Response to “Forecasting Elections with Non-Representative Polls”

The paper by Wang et al is an excellent example of how one can use non-representative polls to make accurate election predictions after careful post-stratification of the data.

Of the eight variables reported from the respondents, the three least representative variables appear to be age, sex, and education. State, race and 2008 vote appear to be most representative. This can be observed from Fig 1 of the paper, which provides us with a comparison of the data from the Xbox 2012 dataset and the 2012 electorate. The discrepancy from the actual population can be explained by the demography of people owning an Xbox and playing video games, as they would likely be male, between the ages of 18- 29, and college educated.

The 2008 exit poll data is used to create the post stratification weights. The authors also mention using Current Population Survey (CPS) data to attempt to create the post stratification weights, however it is missing important post stratification features, one of which is party identification. The 2012 exit poll data is also mentioned but not used, and the authors mention that while combining more data may create more accurate forecasting, it is not used so as to maintain transparency and simplicity.

Post-stratification is an important part of any kind of non-representative sampling. The authors themselves mention that the uncorrected Xbox data predicts a landslide victory for Mitt Romney, and we can see this happen in Fig 2. We use the date of October 22nd as our reference to start checking, and it is only 2 days between October 22nd and 29th that it remains uncertain, predicting a Romney win all the other days leading up to the election.

The pollster data is more uncertain, and predicts Obama before and continues to be uncertain for a bit before predicting Romney, and then finally going on to predict Obama right before the election.

The MRP adjusted Xbox data on the other hand strongly predicts an Obama victory the whole of the last 3 weeks.