Revisiting the forecast combination puzzle with different data types: An empirical study

A research proposal submitted for the degree of

Bachelor of Commerce (Honours)

by

Xiefei Li

30204232

Department of Econometrics and Business Statistics

Monash University

Australia

April 2023

Contents

Introduction

1.1 Research Question and Objective

This paper aims to demonstrate the presence of the forecast combination puzzle in various settings besides the time series domain and to examine the general solution to the forecast combination puzzle empirically. The combination puzzle refers to the empirical finding that the simple average combination method often out-performs sophisticated combination methods. Over the past 50 years, the empirical study undertaken so far has been limited, in that most attention have been focused on different time series datasets. Therefore, it is necessary to explore whether the forecast combination puzzle is present in other data types. Furthermore, the general solution for the forecast combination puzzle is still lack of empirical support. The empirical study will be extended to examine the application of the general explanation and solution proposed by Zischke et al. (2022) and Frazier et al. (2023).

1.2 Motivation

The forecasting accuracy is always a concern when forecasts are used in the decision-making. Under the classical frequentist approach, forecasters often choose only one "best model" to mimic the actual data generating process and then predict the future values. However, that single model could be misleading as it may not capture all the important features of the data. Bates and Granger (1969) discussed the idea of forecast combination and empirically showed the accuracy improvement with a number of combination techniques. This influential finding has attracted increasing attention and massive contributions to combination methods; see Wang et al. (2022) for a modern literature review over the past 50 years.

In short, forecast combinations involve producing point or density forecasts, and then combining them based on a rule or a weighting scheme. Multiple characteristics of the true data generating process could be incorporated via this flexible combination process, which will mitigate different sources of uncertainties and therefore, imporve the accuracy. However, issues arise with careless implementation of methods.

One surprising phenomenon in many empirical study, coined by Stock and Watson (2004), is the so-called "forecast combination puzzle" - "theoretically sophisticated weighting schemes should provide more benefits than the simple average from forecast combination, while empirically the simple average has been continuously found to dominate more complicated approaches to combining forecasts". This counter-intuitive result is mainly discovered in the time series settings, then what will happen when working with datasets such as surveys of professional forecasters, dynamic panels, and pure cross-sectional? Following section 2 of Geweke and Amisano (2011), the weighted linear combinations of two different prediction models will be evaluated via the log predictive scoring rule.

If the forecast combination puzzle occurs in all settings, then it is essential to solve it in a general way without any specific restrictions. Frazier et al. (2023) demonstrated that, in theory, the cause of the puzzle is the two-step approach of estimating unknown parameters and combination weights separately. If forecasts are produced by estimating parameters and weights at the same time, the sophisticated weighting schemes should (asymptotically) be superior.

Methodology

The first goal is to combine any two parametric and suitable models for a given dataset in a linear way. Then, I will show that the forecast combinations indeed improve the forecast accuracy by assessing the log predictive score function.

Before explaining the details, I introduce the following notations used throughout the paper. For a probability density measure p, a vector time series $\boldsymbol{y}_t$ and its past information $\boldsymbol{Y}_{t-1}$, I use $p(\boldsymbol{y}_t|\boldsymbol{Y}_{t-1})$ to represent the probability density for $\boldsymbol{y}_t$ with known history until $t-1$.

The unknown parameters are denoted as a vector $\boldsymbol{\theta}$ and the two-model prediction pool is denoted as $\boldsymbol{M} = \{M_1, M_2\}$.

Assuming that the sample has T available observations, it will be partitioned into two periods with rough proportion, R in-sample periods and P out-of-sample periods, where $R + P = T$.

2.1 Forecast Combination Method

Following section 1.2 of Geweke and Amisano (2011), the forecasts are based on a two-model linear pool M of two constituent predictive densities M_1 and M_2:

$$p(\boldsymbol{y}_t|\boldsymbol{Y}_{t-1}, M) = wp(\boldsymbol{y}_t^{obs}|\boldsymbol{Y}_{t-1}, M_1) + (1-w)p(\boldsymbol{y}_t^{obs}|\boldsymbol{Y}_{t-1}, M_2) \tag{2.1}$$

$p(\boldsymbol{y}_t|\boldsymbol{Y}_{t-1}, M_j)$, $j = 1, 2, ..., J$ where J is the total number of considered models

M_1 and M_2 where $M_1 \neq M_2$. The two-model pool is then defined as $M = \{M_1, M_2\}$.

the weighted linear combinations of two different prediction models will be evaluated via the log predictive scoring rule.

How to combine forecasts

Linear pools

Assumed models are

Combined weights

2.2 Evaluation of Weighted Forecast Combinations

Following section 1.1 of Geweke and Amisano (2011),

2.3 A Motivating Example

2.4 Citations

All citations should be done using markdown notation as shown below. This way, your bibliography will be compiled automatically and correctly. (Frazier et al., 2023; Bates and Granger, 1969). (see Frazier et al., 2023).

The methods were less popular in academic circles until Frazier et al. (2023) introduced a state space formulation of some of the methods, which was extended in **HKSG02** to cover the full range of exponential smoothing methods.

Preliminary Results

3.1 Tables

Let's assume future advertising spend and GDP are at the current levels. Then forecasts for the next year are given in Table **??**.

{r salesforecasts, results="asis"} as.data.frame(fc) %>% select(Quarter, .mean) %>% rename(`Sales forecast` = .mean) %>% knitr::kable(booktabs=TRUE, digits=1, caption = "Forecasts for the next year assuming Advertising budget and GDP are unchanged.")

Again, notice the use of labels and references to automatically generate table numbers.

3.2 Figures

Figure **??** shows time plots of the data we just loaded.

{r deaths, message=FALSE, fig.cap= "Quarterly sales, advertising and GDP data."} sales %>% pivot_longer(Sales:GDP) %>% autoplot(value) + facet_grid(name ~ ., scales = "free_y") + theme(legend.position = "none")

Appendix

$$y_t - y_{t-4} = \beta(x_t - x_{t-4}) + \gamma(z_t - z_{t-4}) + \phi_1(y_{t-1} - y_{t-5}) + \Theta_1 \varepsilon_{t-4} + \varepsilon_t \qquad \text{(A.1)}$$

Bibliography

Bates, JM and CW Granger (1969). The combination of forecasts. *Journal of the operational research society* **20**(4), 451–468. DOI: https://doi.org/10.1057/jors.1969.103.

Frazier, DT, R Zischke, GM Martin, and DS Poskitt (2023). Solving the Forecast Combination Puzzle. [In preparation].

Geweke, J and G Amisano (2011). Optimal prediction pools. *Journal of Econometrics* **164**(1), 130–141. DOI: https://doi.org/10.1016/j.jeconom.2011.02.017.

Stock, JH and MW Watson (2004). Combination forecasts of output growth in a seven-country data set. *Journal of forecasting* **23**(6), 405–430. DOI: https://doi.org/10.1002/for.928.

Wang, X, RJ Hyndman, F Li, and Y Kang (2022). Forecast combinations: an over 50-year review. *International Journal of Forecasting*. DOI: https://doi.org/10.48550/arXiv.2205.04216.

Zischke, R, GM Martin, DT Frazier, and DS Poskitt (2022). The Impact of Sampling Variability on Estimated Combinations of Distributional Forecasts. *arXiv preprint arXiv:2206.02376*. DOI: https://doi.org/10.48550/arXiv.2206.02376.