Discussion

Forecast reconciliation with clustering structure: application to stock prices

Tommaso Di Fonzo

Department of Statistical Sciences, University of Padova

@ difonzo@stat.unipd.it

Prato – September 8^{th} 2023

IIF Workshop on Forecast Reconciliation

Overview

- One of the first applications of **regression-based** forecast reconciliation to **financial time series**: stock prices of the DJIA (for Realized Variance daily forecasts through forecast reconciliation, see Caporin et al., 2023)
- Application of time series clustering to discover and exploit new **data-based hierarchies** to be used in the reconciliation phase: **extension** of an idea formerly used in **bottom-up** and **top-down** frameworks (Darrough and Russell, 2002)
- The authors correctly stress that this can be seen as an **original use of forecast combination** (Di Fonzo and Girolimetto, 2022) for financial time series forecasting
- **Very promising conclusion**:

If the hierarchical structure of the stock market is properly included in the forecasting problem, we can obtain <u>more accurate forecasts than random walk</u>

Main issues considered in the discussion

Not exhaustive, only topics closer to the issues I usually deal with

- The paper considers a special case, the DJIA and its components, that is a **price-weighted index**, where generally equity indices are **value weighted**
- The DJIA divisor d_t in $y_t = \frac{\sum_{i=1}^{N} p_{it}}{d_t}$ (maybe a **minor issue** in the time period covered by the forecasting experiment)
- The industry hierarchy is **unbalanced** (in 15 cases out of 30, only one stock per industry). Possible **impacts** on point reconciled forecasts and their variances
- **Editorial warning**: the same symbol d_t is used to denote both the DJIA divisor (expression (1), page 4) and the **error differential** $d_t = g(e_{1,t}) - g(e_{2,t})$ between two forecasting approaches, used to build the DM test (expression (14), page 9)

DJIA merits and failings (Lin et al., 2021)

- **The literature on the merits and failings of the DJIA is almost as old as the index itself**
  Comer (1952), Milne (1966), and Rudd (1979) discuss the role of stock splits, the weighting methods, and the industry composition of the DJIA. The importance of dividend payments for the DJIA has been discussed by Clarke and Statman (2000) and Shoven and Sialm (2000)
- **Each company in the index is weighted by the price of its stock**
  The importance of each company in the index does not depend on the total market capitalization (a measure of the size) of the company Instead, a highly priced stock has a higher weight than a lower priced stock. Each time a company in the DJIA splits, the weight of this company decreases because the stock price falls by the ratio of the split
- **The constituents of the DJIA are not representative of the market as a whole**
  The 30 firms are chosen among large-capitalization firms to represent different industries, but they are not chosen according to fixed pre-determined rules. In particular, the DJIA is not an index of the 30 largest companies in the US. Furthermore, a more representative index would include a much larger number of companies

The DJIA divisor

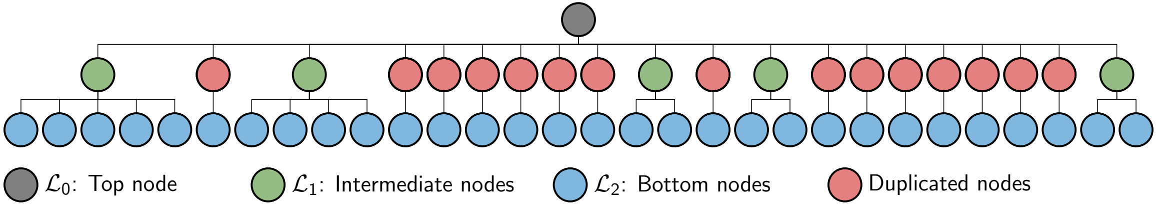
Just a matter of my curiosity, maybe a minor issue for this paper

- Since its inception on May 26, 1896, the components of the DJIA **have changed 57 times**.
- The Divisor approach worked well for the first few decades but, in recent years, the Divisor has started to become **very small**. In 1986, the Dow Divisor fell below 1.0 for the first time – effectively becoming the "Dow Multiplier"
- As of November 4, 2021, the current divisor for DJIA is 0.1517. What about the time period **(1-09-2020 to 30-09-2022)** considered in the forecasting experiment?
- It would be interesting (useful for forecasting purposes?) to know how the **Dow Divisor changed over the time**

