Forecast reconciliation: Methodological issues and applications

Chapter 1 - Cross-temporal forecast reconciliation:
Optimal combination method and heuristic alternatives[1]

Online appendix

Candidate: Daniele Girolimetto
Supervisors: Prof. Tommaso Di Fonzo, Prof. George Athanasopoulos and Prof. Rob J Hyndman

Contents

[1]Di Fonzo, T. and Girolimetto, D. (2023a) Cross-temporal forecast reconciliation: Optimal combination method and heuristic alternatives. International Journal of Forecasting 39(1), 39–57. doi:10.1016/j.ijforecast.2021.08.004

A Balanced and unbalanced hierarchies

A simple three-level hierarchy is shown in the right panel of figure A.1, where variable C at the second level of the hierarchy has no 'children', and thus is considered as a bottom variable too, at level three of the hierarchy.

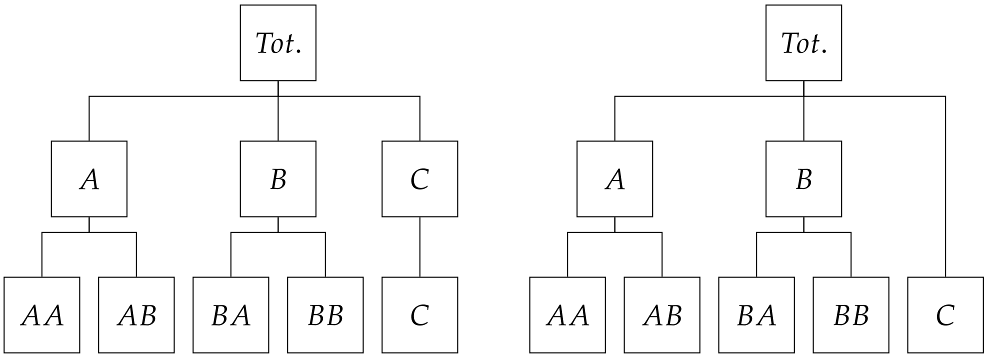


Figure A.1: *A simple unbalanced hierarchy (right) and its balanced version (left)*

The left panel shows the 'balanced version' of the same hierarchy, where variable C is (duplicated and) present at both levels two and three.

The aggregation relationships linking the component series can be expressed as follows:

$$\begin{array}{rcl} y_{Tot} & = & y_{AA} + y_{AB} + y_{BA} + y_{BB} + y_C \\ y_A & = & y_{AA} + y_{AB} \\ y_B & = & y_{BA} + y_{BB} \\ y_C & = & y_C \end{array},$$

where the last equality has merely the function of making the hierarchy balanced. The corresponding contemporaneous aggregation matrix $\boldsymbol{C}$ is given by:

$$\boldsymbol{C} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The redundant relationship ($y_C = y_C$) makes the last row of matrix $\boldsymbol{C}$ equal to the last row of the contemporaneous summing matrix $\boldsymbol{S} = \begin{bmatrix} \boldsymbol{C} \\ \boldsymbol{I}_5 \end{bmatrix}$. This redundancy can be easily eliminated by considering the new contemporaneous aggregation matrix $\tilde{\boldsymbol{C}}$, which in the case of an unbalanced hierarchy has clearly one row less than in the balanced version:

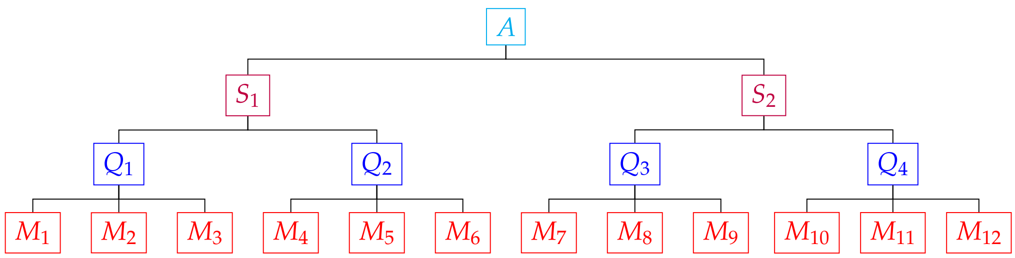
$$\tilde{\boldsymbol{C}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}.$$

The new contemporaneous summing matrix is thus given by $\tilde{\boldsymbol{S}} = \begin{bmatrix} \tilde{\boldsymbol{C}} \\ \boldsymbol{I}_5 \end{bmatrix}$, which has dimension (8×5) instead of (9×5) as for matrix $\boldsymbol{S}$. In a complex hierarchy, mostly when contemporaneous and temporal hierarchies are simultaneously considered, this fact should be carefully considered in order to save memory space and computing time.

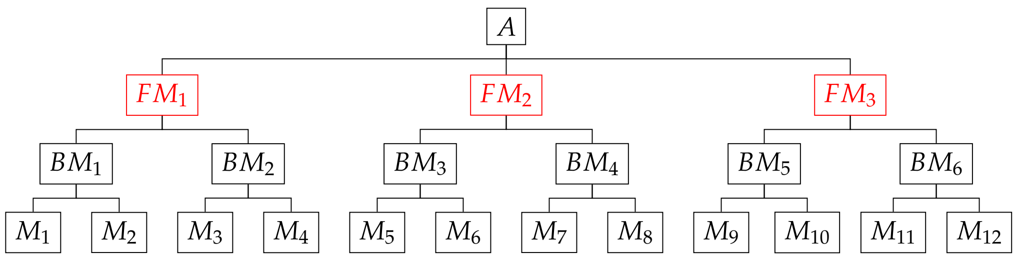
The R package `hts` (Hyndman et al. 2020) manages only balanced hierarchies, and thus makes use of matrix S instead of $\tilde{S}$. Due to this fact, large cross-sectional hierarchies might require computational efforts larger than necessary, and could face numerical problems when more sofisticated reconciliation strategies are applied. For example, the grouped time series of the Australian Tourism Demand (Visitor Nights) analyzed by Wickramasuriya et al. (2019) (see also Ashouri et al. 2021; Bertani et al. 2020; Wickramasuriya et al. 2020), contains 30 duplicated time series, since it comes from two unbalanced hierarchies with only 525 'unique' time series (304 bts and 221 uts), as compared to the 555 time series of the balanced version. Similar, though less pronounced cases found in literature are (i) the reduced version (an unbalanced geographical hierarchy of 105 'unique' time series out of 111) of the Australian Visitor Nights dataset analyzed by Kourentzes & Athanasopoulos (2019), and (ii) the Australian Tourism Demand (Overnight Trips) dataset considered by Panagiotelis et al. (2021), which consists of 104 'unique' time series out of 110 for the balanced hierarchy.

B Monthly and hourly temporal hierarchies

For monthly data, the aggregates of interest are for $k \in \{12, 6, 4, 3, 2, 1\}$. Hence the monthly observations are aggregated to annual, semi-annual, four-monthly, quarterly and bi-monthly observations. These can be represented in two separate hierarchies, as shown in Fig. B.2, which means that the temporal hierarchies form a grouped series, sharing the 'top level' (annual) aggregate, and the same twelve 'bottom' nodes, one for each month of the original temporally disaggregated time series.



(a) *Monthly - Quarterly - Semi-Annual - Annual frequencies*



(b) *Monthly - Bi-Monthly - Four-Monthly - Annual frequencies*

Figure B.2: *The two temporal hierarchies induced by a monthly time series.*

However, the (16×12) temporal aggregation matrix $\boldsymbol{K}_1$ for this case is easily obtained:

$$\boldsymbol{K}_1 = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1
\end{bmatrix},$$

and thus $\boldsymbol{R}_1 = \begin{bmatrix} \boldsymbol{K}_1 \\ \boldsymbol{I}_{12} \end{bmatrix}$, $\boldsymbol{Z}_1' = [\boldsymbol{I}_{16} \ \ -\boldsymbol{K}_1]$, $\boldsymbol{x}_\tau = \boldsymbol{R}_1 \boldsymbol{x}_\tau^{[1]}$, and $\boldsymbol{Z}_1' \boldsymbol{x}_\tau = \boldsymbol{0}$, $\tau = 1, \ldots, N$, where $\boldsymbol{x}_\tau = \left[x_\tau^{[12]}, \boldsymbol{x}_\tau^{[6]'}, \boldsymbol{x}_\tau^{[4]'}, \boldsymbol{x}_\tau^{[3]'}, \boldsymbol{x}_\tau^{[2]'}, \boldsymbol{x}_\tau^{[1]'} \right]'$ is the (28×1) vector containing all temporal aggregates of variable X at the observation index τ (i.e., within the complete τ-th cycle).

