Say Yes to the Guess: Tailoring Elegant Ensembles on a Tight (Data) Budget

Jacob M. Montgomery [1] Florian M. Hollenbach [2] Michael D. Ward [2]

[1]Department of Political Science
Washington University in St. Louis

[2]Department of Political Science
Duke University

January 17, 2013

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

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

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, Rafery and Gneiting (2010) approach to missing data

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 $c \in [0, 1]$

Adjusting EBMA to Forecasting in the Social Sciences

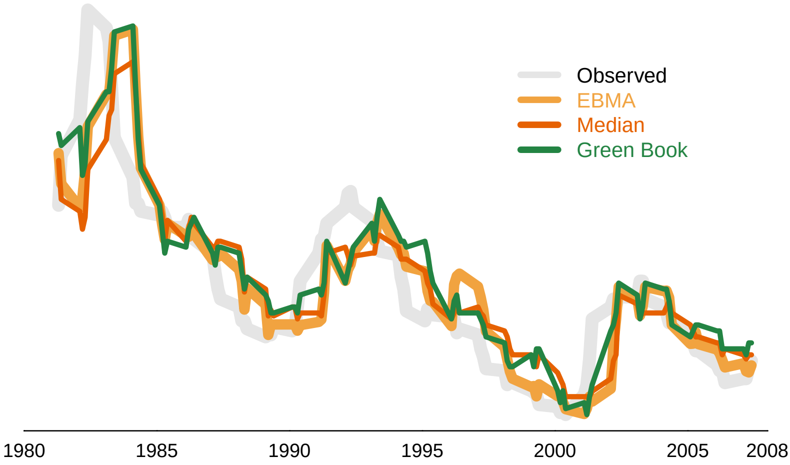
Our solutions so far:

- 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	**0.73**	**27.36**
c=0.05	**0.54**	0.74	**0.37**	**0.093**	**8.33**	**6.30**	0.75	**27.36**
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	**0.73**	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

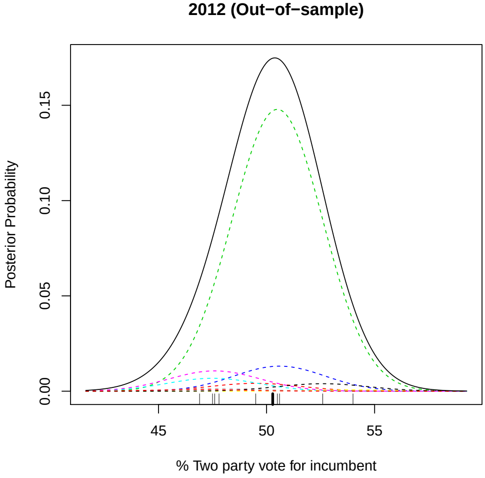
Predicting the Incumbent Voteshare in the 2012 Presidential Election – Calibration Sample

	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
Cuzàn	0.00	1.27	0.95

Predictive Ensemble PDFs of Incumbent-Party Vote Share

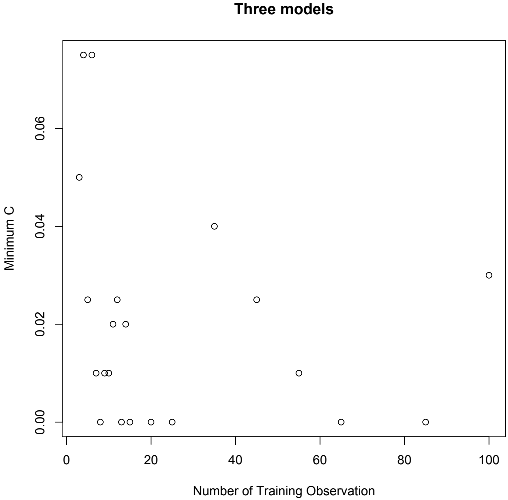


Matrix Fun

Table: Number of Models vs. Training Set Obs

	3	5	7	9	11	13	15
3	0.05	0.03	0.03	0.30	0.30	0.50	0.30
4	0.07	0.02	0.15	0.15	0.50	0.50	0.30
5	0.03	0.00	0.02	0.15	0.20	0.30	0.20
6	0.07	0.07	0.00	0.00	0.10	0.07	0.50
7	0.01	0.01	0.03	0.00	0.00	0.15	0.10
8	0.00	0.00	0.03	0.03	0.15	0.00	0.30
9	0.01	0.04	0.00	0.03	0.03	0.05	0.15
10	0.01	0.01	0.05	0.00	0.00	0.03	0.15
11	0.02	0.04	0.03	0.05	0.00	0.00	0.00
12	0.03	0.01	0.00	0.01	0.00	0.01	0.02
13	0.00	0.03	0.04	0.03	0.05	0.04	0.10
14	0.02	0.00	0.01	0.00	0.01	0.10	0.05
15	0.00	0.00	0.03	0.00	0.02	0.15	0.03
20	0.00	0.04	0.05	0.00	0.02	0.00	0.01
25	0.00	0.01	0.07	0.00	0.00	0.03	0.01
35	0.04	0.01	0.01	0.02	0.03	0.03	0.07
45	0.03	0.01	0.03	0.01	0.00	0.00	0.03
55	0.01	0.05	0.03	0.03	0.03	0.03	0.01
65	0.00	0.03	0.00	0.00	0.01	0.00	0.03
85	0.00	0.03	0.02	0.01	0.02	0.04	0.03
100	0.03	0.00	0.00	0.05	0.00	0.00	0.05

An Example



Three models
Minimum C
0.00
0.02
0.04
0.06
0
20
40
60
80
100
Number of Training Observation