt für Innere Verwaltung). 2022. Mecklenburg-Vorpommern, A133G Bevölkerung nach Alter und Geschlecht, Teil 2: Gemeindeergebnisse. URL: https://www.laiv-mv.de/Statistik/Zahlen-und-Fakten/Gesellschaft-&-Staat/Bevoelkerung. Last accessed: 13 November 2023.

SOEP. 2019. Socio-Economic Panel (SOEP), data for years 1984-2018, version 35. doi:10.5684/soep.v35.

Statistische Ämter des Bundes und der Länder. 2021. Regionaldatenbank:

Bundestagswahlen, 14111-01-04-4. URL: https://www.regionalstatistik.de/genesis//online?operation=table&code=14111-01-04-4&bypass=true&levelindex=1&levelid=1699888581359#abreadcrumb. Last accessed: 13 November 2023.

Statistische Ämter des Bundes und der Länder. 2021a. Regionaldatenbank: Landtagswahlen, 143. URL: https://www.regionalstatistik.de/genesis/online?operation=themes&levelindex=0&levelid=1699957750509&code=14#abreadcrumb. Last accessed: 14 November 2023.

Statistische Ämter des Bundes und der Länder. 2021b. 82111-01-05-4: Bruttoinlandsprodukt/Bruttowertschöpfung nach Wirtschaftsbereichen. URL: https://www.regionalstatistik.de/genesis//online?operation=table&code=82111-01-05-4&bypass=true&levelindex=1&levelid=1699961885301#abreadcrumb. Last accessed: 14 November 2023.

Statistische Ämter des Bundes und der Länder. 2021c. 13312-01-05-4: Erwerbstätige nach Wirtschaftszweigen. URL: https://www.regionalstatistik.de/genesis//online?operation=table&code=13312-01-05-4&bypass=true&levelindex=1&levelid=1699963805799#abreadcrumb. Last accessed: 14 November 2023.

Statistische Ämter des Bundes und der Länder. 2021d. 12411-01-01-4: Bevölkerung nach Geschlecht. URL: https://www.regionalstatistik.de/genesis//online?operation=table&code=12411-01-01-4&bypass=true&levelindex=1&levelid=1699963887098#abreadcrumb. Last accessed: 14 November 2023.

Statistisches Landesamt Sachsen. 2002. Bundestagswahl 2002, Ergebnis 14111-093M. Datafile version: https://www.statistik.sachsen.de/genonline/online?operation=previous&levelindex=2&step=2&titel=Ergebnis&levelid=1699885011817&acceptscookies=false#abreadcrumb. Last accessed: 13 November 2023.

Statistisches Landesamt Sachsen. 2004. Landtagswahl 2002, Ergebnis 14311-093M. Datafile version: https://www.statistik.sachsen.de/genonline//online?operation=table&code=14311-093M&bypass=true&levelindex=0&levelid=1699886553702#abreadcrumb. Last accessed: 13 November 2023.

Statistisches Landesamt Sachsen. 2005. Bundestagswahl 2005, Ergebnis 14111-091M. Datafile version: https://www.statistik.sachsen.de/genonline//online?operation=table&code=14111-091M&bypass=true&levelindex=0&levelid=1699884959918#abreadcrumb. Last accessed: 13 November 2023.

Statistisches Landesamt Sachsen-Anhalt. 2002. Bundestagswahl 2002, Endgültige Ergebnisse in den Gemeinden. Datafile version: https://wahlergebnisse.sachsen-anhalt.de/wahlen/bt02/and/BT2002_GEM.csv. Last accessed: 13 November 2023.

Statistisches Landesamt Sachsen-Anhalt. 2002. Landtagswahl 2002, Endgültige Ergebnisse in den Gemeinden. Datafile version: https://wahlergebnisse.sachsen-anhalt.de/wahlen/lt02/and/LT2002_GEM.csv. Last accessed: 13 November 2023.

Statistisches Landesamt Sachsen-Anhalt. 2005. Bundestagswahl 2005, Endgültige Ergebnisse in den Gemeinden. Datafile version: https://wahlergebnisse.sachsen-anhalt.de/wahlen/bt05/and/BT2005_GEM.csv. Last accessed: 13 November 2023.

Statistisches Landesamt Sachsen-Anhalt. 2006. Landtagswahl 2006, Endgültige Ergebnisse in den Gemeinden. Datafile version: https://wahlergebnisse.sachsen-anhalt.de/wahlen/lt06/and/LT2006GEM.csv. Last accessed: 13 November 2023.