Let's conclude with considering the case of an hourly time series with diurnal periodicity. In this case it is $m = 24$, $k^* = 36$, and $\boldsymbol{K}_N$ is the $(36N \times 24N)$ matrix

$$\boldsymbol{K}_N = \begin{bmatrix}
\boldsymbol{I}_N \otimes \boldsymbol{1}_{24}' \\
\boldsymbol{I}_{2N} \otimes \boldsymbol{1}_{12}' \\
\boldsymbol{I}_{3N} \otimes \boldsymbol{1}_{8}' \\
\boldsymbol{I}_{4N} \otimes \boldsymbol{1}_{6}' \\
\boldsymbol{I}_{6N} \otimes \boldsymbol{1}_{4}' \\
\boldsymbol{I}_{8N} \otimes \boldsymbol{1}_{3}' \\
\boldsymbol{I}_{12N} \otimes \boldsymbol{1}_{2}'
\end{bmatrix},$$

which converts single hour values into the sum of 2, 3, 4, 6, 8, 12, and 24 hours data, respectively, and $\boldsymbol{Z}_N' = [\boldsymbol{I}_{36N} \ \ -\boldsymbol{K}_N]$ is a full row-rank $(36N \times 60N)$ matrix.

C A Cross-temporal structural representation

The cross-temporal structural representation can be seen as a generalization from a single time index t to a single cycle index τ (i.e., the low-frequency time index) of the cross-sectional structural representation, extended to cover $n(k^* + m)$ nodes instead of n.

Denote with $\boldsymbol{Y}_\tau$ the $[(n \times (k^* + m)]$ data matrix available at cycle τ:

$$\boldsymbol{Y}_\tau = \begin{bmatrix} \boldsymbol{A}_\tau \\ \boldsymbol{B}_\tau \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_\tau^{[m]} \; \boldsymbol{A}_\tau^{[k_{p-1}]} \; \dots \; \boldsymbol{A}_\tau^{[k_2]} \; \boldsymbol{A}_\tau^{[1]} \\ \boldsymbol{B}_\tau^{[m]} \; \boldsymbol{B}_\tau^{[k_{p-1}]} \; \dots \; \boldsymbol{B}_\tau^{[k_2]} \; \boldsymbol{B}_\tau^{[1]} \end{bmatrix}, \quad \tau = 1, \dots N,$$

and let $\check{\boldsymbol{S}}$ be the $(n(k^* + m) \times n_b m)$ cross-temporal summation matrix

$$\check{\boldsymbol{S}} = \begin{bmatrix} \check{\boldsymbol{C}} \\ \boldsymbol{I}_{n_b m} \end{bmatrix},$$

where $\check{\boldsymbol{C}}$ denotes a $(n_a^* \times n_b m)$ cross-temporal aggregation matrix mapping the hf-bts into the uts and lf-bts ones ($n_a^* = n_a(k^* + m) + n_b k^*$). Denote with

$$\boldsymbol{a}_\tau^* = \begin{bmatrix} \mathrm{vec}\left(\boldsymbol{A}_\tau'\right) \\ \mathrm{vec}\left(\boldsymbol{B}_\tau^{*\prime}\right) \end{bmatrix}, \quad \tau = 1, \dots, N,$$

the $(n_a^* \times 1)$ vector of 'cross-temporal upper series', containing the uts and lf-bts data at the low-frequency time index τ, where $\boldsymbol{B}_\tau^* = \left[\boldsymbol{B}_\tau^{[m]} \; \boldsymbol{B}_\tau^{[k_{p-1}]} \; \dots \; \boldsymbol{B}_\tau^{[k_2]}\right]$, $\tau = 1, \dots, N$, and with

$$\boldsymbol{b}_\tau^{[1]} = \mathrm{vec}\left(\boldsymbol{B}_\tau^{[1]\prime}\right), \quad \tau = 1, \dots, N,$$

the $(n_b m \times 1)$ vector of 'cross-temporal bottom series', containing the hf-bts data. The structural representation of a cross-temporal system of n time series takes the form

$$\check{\boldsymbol{y}}_\tau = \check{\boldsymbol{S}} \boldsymbol{b}_\tau^{[1]}, \quad \tau = 1, \dots, N, \tag{C.1}$$

where $\check{\boldsymbol{y}}_\tau$ is a $[n(k^* + m) \times 1]$ vector where we place all the uts and the lf-bts at the top, and all the hf-bts at the bottom:

$$\check{\boldsymbol{y}}_\tau = \begin{bmatrix} \boldsymbol{a}_\tau^* \\ \boldsymbol{b}_\tau^{[1]} \end{bmatrix}, \quad \tau = 1, \dots, N. \tag{C.2}$$

D Commutation matrix and the relationships linking vectors and matrices of bottom and upper time series

Given an $(r \times c)$ matrix $\boldsymbol{X}$, denote with $\boldsymbol{C}_{r,c}$ the $(rc \times rc)$ commutation matrix (Magnus & Neudecker 2019) which maps $\mathrm{vec}(\boldsymbol{X})$ into $\mathrm{vec}(\boldsymbol{X}')$:

$$\boldsymbol{C}_{r,c} \mathrm{vec}(\boldsymbol{X}) = \mathrm{vec}\left(\boldsymbol{X}'\right).$$

This matrix is a special type of permutation matrix, obtained by simple exchanges of rows of the identity matrix, and is therefore orthogonal, that is:

$$\boldsymbol{C}_{r,c}^{-1} = \boldsymbol{C}_{r,c}' = \boldsymbol{C}_{c,r}.$$

D.1 Cross-sectional case

Denoting $\boldsymbol{b}^* = \text{vec}(\boldsymbol{B})$, $\boldsymbol{b} = \text{vec}(\boldsymbol{B}')$, $\boldsymbol{a}^* = \text{vec}(\boldsymbol{A})$, $\boldsymbol{a} = \text{vec}(\boldsymbol{A}')$, the mappings of $\boldsymbol{b}^*$ into $\boldsymbol{b}$ and $\boldsymbol{a}^*$ into $\boldsymbol{a}$, respectively, can be expressed as

$$\boldsymbol{P}_b \boldsymbol{b}^* = \boldsymbol{b}, \quad \boldsymbol{P}_a \boldsymbol{a}^* = \boldsymbol{a},$$

where $\boldsymbol{P}_b = \boldsymbol{C}_{n_b T, n_b T}$ and $\boldsymbol{P}_a = \boldsymbol{C}_{n_a T, n_a T}$ are $(n_b T \times n_b T)$ and $(n_a T \times n_a T)$, respectively, commutation matrices. Since both $\boldsymbol{P}_b$ and $\boldsymbol{P}_a$ are orthogonal, it is:

$$\boldsymbol{b}^* = \boldsymbol{P}_b' \boldsymbol{b}, \quad \boldsymbol{a}^* = \boldsymbol{P}_a' \boldsymbol{a}.$$

The index k, $k = 1, \dots, n_b T$, of the generic element of vector $\boldsymbol{b}$ can be expressed in terms of the row and column indices of the corresponding element of matrix $\boldsymbol{B}'$:

$$\text{vec}(\boldsymbol{B}') = \boldsymbol{b} = \{b_k\}, \quad b_k = b_{ti}, \text{ with } k = t + (i-1)T.$$

As for the index l, $l = 1, \dots, n_a T$, of the generic element of vector $\boldsymbol{a}^*$, we have:

$$\text{vec}(\boldsymbol{A}') = \boldsymbol{a} = \{a_l\}, \quad a_l = a_{tj}, \text{ with } l = t + (j-1)T.$$

A numerical example

Assuming that $n = 2$ variables and $T = 3$ time periods are considered, matrix $\boldsymbol{X} = \begin{bmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \end{bmatrix}$ can be vectorized either as

$$\text{vec}(\boldsymbol{X}) = \boldsymbol{x}^* = \begin{bmatrix} 11 & 21 & 12 & 22 & 13 & 23 \end{bmatrix}'$$

or

$$\text{vec}(\boldsymbol{X}') = \boldsymbol{x} = \begin{bmatrix} 11 & 12 & 13 & 21 & 22 & 23 \end{bmatrix}'.$$