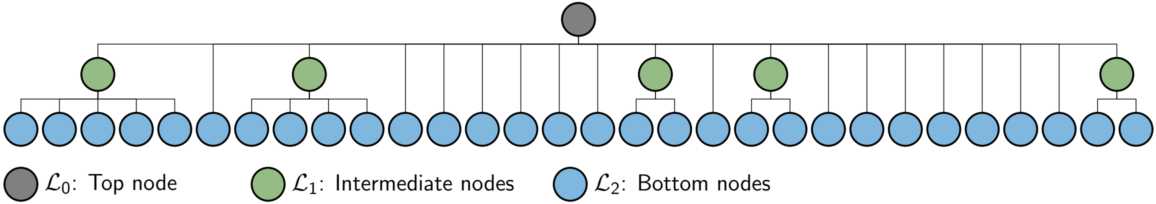
Balanced and unbalanced hierarchy

Hierarchy of the DJIA components (30) according to Industrial-based clusters (20): IND

Balanced hierarchy



Unbalanced hierarchy



Aggregation matrix A_B $(20 \times 30) \Rightarrow S_B = [A_B' \; I_{30}]'$ (50×30)

Balanced hierarchy

	AAPL	CRM	CSCO	IBM	MSFT	**AMGN**	AXP	GS	JPM	V	**BA**	**CAT**	**CVX**	**DIS**	**DOW**	**HD**	HON	MMM	**INTC**	JNJ	MRK	**KO**	**MCD**	**NKE**	**PG**	**TRV**	**UNH**	**VZ**	WBA	WMT
Information technology	1	1	1	1	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Biopharmaceutical	.	.	.	.	.	**1**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Financial services	.	.	.	.	.	.	1	1	1	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Aerospace and defense	.	.	.	.	.	.	.	.	.	.	**1**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Construction and Mining	.	.	.	.	.	.	.	.	.	.	.	**1**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Petroleum industry	.	.	.	.	.	.	.	.	.	.	.	.	**1**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Broadcasting and entertainment	.	.	.	.	.	.	.	.	.	.	.	.	.	**1**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Chemical industry	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**1**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Home Improvement	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**1**	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Conglomerate	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	1	.	.	.	.	.	.	.	.	.	.	.	.
Semiconductor industry	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**1**	.	.	.	.	.	.	.	.	.	.	.
Pharmaceutical industry	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	1	.	.	.	.	.	.	.	.	.
Drink industry	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**1**	.	.	.	.	.	.	.	.
Food industry	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**1**	.	.	.	.	.	.	.
Clothing industry	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**1**	.	.	.	.	.	.
Fast-moving consumer goods	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**1**	.	.	.	.	.
Insurance	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**1**	.	.	.	.
Managed health care	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**1**	.	.	.
Telecommunications industry	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	**1**	.	.
Retailing	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	1

Aggregation matrix A (5×30) $\Rightarrow$ $S = [A' \; I_{30}]'$ (35×30)

Unbalanced hierarchy

	AAPL	CRM	CSCO	IBM	MSFT	**AMGN**	AXP	GS	JPM	V	**BA**	**CAT**	**CVX**	**DIS**	**DOW**	**HD**	HON	MMM	**INTC**	JNJ	MRK	**KO**	**MCD**	**NKE**	**PG**	**TRV**	**UNH**	**VZ**	WBA	WMT
Information technology	1	1	1	1	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Financial services	.	.	.	.	.	.	1	1	1	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Conglomerate	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	1	.	.	.	.	.	.	.	.	.	.	.	.
Pharmaceutical industry	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	1	.	.	.	.	.	.	.	.	.
Retailing	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	1

Issues with unbalanced hierarchy

If not appropriately dealt with

- In general, bts point forecasts are **different**: $\widetilde{\boldsymbol{b}}_B \neq \widetilde{\boldsymbol{b}}$, with $\widetilde{\boldsymbol{b}}_B = \widetilde{\boldsymbol{b}} + \boldsymbol{d}$
- However, provided $\widehat{\boldsymbol{y}}$ is unbiased, $\widetilde{\boldsymbol{b}}_B$ (and $\widetilde{\boldsymbol{y}}_B$ as well) is still **unbiased** (since $E\left(\boldsymbol{d}\right) = \boldsymbol{0}$)
- $\widetilde{\boldsymbol{b}}_B = \widetilde{\boldsymbol{b}}$ only when **all variables are replicated** (i.e., $\widehat{\boldsymbol{y}}_B = [\widehat{\boldsymbol{y}}' \ \widehat{\boldsymbol{y}}']'$, $\widehat{\boldsymbol{y}}_B = [\widehat{\boldsymbol{y}}' \ \widehat{\boldsymbol{y}}' \ \widehat{\boldsymbol{y}}']' \ldots$)
- $\widetilde{\boldsymbol{y}}_B$ contains \`**duplicated values**', then $\widetilde{\boldsymbol{y}}_B$ and $\widetilde{\boldsymbol{y}}$ have **different dimensions**, but in both cases the complete vector of point reconciled forecasts is **coherent**:

$$\widetilde{\boldsymbol{y}}_B = \boldsymbol{S}_B \widetilde{\boldsymbol{b}}_B \quad \text{and} \quad \widetilde{\boldsymbol{y}} = \boldsymbol{S}\widetilde{\boldsymbol{b}}$$

