Available online at www.sciencedirect.com
ScienceDirect
International Journal of Forecasting 23 (2007) 1 – 13
www.elsevier.com/locate/ijforecast

Combining density forecasts ☆

Stephen G. Hall [a,b], James Mitchell [b,*]

[a] Department of Economics, Leicester University, United Kingdom
[b] National Institute of Economic and Social Research, United Kingdom

Abstract

This paper brings together two important but hitherto largely unrelated areas of the forecasting literature, density forecasting and forecast combination. It proposes a practical data-driven approach to the direct combination of density forecasts by taking a weighted linear combination of the competing density forecasts. The combination weights are chosen to minimize the 'distance', as measured by the Kullback–Leibler information criterion, between the forecasted and true but unknown density. We explain how this minimization both can and should be achieved but leave theoretical analysis to future research. Comparisons with the optimal combination of point forecasts are made. An application to simple time-series density forecasts and two widely used published density forecasts for U.K. inflation, namely the Bank of England and NIESR "fan" charts, illustrates that combination can but need not always help.

JEL classification: C53; E37

Keywords: Density forecasts; Uncertainty; Combining forecasts; Evaluating forecasts; Inflation forecasting

1. Introduction

Measures of uncertainty surrounding a "central tendency" (the point forecast) can enhance its usefulness; e.g. see Garratt, Lee, Pesaran, and Shin (2003). So-called "density" forecasts are being used increasingly since they provide commentators with a full impression of the uncertainty associated with a forecast; see Tay and Wallis (2000) for a review. More formally, density forecasts of inflation, say, provide an estimate of the probability distribution of its possible future values.

It is well established that combining competing individual point forecasts of the same event can deliver more accurate forecasts, in the sense of a lower root mean squared error (RMSE); e.g. see Stock and Watson (2004). The success of combination follows from the fact that individual forecasts may be based on misspecified models, poor estimation or non-stationarities; e.g. see Hendry and Clements (2004).

This paper takes the natural next step of considering density forecast combination, to-date a relatively unexplored area. This brings together two important but hitherto largely unrelated areas of the forecasting

☆ Thanks to John Haywood, Ken Wallis, Martin Weale, an Associate Editor and two anonymous referees for the helpful comments. We gratefully acknowledge the financial support from the ESRC (Award Reference: RES-000-22-0610).
* Corresponding author. Tel.: +44 207 654 1926.
E-mail address: j.mitchell@niesr.ac.uk (J. Mitchell).
URL: http:/www.niesr.ac.uk. (J. Mitchell).

0169-2070/$ - see front matter © 2006 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.ijforecast.2006.08.001

literature in economics, density forecasting and forecast combination. We propose a simple and practical data-driven approach to combine density forecasts directly, using what we call "optimal" weights. These optimal weights are those weights that minimize the 'distance', as measured by the Kullback–Leibler information criterion [KLIC], between the combined forecast density and the true (but unknown) density. However, optimality theory is not provided and theoretical analysis is left to future research.[1]

While Clements (2006) and Granger, White, and Kamstra (1989) have considered respectively the combination of event and quantile forecasts, which inevitably involve a loss of information compared with consideration of the 'whole' density, the combination of density forecasts has been relatively neglected. Indeed Clements (2003) identifies this as "an area waiting investigation" (p.2).

However, the finite mixture distribution, which takes a weighted linear combination of the competing density forecasts, offers a well understood and much exploited means of combining density forecasts; see Wallis (2005). For example, the Survey of Professional Forecasters [SPF], previously the ASA-NBER survey, has essentially used it since 1968 to publish a combined density forecast of inflation, amongst other things. Since respondents to the SPF supply density forecasts in the form of histograms, the average or combined density forecast is defined as the mean density forecast across respondents.[2] Despite this long history, to-date little attention has been paid to how the weights on the competing density forecasts in the finite mixture should be determined. But as experience of combining point forecasts has taught us, irrespective of its performance in practice, use of equal weights is only one of many options. For example, one popular alternative to equal weights in the point forecast literature, the so-called regression approach, is to tune the weights to reflect the historical performance of the competing forecasts; e.g. see Granger and Ramanathan (1984) [GR].

The way in which we measure the accuracy of forecasts is central to how we choose to combine them. Point forecasts are traditionally evaluated on the basis of their RMSE relative to the subsequent realizations of the variable. Then point forecasts can be optimally combined to achieve the most "accurate" combined forecast, in the sense of minimum RMSE; this amounts to choosing the optimal weights via OLS estimation of the realizations of the variable on the competing point forecasts. Our methodology for combining density forecasts extends this logic and is motivated by the desire to obtain the most "accurate" density forecast, in a statistical sense.[3] This is defined as that set of weights in the finite mixture that minimize the KLIC distance between the combined density forecast and the true but unknown density of the variable to be forecast. Practically, and conveniently, this minimization can be achieved using the logarithmic scoring rule. Scoring rules evaluate the quality of density forecasts by assigning a numerical score based on the forecast and the subsequent realization of the variable. The use of scoring rules is attractive as it circumvents the need to specify/estimate either the unknown true density of the variable to be forecast or the density of the probability integral transforms of the realization of the variable with respect to the forecast densities.

The plan of this paper is as follows. Section 2 discusses some characteristics of combined density forecasts, and Section 3 proposes a simple approach to choose the combining weights optimally. Comparisons with the optimal combination of point forecasts are made. Section 4 then provides an application to UK inflation. One-year ahead density forecasts of UK inflation are now published each quarter both by the Bank of England in its "fan" chart and the National Institute of Economic and Social Research (NIESR) in its quarterly forecast, and have been for the last ten years. The fan chart is central to the setting of monetary policy by the Monetary Policy Committee at the Bank of England. We examine whether improved density forecasts for inflation might have been obtained in practice if one had optimally combined these competing forecasts with a simple time-series density forecast. Section 5 concludes.

[1] Bayesian Model Averaging offers a natural framework in which to consider density forecast combination; see Section 2 for a further brief discussion and references.

[2] The SPF survey has been analysed *inter alia* by Zarnowitz and Lambros (1987), Diebold, Tay, and Wallis (1999), Giordani and Söderlind (2003) and Clements (2006).

[3] It can be contrasted with economic approaches to evaluation, which evaluate forecasts in terms of their implied economic value; see Granger and Pesaran (2000) and Clements (2004).