

## Post-election Analysis of Presidential Election/Poll Data: Liars Continued to Lie But What Have Changed?

### Abstract:

We carry out analyses of the 2016 and 2020 U. S. presidential election data, including the pre-election polls and the actual elections. The analyses provide statistical evidence of discrepancy between the polls and real elections that is consistent across these two elections. Specifically, the polls had consistently over-estimated advantage of the Democratic candidates (or, equivalently, under-estimated the true population support of the Republican candidate, Donald Trump. The analysis is done at the state level, which is realistic given the U. S. election system, using small area estimation methods. Our analyses have found consistent patterns suggesting that the polls have been under estimating the Republican candidate, especially in the states that mattered (the battleground states). The analysis also suggests improvement in the 2020 polls in terms of the underestimation. Furthermore, we show that a small-area model built upon data from one election can provide better prediction than the polls about outcomes of another election involving the same Republican candidate. This work is joint with Yuanyuan Li of University of California, Davis and Peter X.K. Song of University of Michigan.