

Description of the data and codes for the paper “Regional Heterogeneity and U.S. Presidential Elections: 2020 Real-Time Forecasts and Evaluation” by Rashad Ahmed and M. Hashem Pesaran

Date: May 2021

Data:

1. microfit_county_panel_2020_08_28.csv - County-level panel data set of voting, socioeconomic and demographic variables from 2000-2016.
2. TO_activeset.csv - Active set variables for the voter turnout equation.
3. dLRO_activeset.csv - Active set variables for the change in log Republican odds equation.
4. 2020_act_TO_OCT.csv - Active set variables for the voter turnout equation with data updated through October 2020 to forecast 2020 voter turnout.
5. 2020_act_dLRO_OCT.csv - Active set variables for the change in log Republican odds equation with data updated through October 2020 to forecast 2020 change in log Republican odds.

Code:

1. OCMT_function_public.R - R code to run OCMT algorithm used to forecast 2020 presidential election using pooled-OCMT and regional-OCMT approaches.
2. 2020_OCMT_forecast_recursive_public.Rmd - R code which generates 2020 presidential election forecasts using pooled-OCMT and regional-OCMT approaches.
3. 2020_LASSO_forecast_recursive_public.Rmd - R code which generates 2020 presidential election forecasts using pooled-Lasso and regional-Lasso approaches.

How to generate 2020 forecasts using LASSO and OCMT approaches:

First, make sure all code and data files are stored in the same directory and all necessary packages are loaded in the R workspace. Make sure to run a recent or the latest R version.

To generate Lasso forecasts, run the script “2020_LASSO_forecast_recursive_public.Rmd”. This produces 2020 forecasts under both pooled-Lasso and regional-Lasso. The “popular_preds” variable returns the 2-party U.S. mainland national vote predictions, and “tab_final” provides state-level forecasts of Republican vote share and electoral votes. To generate OCMT forecasts, run the script “2020_OCMT_forecast_recursive_public.Rmd”. The same “popular_preds” and “tab_final” variables return respective 2020 forecasts.