What will happen in Indonesia's 2019 legislative election?

Seth Soderborg
Department of Government, Harvard University
soderborg@g.harvard.edu

In Indonesia's multiparty system, it is difficult to turn polling results into predictions about seats in parliament. This is a problem for observers interested in how upcoming elections might affect political outcomes. My research covers one way to do solve the problem of predicting seats from votes in the Indonesian context.

Adapted from, Soderborg, Seth. 2018. "Predicting Multiparty Elections with Limited Data: A Bayesian Model Applied to Indonesia." Draft manuscript.

Preliminary results:

- Current national polling averages give PDI-P an extensive lead over all other parties in the race for parliament. An unweighted polling average as of April 22, 2018 gives PDI-P 23 percent of the vote, Golkar 12 percent, Gerindra 12 percent, Demokrat 6 percent, and PKB 5 percent.
- The new parliamentary threshold of four percent is likely to prevent all four new parties (Perindo, PSI, Berkarya, and Garuda) from entering parliament.
- Four incumbent parties are in danger of failing to enter parliament: NasDem, PPP, PAN, and Hanura. PBB, which is returning to national competition but failed to cross the 2014 threshold, is also unlikely to meet the minimum standard.
- In fifty percent of simulations, parties win at least this many[1] seats: PDI-P, 149; Golkar, 107; Gerindra, 99; Demokrat, 53; PKB, 44; PKS, 33; PPP, 23; NasDem, 23; PAN, 23.
- New election rules privilege Golkar and Gerindra and punish PDI-P. Although PDI-P can expect to win more seats than it did in 2014, its gains will be severely limited by the introduction of the Saint-Lagüe quota system for allocating seats.
- Despite the reduced number of parties expected to hold seats in parliament, the distribution of power within the parliament will not change as dramatically as the vote share. The effective number of political parties is estimated to remain near 6.3.

Why are multiparty elections hard to predict?

Multiparty elections tend to be more unstable, historically, than two-party elections, and the presence of multiple parties means modelers have to use stronger statistical assumptions. After those assumptions are made, the modeler has to apply whichever seat-allocation algorithm will be used to turn votes into seats. This produces a single number, without an accompanying estimate of uncertainty. Good predictions include uncertainty estimates. I use bootstrapped simulations to incorporate uncertainty estimates into my seat total predictions.

[1] The model was recently altered to include 15 additional seats in three new districts, following announcements made last week. The 2019 parliament will have 575 members elected in 80 districts.

What should I tell my boss about the Saint-Lagüe quota?

While your boss might be interested in the fact that the Saint-Lagüe quota allocates seats according to the equation $seats_i = \frac{V}{2_{s(i-1)+1}}$, the important thing to know about this equation is that the number in the denominator, the two, makes it very hard for any party to win more than three seats in any legislative district. That number penalizes parties that win very large shares of the vote in a district. That two makes a second seat three times harder to win than a first seat, a third seat five times harder, and a fourth seat seven times harder. The only party that does very well in more than a handful of districts is PDI-P. For that reason, they will bear the brunt of this rule change.

What useful things can this model do?

In addition to predicting the number of seats each party will win in the parliament as a whole, this model can give probabilistic estimates of the number of seats each party will win in each electoral district. This model can also change in response to new data. It can even use data other than polls and past elections, although the model currently does not incorporate those sources.

How does this model work?

The model begins with a weighted average of national polls (the weights come from the pollsters' past performance—think Nate Silver and 538). This provides an estimate of each parties' nationwide popularity. The model then looks at the difference in vote share won by each party in each legislative district and each parties' national vote share in previous elections. This value, *national vote share – district vote share* is the deviation. The model uses the deviation from 2014 as the key ingredient in each district's estimated vote share, as well as the historical variance in the deviation, as an estimate of uncertainty. Each district is then modeled by taking the current national polling average and the 2014 deviation to get an estimated mean vote share in the district in 2019. The model then generates thousands of possible vote scenarios that correspond to randomized versions of these mean vote shares. For each scenario, the model estimates the number of seats each party wins in each district. The marquee predictions of this model are the median of the resulting distribution of numbers of seats won.