In this case, the permutation matrix $\boldsymbol{P}$ mapping $\boldsymbol{x}$ into $\boldsymbol{x}^*$, such that $\boldsymbol{x}^* = \boldsymbol{P}\boldsymbol{x}$ (and $\boldsymbol{x} = \boldsymbol{P}'\boldsymbol{x}^*$), is given by:

$$\boldsymbol{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The following **R** script performs the calculation of matrix $\boldsymbol{P}$:

```
n <- 2;
t <- 3;
I <- matrix(1:(n*t), n, t, byrow = T)
I <- as.vector(I) # vectorize the required indices
P <- diag(n*t)    # Initialize an identity matrix
P <- P[I,]        # Re-arrange the rows of the identity matrix
```

```
# A numerical example
X <- matrix(c(11,12,13,21,22,23), byrow=T, nrow=2) # (2 x 3) matrix
Xt <- t(X)
xstar <- as.vector(X)  # xstar = vec(X)
x <- as.vector(Xt)     # x = vec(X')
xstarnew <- P%*%x      # vector x is mapped into vector xstarnew
norm(xstar - xstarnew) # check: the norm of the difference should be zero
xnew <- t(P)%*%xstar   # vector xstar is mapped into vector xnew
norm(x - xnew)         # check: the norm of the difference should be zero
```

D.2 Cross-temporal case

In section A.3 we have shown that the vectorization of matrix $\boldsymbol{Y}$ used in the cross-temporal structural representation, $\check{\boldsymbol{y}}$, is different from $\boldsymbol{y} = \text{vec}(\boldsymbol{Y}')$, which is used in the optimal combination cross-temporal reconciliation formula developed in the paper. In what follows, we determine the permutation matrix $\boldsymbol{Q}$ mapping $\check{\boldsymbol{y}}$ into $\boldsymbol{y}$.

Assuming $h = 1$, denote with $\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{bmatrix}$ the $[n \times (k^* + m)]$ matrix of the target forecasts at any temporal frequency. The $[n_b \times (k^* + m)]$ submatrix $\boldsymbol{B}$, which contains the target forecasts of the bottom time series, can be written as:

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{B}^{[m]} \ \boldsymbol{B}^{[k_{p-1}]} \ \dots \ \boldsymbol{B}^{[k_2]} \ \boldsymbol{B}^{[1]} \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}^* \ \boldsymbol{B}^{[1]} \end{bmatrix},$$

where the $(n_b \times k^*)$ matrix $\boldsymbol{B}^* = \begin{bmatrix} \boldsymbol{B}^{[m]} \ \boldsymbol{B}^{[k_{p-1}]} \ \dots \ \boldsymbol{B}^{[k_2]} \end{bmatrix}$, and matrix $\boldsymbol{B}^{[1]}$ contain the target forecasts for, respectively, the temporally aggregated time series (lf-bts) and the high-frequency ones (hf-bts). The following relationships hold:

$$\boldsymbol{C}_{n_b,(k^*+m)} \left[\text{vec}\left(\boldsymbol{B}\right)\right] = \text{vec}\left(\boldsymbol{B}'\right),$$

$$\boldsymbol{C}_{n_b,k^*} \left[\text{vec}\left(\boldsymbol{B}^*\right)\right] = \text{vec}\left[(\boldsymbol{B}^*)'\right],$$

$$\boldsymbol{C}_{n_b,m} \left[\text{vec}\left(\boldsymbol{B}^{[1]}\right)\right] = \text{vec}\left[(\boldsymbol{B}^{[1]})'\right].$$

Since $\text{vec}\left(\boldsymbol{B}\right) = \begin{bmatrix} \text{vec}\left(\boldsymbol{B}^*\right) \\ \text{vec}\left(\boldsymbol{B}^{[1]}\right) \end{bmatrix}$, we can write:

$$\boldsymbol{C}_{n_b,(k^*+m)}\text{vec}\left(\boldsymbol{B}\right) = \boldsymbol{C}_{n_b,(k^*+m)} \begin{bmatrix} \boldsymbol{C}_{k^*,n_b} & \boldsymbol{0}_{(n_b k^* \times n_b m)} \\ \boldsymbol{0}_{(n_b m \times n_b k^*)} & \boldsymbol{C}_{m,n_b} \end{bmatrix} \begin{bmatrix} \text{vec}\left[(\boldsymbol{B}^*)'\right] \\ \text{vec}\left[(\boldsymbol{B}^{[1]})'\right] \end{bmatrix},$$

that is:

$$\text{vec}\left(\boldsymbol{B}'\right) = \widetilde{\boldsymbol{Q}} \begin{bmatrix} \text{vec}\left[(\boldsymbol{B}^*)'\right] \\ \text{vec}\left[(\boldsymbol{B}^{[1]})'\right] \end{bmatrix},$$

where

$$\widetilde{\boldsymbol{Q}} = \boldsymbol{C}_{n_b,(k^*+m)} \begin{bmatrix} \boldsymbol{C}_{k^*,n_b} & \boldsymbol{0}_{(n_b k^* \times n_b m)} \\ \boldsymbol{0}_{(n_b m \times n_b k^*)} & \boldsymbol{C}_{m,n_b} \end{bmatrix}. \tag{D.3}$$

As the $[n(k^*+m)\times 1]$ vector $\check{\boldsymbol{y}}$ can be written as:

$$\check{\boldsymbol{y}} = \begin{bmatrix} \operatorname{vec}(\boldsymbol{A}') \\ \operatorname{vec}\left[(\boldsymbol{B}^*)'\right] \\ \operatorname{vec}\left[\left(\boldsymbol{B}^{[1]}\right)'\right] \end{bmatrix},$$

$\operatorname{vec}(\boldsymbol{Y}')$ can be expressed in terms of $\check{\boldsymbol{y}}$ as:

$$\operatorname{vec}\left(\boldsymbol{Y}'\right) = \boldsymbol{Q}\check{\boldsymbol{y}}, \tag{D.4}$$

where

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{I}_{n_a(k^*+m)} & \boldsymbol{0}_{[n_a(k^*+m)\times n_b m]} \\ \boldsymbol{0}_{[n_b(k^*+m)\times n_a(k^*+m)]} & \widetilde{\boldsymbol{Q}} \end{bmatrix}$$

is a permutation matrix, and thus $\check{\boldsymbol{y}} = \boldsymbol{Q}'\boldsymbol{y}$.

E Bottom-up cross-temporal forecast reconciliation

Cross temporal reconciled forecasts for all series at any temporal aggregation level can be easily computed by appropriate summation of the hf-bts base forecasts $\widehat{\boldsymbol{b}}_i^{[1]}$, $i = 1,\dots,n_b$, according to a bottom-up procedure like what is currently done in both the cross-sectional and temporal frameworks.

Denoting by $\ddot{\boldsymbol{Y}}$ the $[n\times(k^*+m)]$ matrix containing the bottom-up cross temporal reconciled forecasts, it is:

$$\ddot{\boldsymbol{Y}} = \begin{bmatrix} \ddot{\boldsymbol{A}} \\ \ddot{\boldsymbol{B}} \end{bmatrix} = \begin{bmatrix} \ddot{\boldsymbol{A}}^{[m]} & \ddot{\boldsymbol{A}}^{[k_{p-1}]} & \dots & \ddot{\boldsymbol{A}}^{[k_2]} & \ddot{\boldsymbol{A}}^{[1]} \\ \ddot{\boldsymbol{B}}^{[m]} & \ddot{\boldsymbol{B}}^{[k_{p-1}]} & \dots & \ddot{\boldsymbol{B}}^{[k_2]} & \ddot{\boldsymbol{B}}^{[1]} \end{bmatrix}.$$

Since the hf-bts reconciled forecasts are by definition equal to the hf-bts base forecasts, i.e. $\ddot{\boldsymbol{B}}^{[1]} = \widehat{\boldsymbol{B}}^{[1]}$, the bottom-up forecast reconciliation procedure consists of the following steps:

1. compute the hf-uts reconciled forecasts using the cross-sectional aggregation relationship:

$$\ddot{\boldsymbol{A}}^{[1]} = \boldsymbol{C}\widehat{\boldsymbol{B}}^{[1]};$$

2. compute the lf-bts reconciled forecasts according to the temporal aggregation relationships:

$$\begin{bmatrix} \left(\ddot{\boldsymbol{B}}^{[m]}\right)' \\ \vdots \\ \left(\ddot{\boldsymbol{B}}^{[k_2]}\right)' \end{bmatrix} = \boldsymbol{K}_1\left(\widehat{\boldsymbol{B}}^{[1]}\right)' \quad \Rightarrow \quad \left[\ddot{\boldsymbol{B}}^{[m]}\ \ddot{\boldsymbol{B}}^{[k_{p-1}]} \dots \ddot{\boldsymbol{B}}^{[k_2]}\right] = \widehat{\boldsymbol{B}}^{[1]}\boldsymbol{K}_1';$$

