Prediction Using Several Macroeconomic Models

Gianni Amisano
European Central Bank[1]
John Geweke
University of Technology Sydney, Erasmus University, and
University of Colorado

European Central Bank

May 5, 2012

[1] The views expressed do not represent those of the ECB.

Summary

- Motivation: How to arrive at a single predictive distribution from several, knowing that they are all wrong
- This work uses predictive distributions from three mainstream models and a canonical data set
  - Models: DFM, DSGE, VAR
  - Data set: 7 aggregates of Smets and Wouters (2007), updated and extended through 2011
- Methods used: Pooling, Analysis of predictive variance, Probability integral transforms
- Improvements in prediction:
  - Increment 1, for model predictive distributions: 50%
  - Increment 2, for the single predictive distribution: 42%
  - Combined: 113%

Using models in prediction

- All models have fully specified proper prior distributions and likelihood functions
- Estimation and prediction
  - Posterior mode (PM): Substitute mode for unknown parameter vector
  - Full Bayes (FB): Simulate parameters from posterior
  - In each case, simulate the future conditional on the parameter vector(s)

Three models

- Dynamic factor model of Stock and Watson (2005)
  - 5 additional variables, for a total of 12
  - 3 factors
  - Factors and idiosyncracies are all AR(2)
  - 99 free parameters
- Dynamic stochastic equilibrium model of Smets and Wouters (2007)
  - Conventional linearized solution
  - 39 free parameters
- Vector autoregression model of Sims (1980)
  - Differenced (VARD) and levels (VARL) variants
  - Models use Minnesota prior
  - 231 free parameters
- Denote the models A_i $(i = 1, 2, 3)$

Data

- US quarterly, 1951 - 2011, revisions as of February 16, 2012
- Series:
  - Growth rates in real per capita consumption, investment, GDP, hours worked
  - Hours worked index, GDP inflation, Fed funds rate
- Additional series (DFM):
  - S&P 500 growth rate
  - Civilian unemployment rate
  - 3 month / 10 year Treasury return differential
  - BAA / AAA return differential
  - Growth rate in money supply M2

Procedures

- For quarter t : 1966:1 - 2011:4, y_t is the 7×1 observed vector of aggregates
- Y_t *ex ante;* y_t *ex post* (data)
- Use data through quarter $t-1$, denoted $y_{1:t-1}$
  - For $i = 1, 2, 3$ formulate predictive densities
    - Posterior mode (PM): $p\left(y_t \mid \widehat{\theta}_i\left(t-1\right), y_{1:t-1}, A_i\right)$
    - Full Bayes (FB): $p\left(y_t \mid y_{1:t-1}, A_i\right)$
  - Evaluate
    - $p\left(y_t \mid \widehat{\theta}_i\left(t-1\right), y_{1:t-1}, A_i\right)$
    - $p\left(Y_t \mid y_{1:t-1}, A_i\right)$

Evaluations and comparisons

- All of our evaluations and comparisons are based on log scores
  - $LS(A_i; PM) = \sum_{t \in \text{period}} \log p\left(y_t \mid \widehat{\theta}_i(t-1), y_{1:t-1}, A_i\right)$
  - $LS(A_i; FB) = \sum_{t \in \text{period}} \log p\left(y_t \mid y_{1:t-1}, A_i\right)$
- In this presentation, the period is always the full $T = 184$ quarters 1966:1 - 2011:4.
- Interpretations
  - The geometric average of the probability (density) assigned to what actually occurred is

$$\exp\left[LS(A_i; PM) / T\right], \quad \exp\left[LS(A_i; FB) / T\right]$$

  - $LS(A_i; FB)$ is the
    - Log predictive likelihood for 1966:1 - 2011:4 conditional on 1951:1 - 1965:4
    - Log marginal likelihood treating 1951:1 - 1965:4 as part of the prior

Model comparison

- Results:

