Overview

The code in this replication package reproduces all of the tables and figures for Bushnell, Muehlegger and Rapson (2025) "Energy Prices and Electric Vehicle Adoption."

The Stata code included in the archive reads in the raw data, performs all data cleaning and manipulation and creates the Tables and Figures in the paper.

Computational requirements

The code was last run on an 8-core Intel-based laptop with Windows 11 Pro. The code was run with Stata MP, version 18 under default memory settings.

Data Availability and Provenance Statements

This paper draws on both public datasets and datasets that were purchased from third parties under data use agreements that limit the public dissemination of the data. We provide a description and summary of the public availability of each of the data sources in further detail below.

Statement about Rights

- ✓ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.
- ✓ I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package.

Instructions to Replicators

- Download the data files in the folder \Data. The folder \Data includes all non-proprietary data used in the paper.
- Create ResultsOut/Tables, ResultsOut/Figures, Data/Restricted/DMV, Data/Restricted/Experian, Data/Restricted/OPIS. Restricted OPIS, DMV and Experian data should be placed in the respective folders.
- Run the R-code to create distance_bg_2013_sub*.csv.
- Edit Code/master.do if needed to adjust the default path
- Running Code/master.do reads in, cleans and conducts all of the analysis performed in the paper.

Details on each Data Source

- **Vehicle Registration Data (EXPERIAN):** Transaction-level electric vehicle purchases for California from 2014 – 2017, were purchased from Experian (https://www.experian.com/automotive/auto-vehicle-data). The Experian data includes all battery electric and plug-in hybrid vehicle registered in California between January 2014 and December 2017. For each observation, the data includes the full 17-digit VIN, purchase price, vehicle make-model-model-year, odometer reading, the census block group of the owner, power-train type, vehicle segment, body-type and an indicator for whether the vehicle was leased. *As per the signed data agreement, this data cannot be publicly posted and is not included in the archive.*

- **Gasoline Price Data (OPIS)**: Station-level daily historical gasoline price data were purchased from the Oil Price Information Service (OPIS). https://www.opisnet.com/product/pricing/oil-prices-forecast-and-history/retail-datahouse/. The raw data includes reported prices for unleaded regular, midgrade and premium gasoline at the station-day level, the latitude and longitude of the station, the street address, and station characteristics, include the branding of the station (which was not used in this paper). The OPIS data were aggregated up to: zip code average retail prices by month (zip_month_avg_retail_price.dta) and average retail prices by month for all stations within different radii of the geographic centroid of each census block group (GasPriceRadii.dta). *As per the data agreement, this data cannot be publicly posted and is not included in the archive.*

- **Aggregate vehicle ownership data for 2013 (CA DMV)**: We generated aggregate vehicle ownership data for 2013 (i.e., the fraction of hybrid vehicles and luxury vehicles) from registration counts provided through a data use agreement with the California Department of Motor Vehicles. *As per the data agreement, this data cannot be publicly posted and is not included in the archive.*

- **Duration of ownership data from CA DMV (CA DMV)**: For each electric vehicle in the Experian data, we matched the VIN to DMV data tracking initial registration and subsequent re-registration of electric vehicles in California over 2014 – 2019. The data was provided through a data use agreement with the California Department of Motor Vehicles. *As per the data agreement, this data cannot be publicly posted and is not included in the archive.*

- **OpenEI Utility Rate Database (URDB)**: Publicly posted rate scheduled for utliities were downloaded from the URDB (Usurdb.xlsx, Usurdb_new.xlsx) They are provided as part of the archive.