3. compute the lf-uts reconciled forecasts by cross-sectional aggregation of the lf-bts reconciled forecasts obtained in the previous step:

$$\ddot{\boldsymbol{A}}^{[k]} = \boldsymbol{C}\ddot{\boldsymbol{B}}^{[k]}, \quad k \in \mathcal{K} \quad \Rightarrow \quad \left[\ddot{\boldsymbol{A}}^{[m]}\ \ddot{\boldsymbol{A}}^{[k_{p-1}]} \dots \ddot{\boldsymbol{A}}^{[k_2]}\right] = \boldsymbol{C}\widehat{\boldsymbol{B}}^{[1]}\boldsymbol{K}_1'.$$

In summary, the matrix containing the bottom-up reconciled forecasts, solely depending on the hf-bts base forecasts, is given by:

$$\ddot{\boldsymbol{Y}} = \begin{bmatrix} \boldsymbol{C}\widehat{\boldsymbol{B}}^{[1]}\boldsymbol{K}_1' & \boldsymbol{C}\widehat{\boldsymbol{B}}^{[1]} \\ \widehat{\boldsymbol{B}}^{[1]}\boldsymbol{K}_1' & \widehat{\boldsymbol{B}}^{[1]} \end{bmatrix}. \tag{E.5}$$

An equivalent, succint alternative to expression (E.5) consists in exploiting the cross-temporal structural representation (C.1):

$$\ddot{\check{\boldsymbol{y}}} = \check{\boldsymbol{S}}\widehat{\boldsymbol{b}}^{[1]}, \tag{E.6}$$

where $\widehat{\boldsymbol{b}}^{[1]} = \text{vec}\left[\left(\widehat{\boldsymbol{B}}^{[1]}\right)'\right]$, keeping in mind that the elements in $\check{\boldsymbol{y}}$ and in $\boldsymbol{y}$ are differently organized, and in general it is $\ddot{\check{\boldsymbol{y}}} \neq \ddot{\boldsymbol{y}}$, with $\ddot{\boldsymbol{y}} = \text{vec}\left(\ddot{\boldsymbol{Y}}'\right)$. This last issue can be easily dealt with by considering expression (D.4) in section A.4.2, according to which expression (E.6) can be re-stated as $\ddot{\boldsymbol{y}} = \boldsymbol{Q}\check{\boldsymbol{S}}\widehat{\boldsymbol{b}}^{[1]}$. However, the formulation of matrix $\check{\boldsymbol{S}}$, which requires to manage linear relationships across cross-sectional and temporal dimensions, may be tedious and prone to errors, mostly for large collections of time series. In such cases, it might be preferible using formulation (E.5), where $\widehat{\boldsymbol{B}}^{[1]}$, $\boldsymbol{C}$ and $\boldsymbol{K}_1$ are involved in simple matrix products.

Section A.6 describes all these features with reference to a ‘toy example’ of a very simple two-level hierarchy with two quarterly bottom time series.

F Cross-temporal hierarchy: a toy example

Let us consider the relationships linking all the variables implied by a cross-temporal hierarchy for the very simple case of a total quarterly series observed for one year, X, obtained as the sum of two component variables, W and Z, respectively. The contemporaneous (cross-sectional) constraint, $X = W + Z$, must hold at any observation index of all temporal frequencies (annual, semi-annual and quarterly) considered in the temporal hierarchy, as shown in Figure F.3, which gives a graphical view of the the way in which the two dimensions (cross-sectional and temporal) are combined within a complete time cycle (one year).

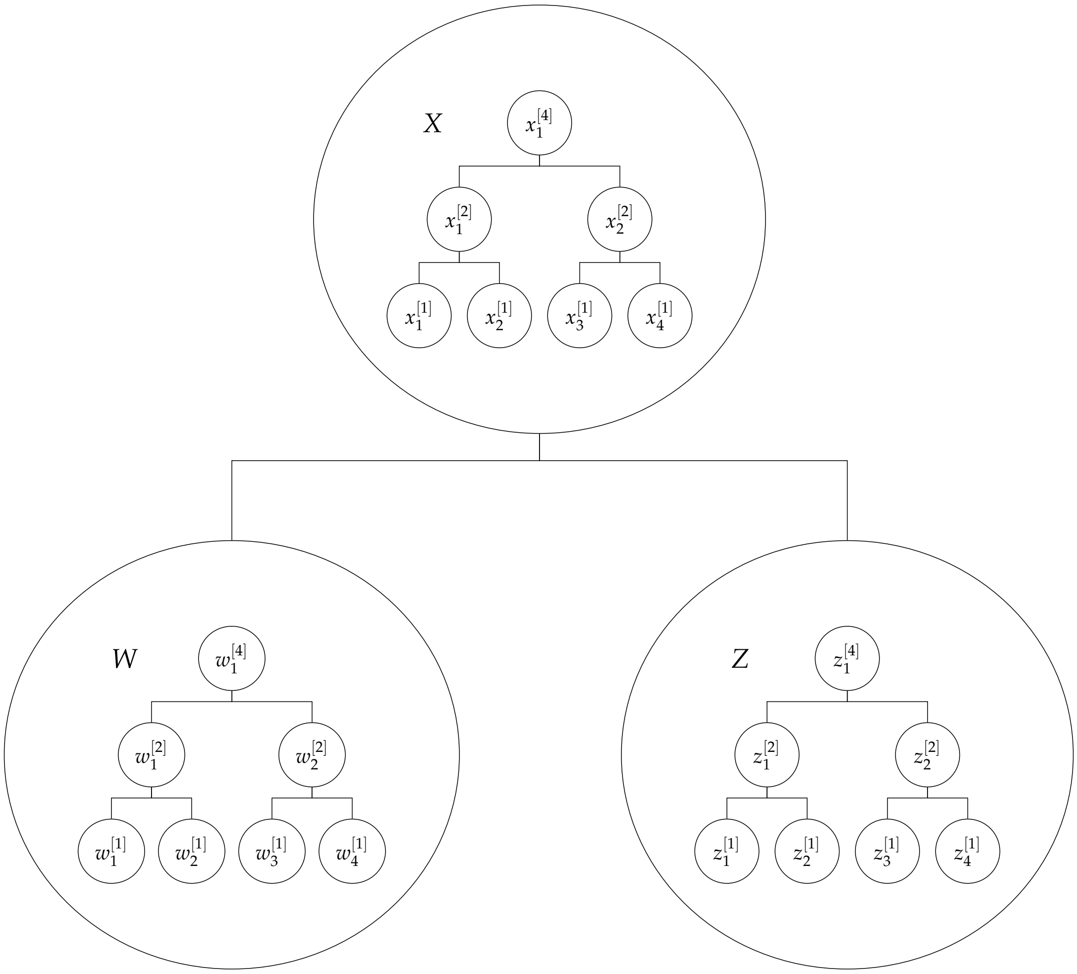


Figure F.3: *A two level cross-temporal hierarchy with quarterly data*

All the nodes in the cross-temporal hierarchy can be expressed in terms of the quarterly bottom time series $w_t^{[1]}$ and $z_t^{[1]}$, $t = 1, \ldots, 4$, according to the structural representation:

$$
\underbrace{\begin{bmatrix} x_1^{[4]} \\ x_1^{[2]} \\ x_2^{[2]} \\ x_1^{[1]} \\ x_2^{[1]} \\ x_3^{[1]} \\ x_4^{[1]} \\ w_1^{[4]} \\ w_1^{[2]} \\ w_2^{[2]} \\ z_1^{[4]} \\ z_1^{[2]} \\ z_2^{[2]} \\ w_1^{[1]} \\ w_2^{[1]} \\ w_3^{[1]} \\ w_4^{[1]} \\ z_1^{[1]} \\ z_2^{[1]} \\ z_3^{[1]} \\ z_4^{[1]} \end{bmatrix}}_{\check{\boldsymbol{y}}} = \underbrace{\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ & & & \boldsymbol{I}_8 & & & & \end{bmatrix}}_{\check{\boldsymbol{S}}} \underbrace{\begin{bmatrix} w_1^{[1]} \\ w_2^{[1]} \\ w_3^{[1]} \\ w_4^{[1]} \\ z_1^{[1]} \\ z_2^{[1]} \\ z_3^{[1]} \\ z_4^{[1]} \end{bmatrix}}_{\boldsymbol{b}},
$$