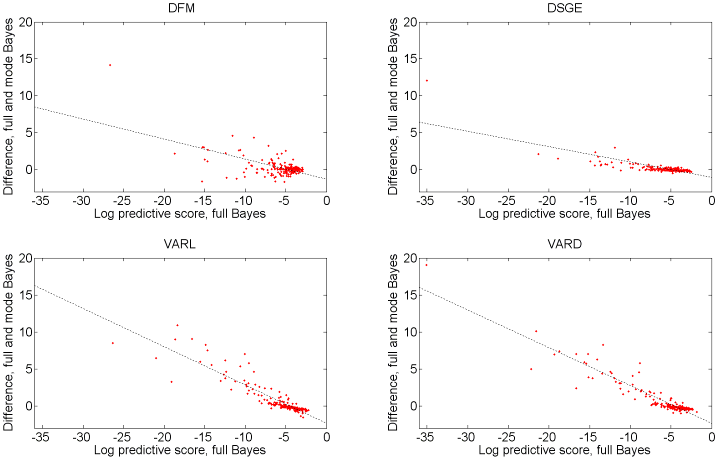
	Log scores		
Model	FB	PM	FB − PM
DFM	-1083.86	-1135.10	51.24
DSGE	-1097.03	-1128.23	31.20
VARD	-1122.43	-1265.46	143.03
Mean	-1176.26	-1101.11	75.15

- Across the three models, geometric mean improvement in FB over PM:

$$\exp\left[75.15/184\right] - 1 = 0.504.$$

- Bayes factor for DFM over DSGE:

$$\exp\left(1097.03 - 1083.86\right) = 5.243 \times 10^5$$


DFM
DSGE
VARL
VARD
Difference, full and mode Bayes
Log predictive score, full Bayes
20
15
10
5
0
-35
-30
-25
-20
-15
-10
-5
0

Model assessment using the probability integral transform

		Inverse cdf at statistic $\mid A_i = D$		
Series	Model	Moments	Autocorrelation	Joint
Consumption	DFM	0.0000	0.4077	0.0001
growth	DSGE	0.0018	0.0040	0.0007
	VARD	0.0003	0.9441	0.0005
Hours	DFM	0.1566	0.0591	0.0651
worked	DSGE	0.0157	0.0107	0.0060
index	VARD	0.3111	0.3340	0.3225
Fed funds	DFM	0.0000	0.0000	0.0000
rate	DSGE	0.0000	0.0000	0.0000
	VARD	0.0000	0.0000	0.0000

Prediction pools

- Predictive densities from alternative models:

$$p\left(Y_t; y_{t-1}, A_i\right) \quad (i = 1, \ldots, n)$$

  - Here, $n = 3$ and $p\left(Y_t; y_{t-1}, A_i\right) = p\left(Y_t \mid y_{t-1}, A_i\right)$.
- A one-step-ahead prediction pool at time t is

$$p\left(Y_t; y_{1:t-1}, w_{t-1}, A_1, \ldots, A_n\right) = \sum_{i=1}^{n} w_{t-1,i}\, p\left(Y_t \mid y_{t-1}, A_i\right).$$

  - The notation w_{t-1} emphasizes the requirement that the prediction pool cannot depend on future data y_{t+s} $(s \geq 0)$.
  - Because Y_t is a vector, the pool must be linear (McConway, 1981).
  - The weight vectors w_{t-1} belong to the $n-$dimensional unit simplex:

$$w_{t-1,i} \in [0,1] \quad (i = 1, \ldots, n), \quad \sum_{i=1}^{n} w_{t-1,i} = 1$$

Three particular pools

- *Equally weighted pool*: $w_{t-1,i} = 1/n \ (i = 1, \ldots, n)$
- *Bayesian model averaging*: The pool is $p\left(Y_t \mid y_{1:t-1}\right) \iff$

  $w_{t-1,i} = p\left(A_i \mid y_{t-1}\right) \propto p\left(A_i\right) \cdot p\left(y_{1:t-1} \mid A_i\right) \quad (i = 1, \ldots, n)$.

