

README for Policy Pulse: Sensing the sentiment of climate policies in the public sphere

Nisan Gorgulu and Penny Mealy

Contents

1. [Overview](#)
2. [Data Availability](#)
3. [Instructions for Replicators](#)
4. [List of Exhibits](#)
5. [Requirements](#)

Overview

This document provides instructions for reproducing the empirical results presented in the paper titled *“Policy Pulse: Sensing the sentiment of climate policies in the public sphere”*, which introduces a novel methodology for assessing the sentiment towards climate policies - particularly carbon taxes and emissions trading schemes (ETS) from news media articles data from the Global Database of Events, Language, and Tone (GDELT) for the period 2017 – 2023.

Folder structure:

Reproducibility Package/

1-Data/

└─ 1-data-raw/

└─ 2-data-processed/

2-Code/

└─ mainStata.do

└─ Main_Python.ipynb

└─ Fig3-YellowVest.ipynb

└─ Fig5-GlobalSentTrends.ipynb

└─ Fig6-BoxandWhisker.ipynb

└─ Fig7a-IncomeTrends.ipynb

└─ Fig7b-IncomeTrends.ipynb

└─ TableA1_ADF.ipynb

3-Output/

└─ Figures/

└─ PolicyPulseMaps_Cleared/

└─ Tables/

Policy Pulse – README.docx

Execution Order:

1. Run mainPython.py >> Generates figures 1 -3, 5-7 and Table A1.
2. Run mainStata.do >> Generates tables 1,2, A3.1

Both scripts automatically save outputs to 3-Output/

Maps A2.1 - A2.4 under PolicyPulseMaps_Cleared are generated separately by the World Bank Cartography Unit using the raw data, which will be described in the Instructions for Replicators section.

Data Availability

All data used in this study are publicly available and included in the repository under **1-Data**. All the raw data is included in the **1-Data/1-data-raw** and processed data are in **1-Data/2-data-processed**.

Data Sources

1. Carbon Pricing Sentiment Data

- a. Filename: GDELT.xlsx, GlobalSentimentTrends.csv (monthly), and CarbonPricing_DailySentiment.xlsx
- b. Source: Global Database of Events, Language, and Tone (GDELT)
- c. URL: <https://www.gdeltproject.org/>
- d. Access year: 2024

2. World Values Survey

- a. Filename: WVS.xlsx
- b. Source: World Values Survey (WVS)
- c. URL: <https://www.worldvaluessurvey.org/WVSContents.jsp>
- d. Reference to download data:
<https://www.worldvaluessurvey.org/WVSContents.jsp>
- e. Access year: 2024

3. IMF-YouGOV Survey

- a. Filename: YouGOV.xlsx
- b. Source: Self-collected from the paper titled “*Public Perceptions of Climate Mitigation Policies: Evidence from Cross-Country Surveys*” by Dabla-Norris *et al.* (2023) (see Figure 6. Support for Mitigation Policies on page 11).
- c. URL of the paper: <https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2023/02/07/Public-Perceptions-of-Climate-Mitigation-Policies-Evidence-from-Cross-Country-Surveys-528057>

4. GDP per capita (current US\$)

- a. Filename: API_NY.GDP.PCAP.CD_DS2_en_excel_v2_1107224.xls
- b. GDP per capita data for the years 2017 – 2023 is extracted manually from the raw data (a) and saved under processed data (GDPpc.xlsx).
- c. Source: World Development Indicators
- d. URL: <https://data.worldbank.org/indicator/NY.GDP.PCAP.CD>
- e. Access year: 2024

5. World Bank Income Group Classification – used to categorize countries into income groups.

- a. Filename: OGHIST_2025_10_07.xlsx
- b. Historical classification by income is downloaded from:
<https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>
- c. Reference year: calendar year 2022 (FY 2024)
- d. Access year: 2024

6. Press Freedom Index Score

- a. Filename: RWB_PFI_SCORE.xlsx
- b. Press Freedom Index Score for the years 2017 – 2023 is extracted manually from the raw data (a) and saved under processed data (PressFreedom.xlsx).
- c. Source: World Bank Data 360 – ID: WB.DATA360:DS_DATA360(1.2)
- d. URL: https://data360.worldbank.org/en/int/indicator/RWB_PFI_SCORE
- e. Access year: 2025

7. Carbon Intensity of GDP (kg CO2e per 2021 PPP \$ of GDP)

- a. Filename: API_EN.GHG.CO2.RT.GDP.PP.KD_DS2_en_excel_v2_128125.xls
- b. Carbon Intensity data for the years 2017 – 2023 is extracted manually from the raw data (a) and saved under processed data (CarbonInt.xlsx).
- c. Source: World Bank based on EDGAR (Emissions Database for Global Atmospheric Research) Community GHG Database
- d. URL: <https://data.worldbank.org/indicator/EN.GHG.CO2.RT.GDP.PP.KD>
- e. Access year: 2025

8. UNDP People’s Climate Vote 2024

- a. Filename: UNDP_ClimateVote.xlsx
- b. % of responses for the public support for country climate commitments represent the responses to the question: “Should your country strengthen or weaken its commitments to address climate change?”. It is reported under the column “weighted mean” in the dataset.
- c. Source: Dataset can be downloaded from <https://peoplesclimate.vote/data-center>
- d. Access year: 2026

Instructions for Replicators

Follow the steps below to reproduce the results.