where $\boldsymbol{a} = \begin{bmatrix} x_1^{[4]} \; x_1^{[2]} \; x_2^{[2]} \; x_1^{[1]} \; x_2^{[1]} \; x_3^{[1]} \; x_4^{[1]} \end{bmatrix}'$, $\boldsymbol{b} = \begin{bmatrix} w_1^{[1]} \; w_2^{[1]} \; w_3^{[1]} \; w_4^{[1]} \; z_1^{[1]} \; z_2^{[1]} \; z_3^{[1]} \; z_4^{[1]} \end{bmatrix}'$, $\check{\boldsymbol{y}} = \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix}$, $\check{\boldsymbol{S}} = \begin{bmatrix} \check{\boldsymbol{C}} \\ \boldsymbol{I}_8 \end{bmatrix}$, and $\check{\boldsymbol{C}}$ is the (13×8) matrix:

$$
\check{C} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}.
$$

The zero constraints valid for the nodes of the cross-temporal hierarchy can be represented through the (13×21) matrix $\check{H}' = \begin{bmatrix} I_{13} & -\check{C} \end{bmatrix}$, which has full row-rank, and is such that:

$$
\check{H}'\check{y} = \mathbf{0}_{(13\times 1)}. \tag{F.7}
$$

According to the notation used so far, it is $n_a = 1$, $n_b = 2$, $T = m = 4$, $N = 1$, $p = 3$, and $\mathcal{K} = \{4, 2, 1\}$. The contemporaneous aggregation matrix C, mapping bts into uts, is simply a (1×2) row vector of ones: $C = [1\ 1]$, and thus U' is the (1×3) row vector $U' = [1\ \ -1\ \ -1]$. Furthermore, the (3×4) temporal aggregation matrix K_1 mapping a quarterly series into its semi-annual and annual counterparts, and the related (3×7) matrix $Z_1' = [I_3\ \ -K_1]$, are given by:

$$
K_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \qquad Z_1' = \begin{bmatrix} 1 & 0 & 0 & -1 & -1 & -1 & -1 \\ 0 & 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & -1 \end{bmatrix}.
$$

The (3×7) matrix Y, collecting all the time series at any observation frequency for a single time cycle, is given by:

$$
Y = \begin{bmatrix} x_1^{[4]} & x_1^{[2]} & x_2^{[2]} & x_1^{[1]} & x_2^{[1]} & x_3^{[1]} & x_4^{[1]} \\ w_1^{[4]} & w_1^{[2]} & w_2^{[2]} & w_1^{[1]} & w_2^{[1]} & w_3^{[1]} & w_4^{[1]} \\ z_1^{[4]} & z_1^{[2]} & z_2^{[2]} & z_1^{[1]} & z_2^{[1]} & z_3^{[1]} & z_4^{[1]} \end{bmatrix},
$$

and then $y = \text{vec}(Y')$ is the (21×1) vector

$$
y = \begin{bmatrix} x_1^{[4]}\ x_1^{[2]}\ x_2^{[2]}\ x_1^{[1]}\ x_2^{[1]}\ x_3^{[1]}\ x_4^{[1]}\ w_1^{[4]}\ w_1^{[2]}\ w_2^{[2]}\ w_1^{[1]}\ w_2^{[1]}\ w_3^{[1]}\ w_4^{[1]}\ z_1^{[4]}\ z_1^{[2]}\ z_2^{[2]}\ z_1^{[1]}\ z_2^{[1]}\ z_3^{[1]}\ z_4^{[1]} \end{bmatrix}',
$$

which is differently organized as compared to $\check{\boldsymbol{y}}$. However, it is easy to show that $\boldsymbol{y} = \boldsymbol{Q}\check{\boldsymbol{y}}$, where $\boldsymbol{Q}$ is the (21×21) permutation matrix

$$
\boldsymbol{Q} = \left[\begin{array}{ccccccccccccccccccccc}
\multicolumn{10}{c}{\boldsymbol{I}_{10}} & \multicolumn{11}{c}{\boldsymbol{0}_{(10\times 11)}} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\multicolumn{17}{c}{\boldsymbol{0}_{(4\times 17)}} & \multicolumn{4}{c}{\boldsymbol{I}_4}
\end{array}\right].
$$

Given the orthogonality of matrix $\boldsymbol{Q}$, it is $\check{\boldsymbol{y}} = \boldsymbol{Q}'\boldsymbol{y}$, and then the constraints (F.7) can be re-stated as $\check{\boldsymbol{H}}'\boldsymbol{Q}'\boldsymbol{y} = \boldsymbol{0}_{(13\times 1)}$, that is

$$
\boldsymbol{H}'\boldsymbol{y} = \boldsymbol{0}_{(13\times 1)},
$$

where $\boldsymbol{H}' = \left(\boldsymbol{Q}\check{\boldsymbol{H}}\right)'$ is a (13×21) full row-rank matrix.

The cross-temporal constraints can be formulated also by simultaneously considering the complete cross-sectional and temporal constraints, as $\check{\boldsymbol{H}}'\boldsymbol{y} = \boldsymbol{0}_{(16\times 1)}$, where matrix

$$
\check{\boldsymbol{H}}' = \begin{bmatrix} \boldsymbol{U}' \otimes \boldsymbol{I}_7 \\ \boldsymbol{I}_3 \otimes \boldsymbol{Z}_1' \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_7 & -\boldsymbol{I}_7 & -\boldsymbol{I}_7 \\ \boldsymbol{Z}_1' & \boldsymbol{0} & \boldsymbol{0}_1 \\ \boldsymbol{0} & \boldsymbol{Z}_1' & \boldsymbol{0}_1 \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{Z}_1' \end{bmatrix}.
$$

has dimension (16×21). It is easy to show that the rank of $\check{\boldsymbol{H}}'$ is 13, which means that the matrix is not full row-rank. The choice of the rows to remove is not unique[2], and in real life applications the elimination of linear dependent relationships from the cross-temporal constraint set might be not as simple as in this toy example. In general, the computation of $\boldsymbol{H}'$ as proposed in section 5.2 of the main paper seems rather quick and effective. In this toy example, the resulting $\boldsymbol{H}'$ matrix according to that procedure is simply matrix $\check{\boldsymbol{H}}'$ without the first three rows, that is:

$$
\boldsymbol{H}' = \begin{bmatrix} \boldsymbol{I}^* & -\boldsymbol{I}^* & -\boldsymbol{I}^* \\ \boldsymbol{Z}_1' & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Z}_1' & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{Z}_1' \end{bmatrix},
$$

where $\boldsymbol{I}^*$ is the (4×7) matrix

$$
\boldsymbol{I}^* = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.
$$

[2]For example, in this simple example a different full row-rank $\boldsymbol{H}'$ can be obtained either by removing the last three rows, or by eliminating rows 5, 10 and 15 from matrix $\check{\boldsymbol{H}}'$.

G An alternative heuristic cross-temporal reconciliation procedure

Let us consider a cross-temporal reconciliation procedure based on the reversal of the order in which the one-dimension forecast reconciliation procedures are applied by KA. The procedure consists in the following steps (it is assumed $h = 1$):

Step 1
Transform $\widehat{\boldsymbol{Y}}$ by computing time-by-time cross-sectional reconciled forecasts $\breve{\boldsymbol{Y}}$ for all the temporal aggregation levels:

$$\widehat{\boldsymbol{Y}} \quad \rightarrow \quad \breve{\boldsymbol{Y}}.$$

The $[n \times (k^* + m)]$ matrix $\widehat{\boldsymbol{Y}}$ can be re-written also as:

$$\widehat{\boldsymbol{Y}} = \left[\widehat{\boldsymbol{Y}}^{[m]} \; \widehat{\boldsymbol{Y}}^{[k_{p-1}]} \ldots \widehat{\boldsymbol{Y}}^{[k_2]} \; \widehat{\boldsymbol{Y}}^{[1]}\right],$$

where $\widehat{\boldsymbol{Y}}^{[k]}$, $k \in \mathcal{K}$, has dimension $(n \times M_k)$. Cross-sectionally reconciled forecasts can be computed by transforming each $\widehat{\boldsymbol{Y}}^{[k]}$ as:

$$\breve{\boldsymbol{Y}}^{[k]} = \boldsymbol{M}^{[k]} \widehat{\boldsymbol{Y}}^{[k]}, \quad k \in \mathcal{K},$$

where $\boldsymbol{M}^{[k]}$ are p transformation matrices, each of dimension $(n \times n)$, given by:

$$\boldsymbol{M}^{[k]} = \boldsymbol{I}_n - \boldsymbol{W}^{[k]} \boldsymbol{U} \left(\boldsymbol{U}' \boldsymbol{W}^{[k]} \boldsymbol{U}\right)^{-1} \boldsymbol{U}', \quad k \in \mathcal{K},$$

and $\boldsymbol{W}^{[k]}$ is a $(n \times n)$ p.d. known matrix. Since it is $\boldsymbol{U}'\boldsymbol{M}^{[k]} = \boldsymbol{0}_{(n_a \times n)}$, $k \in \mathcal{K}$, the reconciled forecasts are cross-sectionally coherent, i.e. $\boldsymbol{U}'\breve{\boldsymbol{Y}} = \boldsymbol{0}_{[n_a \times (k^*+m)]}$, but not temporally: $\boldsymbol{Z}_1' \breve{\boldsymbol{Y}}' \neq \boldsymbol{0}_{(k^* \times n)}$.

