

State Space Models in Analyzing Big Election Data

Wednesday, March 16, 2016 4:40 PM (20 minutes)

This is a proposal to present a paper at the 2016 ISGC Meeting in Academia Sinica, Taiwan ROC. This study employs new methods in analyzing large amount of poll data from all survey firms to detect so called house and mode effects in election surveys. We collect all polling results and identify house (survey firm) and mode (telephone, internet, in-person) effects on British parties' vote shares in nearly 2000 public opinion polls conducted between June 2010 and May 6, 2015. The 2015 British general election resulted in the Conservatives winning a majority of seats in parliament. This outcome was surprising because all major polling houses failed to forecast the parties' vote shares accurately. The widely publicized failure to get the election polling right has reinvigorated debates concerning the accuracy of various survey modes and the British Polling Council is conducting an inquiry concerning what went wrong.

We will estimate house and mode effects using a Bayesian dynamic factor model. The model, pioneered by Jackman (2005), is a state space specification whereby a party's vote intention share on any day t is a latent variable measured by observed polls conducted by one or more survey houses. A Bayesian approach to this state space model is attractive because it easily accommodates polls conducted by various houses at irregular intervals. We extend the model in three ways:

- estimate mode effects as well as house effects. This will provide evidence relevant to the debate about the accuracy of various survey modes—internet, telephone, in-person—currently used for political polling;
- reconceptualize house effects as dynamic entities, the magnitudes of which can vary over time. This will enable us to estimate the magnitude of possible 'herding' effects whereby the vote share estimates of various polling firms converge in the run-up to election day;
- use the dynamic factor model to forecast parties' vote shares in the 2015 election using various time horizons (one week, one month, three months, six months). In addition unconditional forecasts, we will condition on a second factor that measures the public economic mood over the June 2010 - May 2015 period. Survey data indicate that economic evaluations improved markedly in the year preceding the election and individual-level analyses of national election survey data (Clarke et al., 2015, forthcoming) indicate that economic judgments had strong effects on the choices voters made. These results suggest that a time series forecasting model that conditions on the dynamics of the public economic mood may perform relatively well. We will extend this technique to other polling data including the case in Taiwan.

Primary authors: Prof. WENG, Dennis (SUNY Cortland); Prof. CLARKE, Harold D. (University of Texas at Dallas); Prof. HO, Karl (University of Texas at Dallas)

Presenter: Prof. HO, Karl (University of Texas at Dallas)

Session Classification: Humanities, Arts, and Social Sciences Session II

Track Classification: Humanities, Arts, and Social Sciences (HASS) Applications