  - If $p\left(A_i\right) = 1/n \ (i = 1, \ldots, n)$, these are the Bayes factors

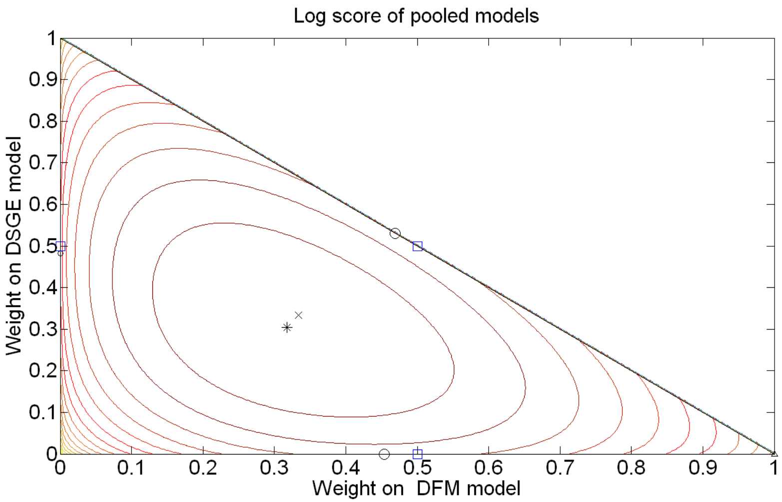
$$p\left(y_{1:t-1} \mid A_i\right) = \prod_{s=1}^{t-1} p\left(y_s \mid y_{1:s-1}, A_i\right);$$

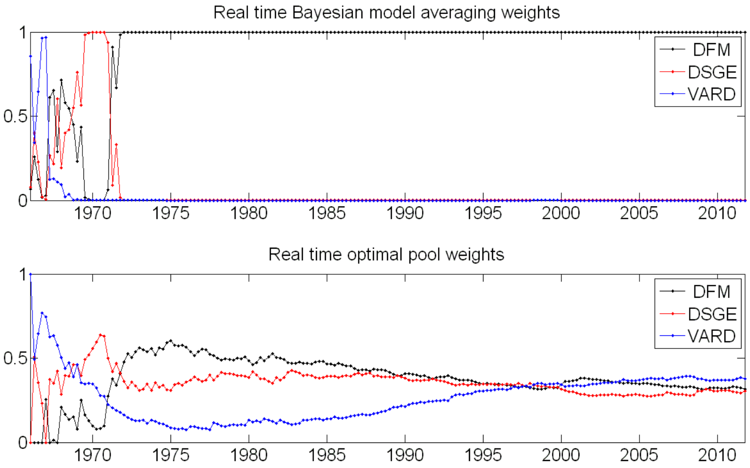
$$p\left(y_t \mid y_{1:t-1}, A_i\right) \approxeq M^{-1} \sum_{m=1}^{M} p\left(y_t \mid y_{1:t-1}, \theta_i^{(m)} A_i\right)$$

  where $\theta_i^{(m)} \sim p\left(\theta_i \mid y_{1t-1}, A_i\right)$

- Optimal pooling

$$w_{t-1} = \arg\max_{w} \sum_{s=1}^{t-1} \log\left[\sum_{i=1}^{n} w_i p\left(y_s \mid y_{1:s-1}, A_i\right)\right]$$


Log score of pooled models
Weight on DSGE model
Weight on DFM model
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


Real time Bayesian model averaging weights
DFM
DSGE
VARD
1
0.5
0
1970
1975
1980
1985
1990
1995
2000
2005
2010
Real time optimal pool weights
DFM
DSGE
VARD
1
0.5
0
1970
1975
1980
1985
1990
1995
2000
2005
2010

Log scores of models and pools

Models:	Log scores
DFM	-1083.86
DSGE	-1097.03
VARD	-1122.43
Mean over models:	-1101.11
Pools:	
Bayesian model averaging	-1084.96
Real time optima	-1043.41
Equally weighted	-1036.72

Summary

- Metric: Percent increase in the geometric mean predictive probability assigned to y_t one quarter before
- *Conclusion 1*, Use the predictive distribution rather than plug in the posterior mode.

Dynamic factor model	32.1%
Dynamic stochastic general equilibrium model	18.5%
Vector autoregression (in differences)	117.6%

- *Conclusion 2*. Pool, but don't Bayesian model average.

Bayesian model averaging	9.2%
Real time optimal	36.8%
Equally weighted	41.9%

(Improvements are relative to the geometric mean predictive probability taken over all models.)