Step 2
For each individual variable, compute the temporally reconciled forecasts $\check{\boldsymbol{Y}}$:

$$\breve{\boldsymbol{Y}} \quad \rightarrow \quad \check{\boldsymbol{Y}}.$$

This result can be obtained by applying the point forecast reconciliation formula according to temporally hierarchies to each column of matrix

$$\breve{\boldsymbol{Y}}' = \begin{bmatrix} \breve{\boldsymbol{t}}_{a_1} & \cdots & \breve{\boldsymbol{t}}_{a_{n_a}} & \breve{\boldsymbol{t}}_{b_1} & \cdots & \breve{\boldsymbol{t}}_{b_{n_b}} \\ \breve{\boldsymbol{a}}_1^{[1]} & \cdots & \breve{\boldsymbol{a}}_{n_a}^{[1]} & \breve{\boldsymbol{b}}_1^{[1]} & \cdots & \breve{\boldsymbol{b}}_{n_b}^{[1]} \end{bmatrix}.$$

The n_a vectors of temporally reconciled forecasts of the uts can be obtained as:

$$\begin{bmatrix} \check{\boldsymbol{t}}_{a_j} \\ \check{\boldsymbol{a}}_j^{[1]} \end{bmatrix} = \boldsymbol{M}_{a_j} \begin{bmatrix} \breve{\boldsymbol{t}}_{a_j} \\ \breve{\boldsymbol{a}}_j^{[1]} \end{bmatrix}, \quad \boldsymbol{M}_{a_j} = \boldsymbol{I}_{k^*+m} - \boldsymbol{\Omega}_{a_j} \boldsymbol{Z}_1 \left(\boldsymbol{Z}_1' \boldsymbol{\Omega}_{a_j} \boldsymbol{Z}_1\right)^{-1} \boldsymbol{Z}_1', \quad j = 1, \ldots, n_a.$$

Likeways, the n_b vectors of temporally reconciled forecasts of the bts are given by:

$$\begin{bmatrix} \check{\boldsymbol{t}}_{b_i} \\ \check{\boldsymbol{b}}_i^{[1]} \end{bmatrix} = \boldsymbol{M}_{b_i} \begin{bmatrix} \breve{\boldsymbol{t}}_{b_i} \\ \breve{\boldsymbol{b}}_i^{[1]} \end{bmatrix}, \quad \boldsymbol{M}_{b_i} = \boldsymbol{I}_{k^*+m} - \boldsymbol{\Omega}_{b_i} \boldsymbol{Z}_1 \left(\boldsymbol{Z}_1' \boldsymbol{\Omega}_{b_i} \boldsymbol{Z}_1\right)^{-1} \boldsymbol{Z}_1', \quad i = 1, \ldots, n_b,$$

where the $n_a + n_b$ matrices $\boldsymbol{M}_{a_j}$ and $\boldsymbol{M}_{b_i}$ have dimension $[(k^* + m) \times (k^* + m)]$, and each $\boldsymbol{\Omega}_{a_j}$, $j = 1, \ldots, n_a$, and $\boldsymbol{\Omega}_{b_i}$, $i = 1, \ldots, n_b$, respectively, is a known p.d. $[(k^* + m) \times (k^* + m)]$ matrix.

The mapping performing the transformation of the base forecasts into the temporally reconciled ones can be expressed in compact form as:

$$\text{vec}\left(\breve{\boldsymbol{Y}}'\right) = \begin{bmatrix} \boldsymbol{M}_{a_1} & \cdots & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{M}_{a_{n_a}} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{M}_{b_1} & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{M}_{b_{n_b}} \end{bmatrix} \text{vec}\left(\breve{\boldsymbol{Y}}'\right).$$

The temporally reconciled forecasts can be then collected in the matrix $\breve{\boldsymbol{Y}}'$:

$$\breve{\boldsymbol{Y}}' = \begin{bmatrix} \breve{\boldsymbol{t}}_{a_1} & \cdots & \breve{\boldsymbol{t}}_{a_{n_a}} & \breve{\boldsymbol{t}}_{b_1} & \cdots & \breve{\boldsymbol{t}}_{b_{n_b}} \\ \breve{\boldsymbol{a}}_1^{[1]} & \cdots & \breve{\boldsymbol{a}}_{n_a}^{[1]} & \breve{\boldsymbol{b}}_1^{[1]} & \cdots & \breve{\boldsymbol{b}}_{n_b}^{[1]} \end{bmatrix} = \begin{bmatrix} (\breve{\boldsymbol{A}}^{[m]})' & (\breve{\boldsymbol{B}}^{[m]})' \\ \vdots & \vdots \\ (\breve{\boldsymbol{A}}^{[k_2]})' & (\breve{\boldsymbol{B}}^{[k_2]})' \\ (\breve{\boldsymbol{A}}^{[1]})' & (\breve{\boldsymbol{B}}^{[1]})' \end{bmatrix},$$

which is in line with the temporal aggregation constraints, i.e. $\boldsymbol{Z}_1'\breve{\boldsymbol{Y}}' = \boldsymbol{0}_{(k^* \times n)}$, but in general it is not in line with the cross-sectional (contemporaneous) constraints: $\boldsymbol{U}'\breve{\boldsymbol{Y}} \neq \boldsymbol{0}_{n_a \times (k^*+m)}$.

Step 3

Transform again the step 1 forecasts $\breve{\boldsymbol{Y}}$, by computing temporally reconciled forecasts for all n variables using the $[(k^* + m) \times (k^* + m)]$ matrix $\overline{\boldsymbol{M}}^{\text{cst}}$, where 'cst' stands for 'cross-sectional-then-temporal', given by the average of the matrices $\boldsymbol{M}_i$ obtained at step 2:

$$\breve{\boldsymbol{Y}} \quad \Rightarrow \widetilde{\boldsymbol{Y}}^{\text{cst}}.$$

Matrix $\overline{\boldsymbol{M}}^{\text{cst}}$ can be expressed as:

$$\overline{\boldsymbol{M}}^{\text{cst}} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{M}_i.$$

The final cross-temporal reconciled forecasts are given by:

$$\widetilde{\boldsymbol{Y}}^{\text{cst}} = \left(\overline{\boldsymbol{M}}^{\text{cst}}\breve{\boldsymbol{Y}}'\right)' = \breve{\boldsymbol{Y}}(\overline{\boldsymbol{M}}^{\text{cst}})'. \tag{G.8}$$

Since $\boldsymbol{U}'\breve{\boldsymbol{Y}} = \boldsymbol{0}_{[n_a \times (k^*+m)]}$, and $\boldsymbol{Z}_1'\overline{\boldsymbol{M}}^{\text{cst}} = n^{-1}\sum_{i=1}^{n}\boldsymbol{Z}_1'\boldsymbol{M}_i = \boldsymbol{0}_{[k^* \times (k^*+m)]}$, the reconciled forecasts (G.8) fulfill both cross-sectional and temporal aggregation constraints:

$$\boldsymbol{U}'\widetilde{\boldsymbol{Y}}^{\text{cst}} = \boldsymbol{U}'\breve{\boldsymbol{Y}}(\overline{\boldsymbol{M}}^{\text{cst}})' = \boldsymbol{0}_{[n_a \times (k^*+m)]},$$

$$\boldsymbol{Z}_1'\left(\widetilde{\boldsymbol{Y}}^{\text{cst}}\right)' = \boldsymbol{Z}_1'\overline{\boldsymbol{M}}^{\text{cst}}\breve{\boldsymbol{Y}}' = \boldsymbol{0}_{(k^* \times n)}.$$

H Average relative accuracy indices for selected groups of variables/time frequencies/forecast horizons, in a rolling forecast experiment

Let

$$\hat{e}^{[k],h}_{i,j,t} = y^{[k]}_{i,t+h} - \hat{y}^{[k],h}_{i,j,t}, \quad \begin{array}{l} i = 1,\dots,n, \\ j = 0,\dots,J, \end{array} \quad t = 1,\dots,q, \quad \begin{array}{l} k \in \mathcal{K}, \\ h = 1,\dots,h_k, \end{array}$$

be the forecast error, where y and $\hat{y}$ are the actual and the forecasted values, respectively, suffix i denotes the variable of interest, j is the forecasting technique, where $j = 0$ is the benchmark forecasting procedure, t is the forecast origin, $\mathcal{K}$ is the set of the time frequencies at which the series is observed, and h is the forecast horizon, whose lead time depends on the time frequency k.

