There are several famous failures of federal election forecasts. In 1936, the popular magazine Literal Digest incorrectly predicted that Franklin Roosevelt would lose, based on a two-million responses mail-in survey [@Forecasting]. In the 2016 United States (US) federal election, most institutions failed to predict that Donald Trump won against Hillary Clinton. Those failed forecasts suggest that the pool of respondents can be highly biased, which would be detrimental to the results of the predictions. Hence, representative polls have been used extensively to make election predictions, where randomly sampled individuals are asked who they would like to vote for. Although this approach has been proved to be effective, it is faced with several challenges: 1) the cost of time and money is becoming enormous; 2) the response rates are keeping decreasing in the past several decades.

As a result, people are seeking alternatives, such as using non-representative polls, together with the benefits of proper statistical adjustment. With this approach, one could make accurate predictions based on faster and less expensive survey sampling methods.

Post-stratification is a well-known method, where the core technique is to partition the population into cells. Particularly, the cells are based on combinations of different demographics such as age groups, states

multilevel regression with post-stratification (MRP)

**Shibboleths**

forecasting

Data

The framework used in this work is multilevel regression with post-stratification (MRP). In general, a model was generated firstly using non-representative survey data, then the post-stratification dataset from . Thus, two datasets were required by this framework and used in this work.

The first dataset was

The second dataset was

Figure 1 in the paper, a comparison of the demographics of the two datasets.

Age and sex are known to be strongly correlated with voting behaviors.

Discussions

The sex and gender problem in modern survey

Given random digit dialing (RDD) as an example, the response rates of RDD have declined to 9% in 2012 [@Forecasting].