Background and Contact

This is the replication package for the following article:

> Gazmararian, Alexander F. 2024. "Sources of Partisan Change: Evidence from the Shale Gas Shock in American Coal Country." *The Journal of Politics*.

For questions, please email Alexander F. Gazmararian at: afg2@princeton.edu; agazmararian@gmail.com

Licensing and Usage Guidelines

This replication package is provided under the MIT License. This license permits the use, modification, distribution, and private use of the content in this package, but with no warranty as stated in the license terms. The datasets, code, and documentation within this package are free for academic, educational, and research purposes. I kindly request that users cite the associated academic article when utilizing or referencing this package in their research or publications. For any commercial use or adaptation of the materials, please contact the author for permission. By using or distributing this replication package, you agree to abide by the terms and conditions of the MIT License.

Software and Package Dependencies

Analyses were performed on a MacBook Pro (M2 chip) with 32 GB of memory running macOS 14.1.1 and using R version 4.3.0 (2023-04-21) [aarch64-apple-darwin20 (64-bit)].

The preamble of each script lists the package dependencies.

Replication Instructions

1. Load the "jop_replication.Rproj" R Project environment. Do not set a working directory. The replication package is self-contained. All file paths are relative, using the `{here}` package.
2. Run the scripts sequentially. Begin with scripts that start with "01", then "02", and so on.
3. Remove objects from the Global Environment between running each script to avoid complications.

The **data** and **output** folders contains data, tables, and figures from running the code, demonstrating the smooth running of all the replication files.

Directory Structure and Scripts

- **code**
  - **01_anthracite** (The scripts in this folder prepare data for the falsification test that exploits variation in anthracite coal's vulnerability to the shale gas shock.)
    - **01_clean_antemp.R** (Prepares the data for creating the treatment group.)
    - **02_validateanthracite.R** (Validates the logic of the falsification test by examining the correlation between natural gas prices and anthracite production.)
  - **01_plantdistance** (The scripts in this folder create the measure of a county's distance to a new natural gas power plant.)
    - **01A_nondistance_columns.R** (Prepares data from 1992 and 1996.)
    - **01B_nondistance_columns_fuzzy_join.R** (Prepares data through 2020.)
    - **02_distance_columns.R** (Calculates distance from county center to new plants.)
    - **03_joining_years.R** (Creates interim dataset to be merged later for analysis.)

- **01_process** (The scripts in this folder prepare raw data to be merged later for the primary analysis.)
  - **01_clean_cbp.R** (Prepares the county employment data for analysis.)
  - **01_clean_census.R** (Downloads and prepares the US Census data for analysis.)
  - **01_clean_countygdp.R** (Prepares county-industry level GDP data for analysis.)
  - **01_clean_debt.R** (Prepares household debt-to-income ratio data for analysis.)
  - **01_clean_eia_annualcoalreport.R** (Prepares coal production data used in subsequent analysis.)
  - **01_clean_eia_electricityfuel.R** (Prepares net-generation data for analysis of shale shock timing.)
  - **01_clean_eia_nautralgasannual.R** (Prepares natural gas production data for analysis of shale shock timing.)
  - **01_clean_elections.R** (Prepares the election data for analysis.)
    - Note: Under the terms of access and copyright for the elections data, I am prohibited from sharing the complete underlying dataset. I still provide the script that processes this data for replication. If you do not have access to Dave Leip's Atlas of U.S. Elections through your institution, I am allowed to share the data with researchers directly upon request.
  - **01_clean_fuelprices.R** (Prepares gas and coal price data for analysis.)
  - **01_clean_mines.R** (Prepares mine data for falsification tests.)
  - **01_clean_powergrants.R** (Prepares POWER grant data for analysis.)
  - **01_clean_union_eia7a.R** (Prepares data on coal mine unionization levels for analysis.)
  - **01_stateshocks.R** (Estimates state-specific shale shocks for analysis.)
- **02_adjcoalemp.R** (Identifies coal counties and adjacent counties to be used to construct the matched dataset for analysis.)
- **02_natunionrates_fig_H4.R** (Downloads data and creates Fig. H4 that depicts national trends in coal miner unionization.)
- **02_shocktiming_fig1.R** (Prepares Fig. 1 that identifies the shale shock's timing.)
- **03_merge.R** (Merges all interim datasets for analysis and constructs variables.)
- **04_matching.R** (Creates matched dataset for analysis.)
- **05_analysis.R** (Performs the primary difference-in-difference analyses, robustness tests, falsification tests, moderating effect of visibility analyses, etc.)
- **05_compositional** (This folder contains scripts that analyze the robustness of the results to compositional changes.)
  - **01_process_migration.R** (Prepares migration data for analysis.)
  - **02_analyze_migration.R** (Analyzes migration data.)
  - **03_composition_counterfactual.R** (Analyzes vote share under counter-factual assumptions about compositional changes.)
- **05_fieldwork** (This folder contains scripts that describe the county fair samples and characterize the interview study site relative to the rest of coal country.)
  - **community_preferences.R** (Analyzes the extent to which survey respondents are embedded in networks with fossil fuel workers.)

        - **fair2021_summary.R** (Analyzes the representativeness of the 2021 county fair sample.)
        - **fair2022_summary.R** (Analyzes the representativeness of the 2022 county fair sample.)
        - **interview_sample.R** (Analyzes the representativeness of the study site county.)
        - **visibilitysurvey.R** (Analyzes the representatives of this subset of respondents for assessing the visibility of natural gas power plants in an applicable county.)
    - **05_laborsalience** (This folder contains scripts that analyze the salience of labor issues in coal country.)
        - **minefatalities.R** (Processes and and analyzes data on coal mine fatalities.)
        - **umwj.R** (Analyzes the text summarizing content from a selection of United Mine Worker Journals.)
    - **05_media** (This folder contains scripts that analyzes media coverage of the shale shock and decline of coal.)
        - **localnews.R** (Analyzes the content of local news in Southwest Pennsylvania.)
        - **tvcoverage.R** (Processes and analyzes cable new coverage of coal and the shale shock.)
    - **05_summarystats.R** (Creates summary statistics describing the primary dataset.)
    - **05_taxrevenue.R** (Analyzes the effect of new gas power plant construction on county tax revenue.)
    - **fun** (This folder contains custom functions called in other scripts.)
- **data**
    - **input** (This folder contains unprocessed data.)
    - **inter** (This folder contains processed data that has not yet been merged into the ultimate dataset.)
    - **output** (This folder contains the ultimate datasets used in the analyses).
- **output** (Location and and figures created by the scripts.)
    - **figures**
    - **tables**