Denote by $\mathrm{A}^{[k],h}_{i,j}$ the forecasting accuracy of the technique j, computed across q forecast origins, for the h-step-ahead forecasts of the variable i at the temporal aggregation level k. For example, $\mathrm{A}^{[k],h}_{i,j} = MSE^{[k],h}_{i,j}$, otherwise we might have $\mathrm{A}^{[k],h}_{i,j} = MAE^{[k],h}_{i,j}$ or $\mathrm{A}^{[k],h}_{i,j} = RMSE^{[k],h}_{i,j}$, where

$$\begin{aligned} MAE^{[k],h}_{i,j} &= \frac{1}{q}\sum_{t=1}^{q}\left|\hat{e}^{[k],h}_{i,j,t}\right| \\ RMSE^{[k],h}_{i,j} &= \sqrt{\frac{1}{q}\sum_{t=1}^{q}\left(\hat{e}^{[k],h}_{i,j,t}\right)^2} \end{aligned}$$

In any case, we consider the relative version of the accuracy index $A^{[k],h}_{i,j}$, given by:

$$r^{[k],h}_{i,j} = \frac{A^{[k],h}_{i,j}}{A^{[k],h}_{i,0}}, \quad i = 1,\dots,n, \quad j = 0,\dots,J, \quad k \in \mathcal{K}, \quad h = 1,\dots,h_k,$$

and use it to compute the Average relative accuracy index of the forecasting procedure j, for given k and h, through the geometric mean:

$$\mathrm{AvgRelA}^{[k],h}_{j} = \left(\prod_{i=1}^{n} r^{[k],h}_{i,j}\right)^{\frac{1}{n}}, \quad j = 0,\dots,J.$$

We may consider the following average relative accuracy indices for selected groups of variables/time frequencies and forecast horizons:

Average relative accuracy indices for a single variable at a given time frequency, for multiple forecast horizons

$$\mathrm{AvgRelA}^{[k],q_1:q_2}_{i,j} = \left(\prod_{h=q_1}^{q_2} r^{[k],h}_{i,j}\right)^{\frac{1}{q_2-q_1+1}}, \quad \begin{array}{ll} i = 1,\dots,n, & k \in \mathcal{K}, \\ j = 0,\dots,J, & 1 \le q_1 \le q_2 \le h_k \end{array}.$$

Average relative accuracy indices for a group of variables (either all, or selected groups, e.g. a: uts, b: bts) at a given time frequency, either for a single forecast horizon or across them

$$
\begin{aligned}
\text{AvgRelA}_{j}^{[k],h} &= \left(\prod_{i=1}^{n} r_{i,j}^{[k],h}\right)^{\frac{1}{n}}, & j = 0,\dots,J,\ k \in \mathcal{K},\ h = 1,\dots,h_k \\
\text{AvgRelA}_{a,j}^{[k],h} &= \left(\prod_{i=1}^{n_a} r_{i,j}^{[k],h}\right)^{\frac{1}{n_a}}, & j = 0,\dots,J,\ k \in \mathcal{K} \\
\text{AvgRelA}_{b,j}^{[k],h} &= \left(\prod_{i=n_a+1}^{n} r_{i,j}^{[k],h}\right)^{\frac{1}{n_b}}, & j = 0,\dots,J,\ k \in \mathcal{K} \\
\text{AvgRelA}_{j}^{[k]} &= \left(\prod_{i=1}^{n}\prod_{h=1}^{h_k} r_{i,j}^{[k],h}\right)^{\frac{1}{nh_k}}, & j = 0,\dots,J,\ k \in \mathcal{K} \\
\text{AvgRelA}_{a,j}^{[k]} &= \left(\prod_{i=1}^{n_a}\prod_{h=1}^{h_k} r_{i,j}^{[k],h}\right)^{\frac{1}{n_a h_k}}, & j = 0,\dots,J,\ k \in \mathcal{K} \\
\text{AvgRelA}_{b,j}^{[k]} &= \left(\prod_{i=n_a+1}^{n}\prod_{h=1}^{h_k} r_{i,j}^{[k],h}\right)^{\frac{1}{n_b h_k}}, & j = 0,\dots,J,\ k \in \mathcal{K}
\end{aligned}
$$

Average relative accuracy indices for a single variable or for a group of variables (all, a: uts, b: bts), across all time frequencies and forecast horizons

$$
\begin{aligned}
\text{AvgRelA}_{i,j} &= \left(\prod_{k\in\mathcal{K}}\prod_{h=1}^{h_k} \mathrm{r}_{i,j}^{[k],h}\right)^{\frac{1}{k^*+m}}, & \begin{array}{l} i = 1,\dots,n \\ j = 0,\dots,J \end{array} \\
\text{AvgRelA}_{j} &= \left(\prod_{i=1}^{n}\prod_{k\in\mathcal{K}}\prod_{h=1}^{h_k} \mathrm{r}_{i,j}^{[k],h}\right)^{\frac{1}{n(k^*+m)}}, & j = 0,\dots,J \\
\text{AvgRelA}_{a,j} &= \left(\prod_{i=1}^{n_a}\prod_{k\in\mathcal{K}}\prod_{h=1}^{h_k} \mathrm{r}_{i,j}^{[k],h}\right)^{\frac{1}{n_a(k^*+m)}}, & j = 0,\dots,J \\
\text{AvgRelA}_{b,j} &= \left(\prod_{i=n_a+1}^{n}\prod_{k\in\mathcal{K}}\prod_{h=1}^{h_k} \mathrm{r}_{i,j}^{[k],h}\right)^{\frac{1}{n_b(k^*+m)}}, & j = 0,\dots,J
\end{aligned}
$$

I AvgRelMSE for selected upper time series

In Table I.1 the AvgRelMSE's for selected upper time series and reconciliation procedures are shown. The series we analyze (see Figure I.4) come from the first three levels of both the Income and Expenditure sides hierarchies:

- Gross Domestic Product
- Total Factor Income (Income side)
- Gross Operating Surplus (Income side)
- Compensation of Employees (Income side)
- Gross National Expenditure (Expenditure side)

Figure I.4: *Quarterly GDP and selected time series from both Income and Expenditure sides: actual values and one-step-ahead base forecasts during the testing period (1994:Q4 - 2018:Q1)*

- Domestic Final Demand (Expenditure side)
- Changes in Inventories (Expenditure sides)
- Final Consumption Expenditures (Expenditure side)
- Gross Fixed Capital Formation (Expenditure side)

The ability of cs-shr to improve on short-term (1 or 2-quarter ahead) base forecasts clearly emerges, with the only exception of the forecasts of the Change in Inventories series, where most indices at quarterly level are greater than 1. However, this bad performance is shared by the other reconciliation procedures as well, and is likely due to the low quality of the base forecasts as compared to the other considered series (see Figure I.4).

More detailed supplementary tables and graphs are available on request from the corresponding author.