- Ensure that all required software and dependencies are installed as listed in the [Requirements](#) section.
- Before running any scripts, users need to set the project root directory to match the local location of the *Reproducibility Package* folder.
- Maps A2.1 - A2.4 are generated by the World Bank Cartography Unit using the raw data in "1-Data/1-data-raw/Data_for_PolicyPulse_Maps.xlsx", which was based on the main data collected from the GDELT. The final maps are in the folder "3-Output/PolicyPulseMaps_Cleared".
- Figures 1 and 2 are illustrative news articles with associated sentiment scores. The samples were compiled by the authors from the Global Database of Events, Language, and Tone (GDELT) (<https://api.gdeltproject.org/api/v2/summary/summary>) using the keywords related to carbon tax and emissions trading scheme in Table 1 in the paper. The links and associated sentiments for each news article are in the folder "2-Data/Processed".
- Figures 4 a-f are prepared manually in Microsoft Excel for visual comparability across survey sources. The figures are produced by matching "Country" name field and applying simple lookup operations (e.g., VLOOKUP) to align survey responses with the corresponding sentiment measures. All input datasets used in these figures are provided in the repository.
 - For the World Values Survey (WVS), responses to the question "*Protecting the environment vs economic growth: Which of them comes closer to your own point of view? 1. Protecting the environment should be given priority, even if it causes slower economic growth and some loss of jobs. 2. Economic growth and creating jobs should be the top priority, even if the environment suffers to some extent*" are aggregated manually in Excel. A total responses column is constructed as the sum of responses (1) protecting environment, (2) economic growth, (3) other answer. The variable "WVS_ProtectEnv (%)" is then computed as the share of response selecting protecting environment relative to this total, multiplied by 100. All calculations are documented via formulas within the Excel file.
 - Data for GDELT and UNDP Climate Vote datasets are downloaded from the websites stated under data sources.

- Data on public support for climate mitigation policies are manually extracted from Figure 6 in “*Support for Mitigation Policies*” by Dabla-Norris *et al.* (2023). The reported indicators correspond to the percentage of survey respondents expressing support for climate mitigation policies.
- Figures 7a and 7b are generated using the files CarbonTax_DailySent_byIncome.csv and ETS_DailySent_byIncome.csv, located in 1-Data/2-data-processed/. These datasets are constructed from the file CarbobPricing_DailySentiment.xlsx (located in 1-Data/1-data-raw/), which contains daily country level sentiment measures for carbon tax and ETS policies. For the processed datasets, the variables Date, CountryCode, and AverageTone are retained. An Income Group variable (LMIC, UMIC, LIC, HI) are based on the World Bank 2022 calendar year (FY2024) income classification (OGHIST_2025_10_07.xlsx; located in 1-Data/1-data-raw/). The income classification is merged using CountryCode as the key in Excel (via standard lookup functions). The resulting datasets are saved as csv files and are used by mainPython.py to produce Figures 7a and 7b.
- Tables 1, 2, and A3.1 are produced in Stata using PolicyPulse_data_for_Tables.dta, which is manually constructed from: GDELT.xlsx, CanbonInt.xlsx, FreedomPress.xlsx, and GDPpc.xlsx.
 - Each dataset has “Country” column and is provided in wide format. They are reshaped manually at the country-year format and merged.
 - The Stata estimation file is a country-year panel dataset constructed from the input files described above. No additional data merging or reshaping is performed within mainStata.do; the script directly uses the prepared panel dataset to replicate the reported results.

List of Exhibits

Exhibit name	Output filename	Script
Figures 1 -2	Figs1-2.pptx	Produced in PPT using 2-data-processed/Figs1-2.xlsx
Figure 3	3- Output/Figures/ Fig3_YellowVest_Sentiment.jpg	2-Code/mainPyhon.py
Figures 4 a-f	Fig4_Surveys.xlsx (in 3- Output/Figures)	Prepared in Excel by hand, using the datasets GDELT.xlsx, WVS.xlsx, and YouGov.xlsx and UNDP_ClimateVote.xlsx
Figure 5	3- Output/Figures/ Fig5_GlobalSentTrends.jpg	2-Code/mainPyhon.py

Figure 6 a-b	3- Output/Figures/ Fig6-BoxandWhisker.jpg	2-Code/mainPyhon.py
Figure 7a	3- Output/Figures/ Fig7a_CarbonTax_byIncome.jpg	2-Code/mainPyhon.py
Figure 7b	3- Output/Figures/ Fig7b_ETS_byIncome.jpg	2-Code/mainPyhon.py
Tables 1 - 2	Table1.csv and Table2.csv	2-Code/mainStata.do
Table A1	3- Output/Tables/ TableA1_ADF_results.csv	2-Code/mainPyhon.py
Table A3.1	TableA3_1.csv	2-Code/mainStata.do
Maps A2.1 - 4	WLD48967.pgj, WLD48968.pgj, WLD48969.pgj, WLD48970.pgj	Generated by the World Bank Cartography Unit based on Data_for_PolicyPulse_Maps.xlsx

Requirements

Computational Requirements

Software Requirements

- ☐ Microsoft Excel
- ☐ Stata Version 18.0
- ☐ Python (Anaconda or version 3.10 recommended)
 - Required Python packages include: pandas, numpy, statsmodels, matplotlib, and geopandas

Memory and Runtime and Storage Requirements

Runtime: Approx. 10 mins