- The variances of the reconciled bts forecasts are **different**, and those coming from the balanced framework are **downward biased** (e.g., halved when all bottom and upper variables are duplicated):

$$Var\left(\widetilde{\boldsymbol{b}}_B - \boldsymbol{b}\right) \leq Var\left(\widetilde{\boldsymbol{b}} - \boldsymbol{b}\right)$$

- Thus, if the balanced framework is used in cases where it is not justified by the data structure, the forecast evaluation/comparison might be **not well grounded** (e.g., confidence intervals will shrink)

Final issues about the forecasting experiment and the evaluation of investment strategies

- Why $h = 1, 3, 6, 12$ are considered instead of the usual (in financial forecasting experiments) $h = 1, 5, 22$?
- Please, check Table 6: when forecasting the top series (DJIA index), for $h = 1$ **RW is better than MinT:ARMA** in terms of RMSE loss (424.80 vs 424.86)
- Same in Table 8: when considering all the constituents series of the DJIA, for $h = 1$ **RW is better than MinT:ARMA** in terms of MAE loss (15.35 *vs* 15.37)
- In Table 9, **no comparison between Base and RW** forecast accuracy (both in terms of absolute and squared errors). Why?
- The procedure proposed to test if two investment strategies lead to statistically different Sharpe ratios **should be updated** according to the robust approach by Ledoit & Wolfe (2008) (however, I don't think that the final results will change that much ...)
- When comparing the Sharpe ratios between two forecast-based investment strategies (Table 10), **RW is not considered**. Why?

Conclusions

- A **very promising** work, where two modern statistical approaches (time series clustering and forecast reconciliation) are considered in an **original way**
- A deeper consideration of the characteristics of the DJIA, and a fine-tuning of the forecasting experiment might **improve the relevance** of the contribution (e.g., usefulness for financial econometricians and investors)
- However, looking at Table 9 of the paper, it seems **hard and questionable** to draw a general conclusion like

  If the hierarchical structure of the stock market is properly included in the forecasting problem, we can obtain more accurate forecasts than random walk

  Rather, it seems to be confirmed that the Efficient Market Hypothesis is **hard to beat**!
- A little **take-home message** to practitioners: when facing unbalanced hierarchies, please **do not balance** them

References I

[1] Caporin, M., Di Fonzo, T., Girolimetto, D. (2023), Exploiting intraday decompositions in realized volatility forecasting: A forecast reconciliation approach, arXiv.2306.02952.

[2] Clarke, R. G., Statman, M. (2000), The DJIA crossed 652,230, *The Journal of Portfolio Management* 26(2), 89–93.

[3] Comer, H.D. (1952), The market-looking backward, *The Financial Analysts Journal* 8(4), 143–144.

[4] Di Fonzo, T., Girolimetto, D. (2022), Forecast combination based forecast reconciliation, *International Journal of Forecasting*, in press.

[5] Ledoit, O., Wolf, M. (2008), Robust performance hypothesis testing with the Sharpe ratio, *Journal of Empirical Finance* 15, 850–859.

[6] Lin, J., Selden, G.C., Shoven, J.B., Sialm, C. (2021), *Replicating the Dow Jones Industrial Average*, National Bureau of Economic Research Working Paper.

[7] Milne, R. D. (1966), The Dow-Jones Industrial Average re-examined, *The Financial Analysts Journal*, 22(6), 83–88.

References II

[8] Rudd, A.T. (1979), The revised Dow Jones Industrial Average: New wine in old bottles?, *The Financial Analysts Journal* 35(6), 57–63.

[9] Shoven, J.B., Sialm, C. (2000), The Dow Jones Industrial Average: The impact of fixing its flaws, *The Journal of Wealth Management* 3(3), 9–18.

Prato – September 8^{th} 2023

IIF Workshop on Forecast Reconciliation

UNIVERSITÀ DEGLI STUDI DI PADOVA

THANK YOU!

Tommaso Di Fonzo

Department of Statistical Sciences

University of Padova (Italy)

@ difonzo@stat.unipd.it