Table I.1: *AvgRelMSE at any temporal aggregation level and any forecast horizon for selected upper time series and reconciliation procedures.*

Reconciliation procedure	Quarterly 1	Quarterly 2	Quarterly 3	Quarterly 4	Quarterly 1-4	Semi-annual 1	Semi-annual 2	Semi-annual 1-2	Annual 1	All
Gross Domestic Product										
cs-shr	**0.9740**	**0.9397**	0.9028	**0.8924**	**0.9267**	0.8382	0.8713	0.8546	0.7116	0.8719
t-acov	**1.0883**	**1.0356**	**1.0108**	1.0000	**1.0331**	0.6539	0.8717	0.7549	0.6047	0.8750
kah-wlsv-shr	**1.1249**	0.9876	0.9068	0.8719	0.9681	0.6510	**0.7698**	0.7079	0.5485	0.8163
ite-acov-shr	**1.0503**	0.9808	0.9027	0.8853	0.9526	0.6281	0.7730	0.6968	0.5427	0.8039
oct-acov	**1.0696**	0.9689	**0.8975**	0.8926	0.9545	**0.6245**	0.7745	**0.6954**	**0.5402**	**0.8039**
Total Factor Income										
cs-shr	**0.8316**	**0.9002**	0.8769	0.8335	**0.8600**	0.8232	0.8760	0.8492	0.7162	0.8348
t-acov	**1.0434**	**1.0927**	0.9971	0.9818	**1.0279**	0.7174	0.9408	0.8215	0.6636	0.9057
kah-wlsv-shr	0.9598	0.9523	0.8696	0.7984	0.8925	0.6353	**0.7870**	0.7071	0.5680	0.7829
ite-acov-shr	0.8995	0.9428	**0.8663**	**0.8134**	0.8792	0.6141	0.7909	0.6969	0.5642	0.7722
oct-acov	0.8819	0.9335	0.8635	0.8131	0.8719	**0.6078**	0.7907	**0.6932**	**0.5603**	**0.7666**
Gross Operating Surplus										
cs-shr	**0.9170**	**0.8834**	0.9140	0.9008	**0.9037**	**1.0425**	**1.0489**	**1.0457**	0.9354	0.9468
t-acov	**1.0180**	0.9768	0.9760	0.9459	0.9789	0.8958	**1.1015**	0.9933	0.8807	0.9682
kah-wlsv-shr	0.9867	0.9139	0.8988	**0.8717**	0.9168	0.8572	**1.0133**	0.9320	0.8134	0.9055
ite-acov-shr	0.9673	0.8943	**0.8985**	0.8810	0.9097	**0.8338**	**1.0147**	**0.9199**	**0.8083**	**0.8973**
oct-acov	0.9524	0.9233	0.9181	0.8826	0.9187	0.8534	**1.0301**	0.9376	0.8262	0.9102
Compensation of Employees										
cs-shr	**0.9416**	**0.9880**	**1.0172**	**1.0112**	**0.9891**	**1.0519**	**1.0820**	**1.0669**	**1.0488**	**1.0192**
t-acov	**1.0635**	**1.0506**	**1.0593**	**1.0365**	**1.0524**	0.7474	0.8618	**0.8026**	0.5876	0.8962
kah-wlsv-shr	**1.0893**	**1.0739**	**1.0886**	**1.0330**	**1.0709**	0.7663	0.8726	0.8177	0.5932	0.9112
ite-acov-shr	**1.0060**	**1.0417**	**1.0778**	**1.0668**	**1.0477**	**0.7326**	0.8859	0.8056	0.5931	0.8961
oct-acov	**1.0585**	**1.0662**	**1.0576**	**1.0251**	**1.0517**	0.7560	**0.8567**	0.8048	**0.5853**	**0.8960**
Gross National Expenditure										
cs-shr	**0.9243**	**0.9407**	0.9212	**0.8897**	**0.9188**	0.9865	0.8728	0.9280	0.9302	0.9230
t-acov	**1.0197**	**1.0367**	**1.0113**	**1.0060**	**1.0184**	0.8447	0.9008	0.8723	0.6630	0.9164
kah-wlsv-shr	0.9966	0.9959	0.9265	0.9017	0.9542	0.8284	0.8113	0.8198	0.6156	0.8583
ite-acov-shr	0.9723	0.9925	**0.9155**	0.9094	0.9467	0.8244	0.8158	0.8201	0.6156	**0.8545**
oct-acov	**1.0071**	**1.0002**	0.9278	0.9031	0.9585	**0.8193**	**0.8075**	**0.8133**	**0.6064**	0.8567
Domestic Final Demand										
cs-shr	**0.8713**	**0.9737**	**1.0182**	0.9958	**0.9631**	0.9787	**1.0192**	0.9988	**1.0038**	0.9789
t-acov	0.9844	**1.0002**	**1.0031**	**0.9851**	0.9932	0.8656	**0.9421**	0.9030	0.6745	0.9146
kah-wlsv-shr	0.9112	**1.0152**	**1.0136**	**1.0039**	0.9850	0.8184	0.9562	0.8846	0.6747	0.9049
ite-acov-shr	0.8843	**1.0014**	**1.0049**	**1.0161**	0.9751	**0.8119**	0.9662	0.8857	0.6758	0.9003
oct-acov	0.9274	**1.0114**	**1.0088**	0.9956	0.9852	0.8142	0.9428	**0.8761**	**0.6608**	**0.8999**
Changes in Inventories										
cs-shr	**1.0791**	**1.0228**	**1.0412**	**0.9250**	**1.0154**	0.7215	0.8134	0.7661	0.8811	0.9181
t-acov	**1.0382**	**1.0609**	**1.0032**	0.9999	**1.0253**	0.6886	0.7098	0.6991	0.8996	0.9020
kah-wlsv-shr	**1.0204**	**1.0339**	**1.0163**	0.9467	**1.0037**	**0.6644**	0.6795	**0.6719**	**0.8369**	0.8720
ite-acov-shr	**1.0317**	**1.0239**	**0.9908**	0.9285	**0.9929**	0.6776	**0.6676**	0.6726	0.8407	**0.8674**
oct-acov	**1.0074**	**1.0401**	**1.0087**	0.9540	**1.0021**	0.6813	0.7051	0.6931	0.9084	0.8894
Final Consumption Expenditures										
cs-shr	**0.8826**	**0.8184**	**0.8216**	**0.8223**	**0.8358**	0.9482	0.9741	0.9611	0.9988	0.8923
t-acov	0.9956	**1.0268**	0.9982	**1.0199**	**1.0100**	0.8978	0.9456	0.9214	0.7540	0.9436
kah-wlsv-shr	0.9370	0.9094	0.9000	0.8956	0.9104	0.8077	0.8395	0.8234	0.6708	0.8469
ite-acov-shr	0.9263	0.8804	0.8963	0.8913	0.8984	**0.7861**	**0.8331**	**0.8093**	**0.6593**	**0.8343**
oct-acov	0.9691	0.9489	0.9310	0.9373	0.9464	0.8277	0.8673	0.8473	0.6888	0.8763
Gross Fixed Capital Formation										
cs-shr	**0.9442**	**0.9828**	**1.0156**	**1.0096**	0.9876	**1.0225**	**1.0185**	**1.0205**	0.9719	0.9946
t-acov	0.9875	**1.0066**	0.9967	0.9653	0.9889	0.8881	**1.0002**	0.9425	0.7258	0.9333
kah-wlsv-shr	0.9875	0.9790	0.9768	0.9663	0.9774	0.8480	0.9829	0.9130	0.7052	0.9149
ite-acov-shr	0.9524	0.9827	0.9651	0.9859	0.9714	0.8453	0.9973	0.9182	0.7097	0.9140
oct-acov	0.9498	0.9511	**0.9539**	**0.9426**	**0.9448**	**0.8149**	**0.9476**	**0.8787**	**0.6726**	**0.8816**

References

Ashouri, M., Hyndman, R. J. & Shmueli, G. (2021), 'Fast forecast reconciliation using linear models', *Journal of Computational and Graphical Statistics* .

Bertani, N., Satopää, V. & Jensen, S. (2020), 'Joint Bottom-Up Method for Hierarchical Time-Series: Application to Australian Tourism', *SSRN Electronic Journal* pp. 1–27.

Hyndman, R. J., Lee, A. J., Wang, E. & Wickramasuriya, S. L. (2020), hts: Hierarchical and Grouped Time Series.
URL: *https://cran.r-project.org/package=hts*

Kourentzes, N. & Athanasopoulos, G. (2019), 'Cross-temporal coherent forecasts for Australian tourism', Annals of Tourism Research **75**, 393–409.

Magnus, J. R. & Neudecker, H. (2019), Matrix Differential Calculus with Applications in Statistics and Econometrics, Wiley, New York.

Panagiotelis, A., Athanasopoulos, G., Gamakumara, P. & Hyndman, R. J. (2021), 'Forecast reconciliation: A geometric view with new insights on bias correction', International Journal of Forecasting (37), 343–359.

Wickramasuriya, S. L., Athanasopoulos, G. & Hyndman, R. J. (2019), 'Optimal Forecast Reconciliation for Hierarchical and Grouped Time Series Through Trace Minimization', Journal of the American Statistical Association **114**(526), 804–819.

Wickramasuriya, S. L., Turlach, B. A. & Hyndman, R. J. (2020), 'Optimal non-negative forecast reconciliation', Statistics and Computing **30**(5), 1167–1182.