- **Hand-collected utility rate data:** For utilities lacking rate schedules posted in the URDB, rate schedules were hand-collected and and supplemented with hand-collected data on utility rate schedules. These hand-collected data are provided as part of the archive. (SCE_rate_schedules_manually_collected.xlsx, SDGE_rate_schedules_manually_collected.xlsx, tiered_residential_rates_compiled.xlsx)

- **Energy Information Administration Form 861:** Utility names were standardized across datasets using the EIA Form 861 data. (Utility_Data_2017.xlsx) These data are provided as part of the archive as are two datasets that reflect hand-cleaned version of the EIA Form 861 Name Database (utility_names_shorthand.xlsx, utility_names_eiaid_masterlist.xlsx)

- **GIS output:** We include several datasets that were constructed using GIS software as part of the archive: (1) the assignment of census block groups to climate zones defined by the California Energy Commission (CBG_climate_zone.csv), (2) matched pairs of block groups that straddle utility boundaries (cbg_boundary_matched), (3) matched pairs of block groups that straddle municipal boundaries (cbg_boundary_muni_matched_named.dta)), (4) the assignment of census block groups to utility area boundaries (overlap_boundaries_proper_new.xlsx), (5) the mapping of zip codes to PGE service territory (RESZIPS.xlsx), and (6) the census block group demographic characteristics across the municipal area boundaries for the placebo test. (cbg_muni.dta).

- **Alternative Fuels Data Center (AFDC)**: Locations of public electric vehicle charging stations were mapped to census block groups. These data are provides as part of the archive. (join_cbg_to_alt_stations.xlsx)

- **Census American Community Survey (ACS)**: Demographics at the census block group level were collected from the ACS. These data are provides as part of the archive. (ACSDT5Y*-Data.csv, CBG_MUDs.csv, Med_income_bg_2013.csv). In addition, we provide provide a cross-walk that overlays the ZCTA shape files and block-group shape-shape files (join_ZCTA_bg.csv).

- **Utility Area ShapeFiles:** Utility Area Shapefiles are provided as part of the archive in Data/Electric_Utility_Service_Areas/ and are used to calculate distances between centroid of census block groups and from utility area boundaries.

Dataset list

Raw Data file	Source	Provided
ElectricityRates/Usurdb.xlsx	OpenEI URDB	Yes
ElectricityRates/Usurdb_new.xlsx	OpenEI URDB	Yes
ElectricityRates/Utility_Data_2017.xlsx	EIA Form 861	Yes
ElectricityRates/utility_names_shorthand.xlsx	Hand-coded	Yes
ElectricityRates/SCE_rate_schedules_manually_collected.xlsx	Hand-collected	Yes
ElectricityRates/SDGE_rate_schedules_manually_collected.xlsx	Hand-collected	Yes
ElectricityRates/tiered_residential_rates_compiled.xlsx	Hand-collected	Yes
ElectricityRates/utility_names_eiaid_masterlist.xlsx	EIA Form 861	Yes
Restricted/OPIS/zip_month_avg_retail_price.dta	OPIS	No
Restricted/OPIS /GasPriceRadii.dta	OPIS	No
Restricted/DMV/duration_data.dta	CA DMV	No
Restricted/DMV/CBG_dmv_2013.csv	Summarized CA DMV data	No
Restricted/Experian/ExperianMaster.dta	Experian	No
Restricted/Experian/experian_CBG_xwalk	Experian / CA DMV	No
Boundaries/CBG_climate_zone.csv	GIS	Yes
Boundaries/cbg_boundary_matched.dta	GIS	Yes
Boundaries/cbg_boundary_muni_matched_named.dta	GIS	Yes
Boundaries/distance_bg_2013_sub*.csv	Created by R-code	No
Boundaries/join_cbg_to_alt_stations.xlsx	AFDC / GIS	Yes
Boundaries/cbg_muni.dta	Hand-constructed	Yes
Boundaries/join_ZCTA_bg.csv	Census / GIS	Yes
Boundaries/overlap_boundaries_proper_new.xlsx	GIS	Yes
Boundaries/RESZIPS.XLS	Hand-constructed	Yes
Electric_Utility_Service_Areas/*	Utility Shapefiles	Yes
ACS/ACSDT5Y*-Data.csv	Census, ACS	Yes
ACS/CBG_MUDs.csv	Census, ACS	Yes
ACS/Med_income_bg_2013.csv	Census, ACS	Yes

Description of programs/code

Main file:

- master.do – runs all of the data cleaning and generates tables and figures, by calling .do files listed and described below.

