Econ 361: Advanced Econometrics

"Election Polls"

2020 Presidential Election Polls

$$T_i = \begin{cases} 1 & \text{if } i \text{ supports Trump} \\ 0 & \text{otherwise} \end{cases} \qquad B_i = \begin{cases} 1 & \text{if } i \text{ supports Biden} \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Let } Y_i = T_i - B_i$$

Election Poll s is essentially $\bar{Y}_s = \frac{1}{N_s} \sum_{i=1}^{N_s} Y_i$ which is the same as the OLS predictor of Y when the only regressor is a constant (a naive OLS predictor of Y)

Now consider a **sample of polls**: $\{ \bar{Y}_s \}_{s=1}^{S}$

How can we use such a sample to arrive at a better predictor of Y, a better estimator of $E[Y]$ which is BP_{MSE} in the naive case

Think **GLS** and **SUR** Models.

2020 Presidential Election Polls

1. Do all election polls contain the same quality of information? Are the information contained in one election poll distinct from that in other election polls?

   - Possible violation(s) of the Spherical Errors assumption

   - Without such violation(s), we could simply use OLS – i.e. simple sample average of the Election Polls

2. What about potential bias in the election polls?

   - Spherical Errors may not be the only G-M assumptions violated

3. Should we look at each state separately? Are election polls from one state also informative about election polls/voting in other states?

   - Cross "equation" parameter restrictions/correlations ?