DEUTSCHE BUNDESBANK
EUROSYSTEM

Freie Universität Berlin

VIESSMANN EUROPEAN RESEARCH CENTRE
LAURIER

Conference on

Forecasting and Monetary Policy

Berlin, 23-24 March 2009

Kenneth F. Wallis
University of Warwick

Discussion of „Density Combination at the Norges Bank“

Conference on "Forecasting and Monetary Policy"

Discussion of "There is more than one weight to skin a cat: Combining densities at Norges Bank"

by Bjornland, Gerdrup, Jore, Smith and Thorsrud

Kenneth F. Wallis

Department of Economics, University of Warwick
http://www.warwick.ac.uk/go/kfwallis

Deutsche Bundesbank **23 March 2009**

Outline

- This paper reports an extensive investigation of different ways of constructing combined forecasts, considering both point forecasts and density forecasts for two Norwegian series

- It is found that you do better according to a particular performance measure to use combination weights based on that measure

- I first make some observations of a technical nature, in the course of a presentation of some of the statistical background to such an exercise

- I then say a couple of things I would have done differently

Combining point forecasts (Bates and Granger, 1969; Smith and Wallis, 2009)

Consider two point forecasts of y_t made *h* periods earlier, f_{1t} and f_{2t}

Forecast errors $e_{it} = y_t - f_{it}$ have variances σ_i^2 and covariance σ_{12}

Combined forecast: $f_{Ct} = kf_{1t} + (1-k)\, f_{2t}$

Optimal weight: $k_{opt} = \dfrac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}$

Combination using k_{opt} has expected squared error $\sigma_{opt}^2 \le \min\left(\sigma_1^2, \sigma_2^2\right)$

Other weights: “inverse MSE” $k' = \dfrac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} = \dfrac{\frac{1}{\sigma_1^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}$; simple average $k = \frac{1}{2}$

Estimated weights: $\text{asy var}\left(\hat{k}'\right) = (1-\rho)^2 \text{asy var}\left(\hat{k}_{opt}\right)$ where $\rho = \sigma_{12}/\sigma_1\sigma_2$

Combining density forecasts

(1) Linear opinion pool/ finite mixture distribution (Wallis, 2005)

Consider *n* individual density forecasts $f_i(y),\ i=1,\dots,n,$ for given *t* and *h*

Combined forecast: $f_C(y)=\sum_{i=1}^{n} w_i f_i(y)$ with weights $w_i \geq 0,\ i=1,\dots,n,\ \Sigma w_i = 1$

If $f_i(y)=N\left(\mu_i,\sigma_i^2\right)$ then $f_C(y)$ is a mixture of normals (Pearson, 1894)

For any individual densities, moments about the origin are mixed in the same way:

$$\mu_1' = \sum_{i=1}^{n} w_i\mu_i = \mu_C, \qquad \mu_2' = \sum_{i=1}^{n} w_i\left(\mu_i^2+\sigma_i^2\right)$$

hence the variance of f_C is $\sigma_C^2 = \mu_2' - \mu_1'^2 = \sum_{i=1}^{n} w_i\sigma_i^2 + \sum_{i=1}^{n} w_i\left(\mu_i-\mu_C\right)^2$

i.e. aggregate uncertainty = average individual uncertainty + disagreement

(2) Logarithmic combination (Bacharach, Hammond)

This is usually written as $f_C(y) = \dfrac{\prod f_i^{w_i}(y)}{\int \prod f_i^{w_i}(y)dy}$

If $f_i(y) = N\left(\mu_i, \sigma_i^2\right)$ then $f_C(y) = N\left(\mu_C, \sigma_C^2\right)$ where

$$\frac{\mu_C}{\sigma_C^2} = \sum_{i=1}^{n} w_i \frac{\mu_i}{\sigma_i^2} \quad \text{and} \quad \frac{1}{\sigma_C^2} = \sum_{i=1}^{n} w_i \frac{1}{\sigma_i^2}$$

In the case of equal weights, σ_C^2 is the harmonic mean of the individual variances; this is less than the arithmetic mean, which is less than the finite mixture variance

(3) Pragmatic combination (Wallis, 2005)

Combine moments as in the finite mixture distribution, and use a normal (or 2PN) distribution with these moments to produce interval forecasts and fan charts

Evaluating combined density forecasts

- For density forecasts we have no results like the Bates-Granger result for the expected squared error of the optimal point forecast combination
- For several scoring rules we have the result that the expected score of the combined density forecast is between that of the best and worst forecasts
- This suggests removing ("trimming") the worst forecasts from the collection before constructing combined forecasts
  - trim the set of forecasting models
  - for point forecasting, remove extremes from the set $\{ f_{it} \}$
- We have individual examples where the actual score of the "best" combination is better than that of the best forecast (in "Optimal prediction pools" by John Geweke and Gianni Amisano, October 2008)

A public health warning about forecast evaluation criteria

Some criteria are not sensitive to the use of the wrong information set

- point forecasts: tests of equality of in-sample and out-of-sample MSE
- density forecasts: tests of uniformity of probability integral transforms
  - any forecast that is the correct conditional distribution for the information on which it is based has uniform PITs
  - extreme example: the unconditional distribution

So we need additional criteria, and comparative testing (“encompassing”)

Some things I would do differently

- give a clear account of the statistical framework and definitions of all statistics employed
- make more use of the estimation sample results to rule out rubbish models
- separate the discussion of the point and density forecasting problems, since they are different problems; only then take a brief look at similarities and differences between the rankings of models/methods in the two problems
- give plots of the target series and selected competing forecasts, so that the practical significance of perceived differences can be assessed
- use more target series
- in computing inverse MSE weights for use in constructing forecast combinations, consider Stock and Watson's variations:
    - raise each 1/MSE term to the power ω: $\omega < 1$ shrinks weights towards equality, $\omega > 1$ increases the weight given to better-performing forecasts
    - increase weight on recent errors via discounted sum of squared errors