Code Called by Main file:

- Data_Electric.do – cleans electricity data from Open EI. Creates Appendix Figures A1, A2.
- Data_GIS.do – cleans and merges data to the census block level.
- summary_statistics_table.do – create summary statistics presented in Table 1.
- Prepare_Panel_data.do – organize data for panel regressions.
- Panel_demog.do – create panel of demographics for panel regressions.
- PanelRegs.do – Panel regressions, creation of Table 2, Appendix Table A2.
- Prepare_SpatialRD_data.do – organize data for spatial discontinuity regressions.
- SpatialRDgraphs.do – Creation of spatial RD graphs: Figures 2, 4A and 4B.
- RD_discontinuity_graphs.do – Creation of graphs of the observables across the municipal / IOU boundary: Figure 1, Figures 3a – 3g.
- SpatialRDregs.do – Spatial RD regressions. Creation of Tables 3 and 4, Figures 5, 6 and 7 Appendix Tables A1 and A3, and Appendix Figures A3, A4 and A5.
- Data_GIS_placebo.do – cleans and merges data for the placebo analysis.
- Prepare_SpatialRD_data_placebo.do – organize data for placebo regressions.
- SpatialRDregs_placebo.do – Placebo regressions. Creation of Table 5.
- Experian_lease_extract.do – summarizes leasing information from Experian data at the CBG-year level.
- DMV_duration.do – Creates duration of ownership graphs and tables: Figure 8, Table 6 and Appendix Table A4.
- Welfare_Policy.do – Creates Figures 9A, 9B and 9C.

R code:

- Distance_data_bg.R – Creates distances between pairs of census block group centroids and creates distances between census block group centroids and utility area boundaries.

List of tables and programs

The code reproduces all figures in the paper.

Figure/Table #	Program	Line Number
Table 1	summary_statistics_table.do	127
Table 2	PanelRegs.do	208
Table 3	SpatialRDregs.do	242
Table 4	SpatialRDregs.do	253
Table 5	SpatialRDregs_placebo.do	157
Table 6	DMV_duration.do	252
Figure 1	RD_discontinuity_graphs.do	156
Figure 2	SpatialRDgraphs.do	38
Figure 3a – 3g	RD_discontinuity_graphs.do	140 - 148
Figure 4A	SpatialRDgraphs.do	150
Figure 4B	SpatialRDgraphs.do	156
Figure 5	SpatialRDregs.do	388
Figure 6	SpatialRDregs.do	525
Figure 7	SpatialRDregs.do	703
Figure 8	DMV_duration.do	159
Figure 9a – 9c	Welfare_Policy.do	78, 107, 136
Appendix Table A1	SpatialRDregs.do	548
Appendix Table A2	PanelRegs.do	228, 237
Appendix Table A3	SpatialRDregs.do	742
Appendix Table A4	DMV_duration.do	306
Appendix Figure A1	Data_Electric.do	1266
Appendix Figure A2	Data_Electric.do	1266
Appendix Figure A3	SpatialRDregs.do	395, 402
Appendix Figure A4	SpatialRDregs.do	535, 543
Appendix Figure A5	SpatialRDregs.do	212

Data References

Energy Information Administration, Annual Electric Power Industry Report, Form EIA-861, https://www.eia.gov/electricity/data/eia861/

Oil Price Information Service (OPIS), https://www.opisnet.com/product/pricing/oil-prices-forecast-and-history/retail-datahouse/

OpenEI, U.S. Utility Rate Database, https://apps.openei.org/USURDB/

US Department of Energy, Alternative Fuel Data Center, Alternative Fueling Station Locator, https://afdc.energy.gov/stations#/find/nearest

US Census Bureau, American Community Survey, https://www.census.gov/data/developers/data-sets/acs-5year.html