

# SAY YES TO THE GUESS: TAILORING ELEGANT ENSEMBLES ON A TIGHT (DATA) BUDGET

Jacob M. Montgomery <sup>1</sup>    Florian M. Hollenbach <sup>2</sup>    Michael D.  
Ward <sup>2</sup>

<sup>1</sup>Department of Political Science  
Washington University in St. Louis

<sup>2</sup>Department of Political Science  
Duke University

August 31, 2012

# Ensemble Bayesian Model Averaging (EBMA)

EBMA is a principled way of combining predictions to increase out-of-sample performance

- Models are weighted based on calibration period
- EBMA is a finite mixture model

# Ensemble Bayesian Model Averaging (EBMA)

EBMA is a principled way of combining predictions to increase out-of-sample performance

- Models are weighted based on calibration period
- EBMA is a finite mixture model

## Predictive EBMA PDF

$$p(y|f_1^{t^*}, \dots, f_K^{t^*}) = \sum_{k=1}^K w_k g_k(y|f_k^{t^*})$$

# Ensemble Bayesian Model Averaging (EBMA)

- $w_k$  estimated using EM algorithm

# Ensemble Bayesian Model Averaging (EBMA)

- $w_k$  estimated using EM algorithm
- Increased weight given for

# Ensemble Bayesian Model Averaging (EBMA)

- $w_k$  estimated using EM algorithm
- Increased weight given for
  - ▶ Accuracy

# Ensemble Bayesian Model Averaging (EBMA)

- $w_k$  estimated using EM algorithm
- Increased weight given for
  - ▶ Accuracy
  - ▶ Uniqueness

# Adjusting EBMA to Forecasting in the Social Sciences

Forecasting efforts in the social sciences are often hampered by data problems

- Missingness
- Non-random missingness
- Many forecasts, few observations



# Predicting the Incumbent Voteshare in the 2012 Presidential Election

	F	A	C	H	LBRT	L	Hol	EW	Cuz
1992	55.7	46.3	49.7	48.9	47.3				
1996	49.5	57.0	55.5	53.5	53.3		57.2	55.6	
2000	50.8	53.2	52.8	54.8	55.4	60.3	60.3	55.2	
2004	57.5	53.7	52.8	53.2	49.9	57.6	55.8	52.9	51.1
2008	48.1	45.7	52.7	48.5	43.4	41.8	44.3	47.8	48.1

# Adjusting EBMA to Forecasting in the Social Sciences

Our solutions so far:

- Fraley, Raftery and Gneiting (2010) approach to missing data

# Adjusting EBMA to Forecasting in the Social Sciences

Our solutions so far:

- Fraley, Raftery and Gneiting (2010) approach to missing data
- Introduce a “wisdom of the crowds” tuning parameter  $c \in [0, 1]$

# Adjusting EBMA to Forecasting in the Social Sciences

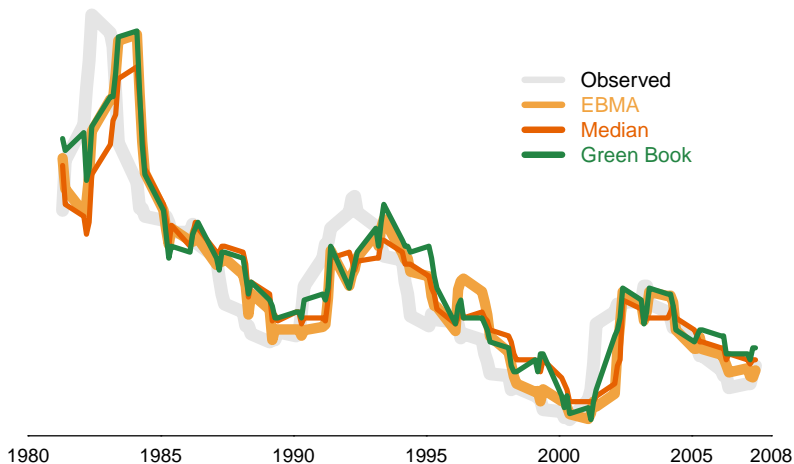
Our solutions so far:

- Fraley, Rafery and Gneiting (2010) approach to missing data
- Introduce a “wisdom of the crowds” tuning parameter  $c \in [0, 1]$
- A minimum ( $\frac{c}{k}$ ) probability that each the observation is “best” predicted by each  $k$  model.

