PROJECT SUMMARY

In order to facilitate better scientific modeling of social phenomena (e.g., election outcomes or domestic crises), we propose to extend a promising statistical method and to develop software that will aid scholars across disciplines to make more accurate forecasts of future events. This project builds on work in the fields of meteorology and statistics to (1) extend the method for application to a wider array of outcomes (e.g., binary data), (2) provide freely available software that implements both maximum likelihood and Bayesian estimation techniques, and (3) provide accessible explanations of the method and social science applications.

Ensemble Bayesian model averaging (EBMA) improves prediction by pooling *all* information from multiple forecasting models to generate ensemble predictions similar to a weighted average of component forecasts. The weight assigned each forecast is based on its predictive accuracy and precision. Thus, EBMA differs from prior work in the social sciences that relies on simple averages of component forecasts and neither calibrates weights based on model performance nor accounts for predictive uncertainty.

The aim of EBMA is not to choose some "best" model, but rather to incorporate the insights and knowledge implicit in various forecasting efforts via statistical post-processing. These component models can be diverse and need not share covariates, functional forms, or error structures. Indeed, the components may not even be statistical models, but may be predictions generated by agent-based models or subject-matter experts. In practice, the method provides superior predictive power relative to any single component and reduces the likelihood of dramatic miss-predictions as it is not reliant on any one data source or methodology.

Intellectual merit: This project will address several substantive questions in political science, with a focus on the fields of international relations, where policymakers have a particular demand for improved forecasting, and elections. Prediction of socially determined outcomes has recently gained favor in political science, but to date all efforts have been directed at finding the "best" models, rather than making the best predictions. In addition, the proposed project will include research on the comparative value of the various forecast comparison metrics for applied analysis in the social science. It will also provide basic research into prior structures that penalize model complexity and ensure that ensemble forecasts reflect uncertainty in component forecasts, data vintage, and missingness. EBMA has received attention in the fields of statistics, meteorology, and (to a lesser extent) economics. Yet, it has not been advanced in the methodological directions herein proposed.

Broader impacts: Testing systematic predictions about future events against observed outcomes is generally seen as the most stringent validity check of statistical *and* theoretical models. Yet, political scientists rarely make predictions about the future; instead, they typically focus on developing and validating theories that explain past events. However, research in political science could gain immensely in its policy relevance if predictions were more common and more accurate. Improved forecasting of important political events would make research more germane to policymakers and the general public who may be less interested in explaining the past than anticipating and altering the future. From a scientific standpoint, greater attention to forecasting would facilitate stringent validation of statistical models since truly causal models should perform better in out-of-sample forecasting. More widespread usage of prediction heuristics should also help anneal theoretical models.

This project will further develop statistical techniques that improve the capabilities of researchers across disciplines to produce more accurate forecasts of future events. The methodological advancement made in this project will be available to the social science community and influence research agendas in multiple fields. This is particularly important as existing research projects and software packages for ensemble forecasting are narrowly tailored to the needs of weather forecasting.

Finally, the principal investigators are active members of the research community at the intersection of statistics and the social sciences. Their work is widely read in political science and other disciplines. Further, at least two graduate students in political science and/or statistical science will be included in the research and will gain experience in large-scale research projects.