# Predicting Quarterly Unemployment in the US

- Predictions from the Survey of Professional Forecasters (SPF)
- Green Book of the Fed
- Predictions four quarters into the future
- Rolling calibration window of ten quarters
- Minimum of five forecasts in the calibration period for model inclusion

# Observed and Forecasted U.S. unemployment (1981-2007)



# Comparing Forecasts of U.S. Unemployment – fit statistics summary

	MAE	RMSE	MAD	RMSLE	MAPE	MEAPE	MRAE	PW
c=0	0.54	0.74	0.37	0.093	8.37	6.49	<b>0.73</b>	<b>27.36</b>
c=0.05	<b>0.54</b>	0.74	<b>0.37</b>	<b>0.093</b>	<b>8.33</b>	<b>6.30</b>	0.75	<b>27.36</b>
c=0.1	0.54	0.74	0.35	0.093	8.40	6.44	0.76	28.30
c=1	0.61	0.80	0.46	0.102	9.72	8.92	0.95	46.23
GB	0.57	<b>0.73</b>	0.43	0.093	9.37	8.81	1.00	45.28
Median	0.62	0.81	0.47	0.103	9.83	8.87	0.98	47.17
Mean	0.61	0.80	0.46	0.102	9.71	9.06	0.93	46.23

# Predicting the Incumbent Voteshare in the 2012 Presidential Election – Calibration Sample

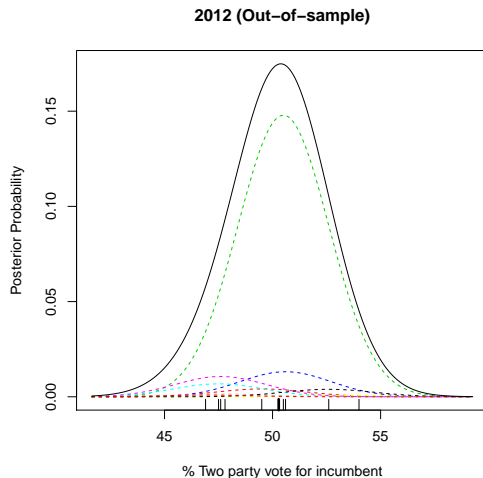
	F	A	C	H	LBRT	L	Hol	EW	Cuz
1992	55.7	46.3	49.7	48.9	47.3				
1996	49.5	57.0	55.5	53.5	53.3		57.2	55.6	
2000	50.8	53.2	52.8	54.8	55.4	60.3	60.3	55.2	
2004	57.5	53.7	52.8	53.2	49.9	57.6	55.8	52.9	51.1
2008	48.1	45.7	52.7	48.5	43.4	41.8	44.3	47.8	48.1



# Model Weights and In-Sample Fit Statistics for EBMA Model of U.S. Presidential Elections (1992-2008)

	EBMA Weight	RMSE	MAE
EBMA		1.92	1.56
Fair	0.02	5.53	4.58
Abramowitz	0.78	2.02	1.72
Campbell	0.07	3.46	2.88
Hibbs	0.04	2.68	2.44
Lewis-Beck, Rice, and Tien	0.06	2.78	2.28
Lockerbie	0.00	7.33	6.97
Holbrook	0.01	5.73	4.77
Erikson and Wlezien	0.02	2.74	2.25
Cuzà	0.00	1.27	0.95

# Predictive Ensemble PDFs of Incumbent-Party Vote Share



# Future Work

- Simulation study to show over-fitting problem

# Future Work

- Simulation study to show over-fitting problem
- Alternative methods for handling missing forecasts

# Future Work

- Simulation study to show over-fitting problem
- Alternative methods for handling missing forecasts
- Estimate  $c$  parameter