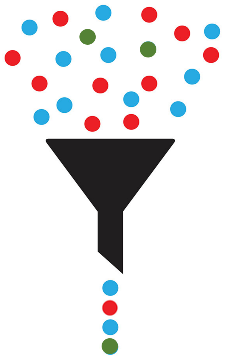
Forecast reconciliation with subset selection



Xiaoqian Wang

In collaboration with: Rob J Hyndman & Shanika Wickramasuriya

Department of Econometrics and Business Statistics

MONASH University

Outline

Outline

Linear forecast reconciliation

$$\tilde{\boldsymbol{y}}_h = \boldsymbol{S}\boldsymbol{G}\hat{\boldsymbol{y}}_h$$

- $\hat{\boldsymbol{y}}_h$: vector of initial h-step-ahead base forecasts made at time T.
- $\boldsymbol{G}$: matrix combining all base forecasts to form bottom-level reconciled forecasts.
- $\boldsymbol{S}$: summing matrix containing the linear constraints.
- $\tilde{\boldsymbol{y}}_h$: vector of coherent linear forecasts.

Single-level approaches

- Bottom-Up: $\boldsymbol{G}_{BU} = \left[\boldsymbol{O}_{n_b \times n_a} \mid \boldsymbol{I}_{n_b}\right]$.
- Top-Down: $\boldsymbol{G}_{TD} = \left[\boldsymbol{p} \mid \boldsymbol{O}_{n_b \times (n-1)}\right]$ and $\sum_{i=1}^{n_b} p_i = 1$.

Minimum trace reconciliation

- Problem: minimizing the trace of the covariance matrix $\mathrm{Var}\left(\boldsymbol{y}_h - \tilde{\boldsymbol{y}}_h\right)$.
- Solution: $\boldsymbol{G} = \left(\boldsymbol{S}'\boldsymbol{W}_h^{-1}\boldsymbol{S}\right)^{-1}\boldsymbol{S}'\boldsymbol{W}_h^{-1}$.
- $\boldsymbol{W}_h$ estimators: OLS, WLSs, WLSv, MinT, MinTs.

Linear forecast reconciliation

$$\tilde{\boldsymbol{y}}_h = \boldsymbol{S}\boldsymbol{G}\hat{\boldsymbol{y}}_h$$

- $\hat{\boldsymbol{y}}_h$: vector of initial h-step-ahead base forecasts made at time T.
- $\boldsymbol{G}$: matrix combining all base forecasts to form bottom-level reconciled forecasts.
- $\boldsymbol{S}$: summing matrix containing the linear constraints.
- $\tilde{\boldsymbol{y}}_h$: vector of coherent linear forecasts.

Question 1

Is there an approach that always dominates the others?

Single-level approaches

- Bottom-Up: $\boldsymbol{G}_{BU} = \left[\boldsymbol{O}_{n_b \times n_a} \mid \boldsymbol{I}_{n_b}\right]$.
- Top-Down: $\boldsymbol{G}_{TD} = \left[\boldsymbol{p} \mid \boldsymbol{O}_{n_b \times (n-1)}\right]$ and $\sum_{i=1}^{n_b} p_i = 1$.

Minimum trace reconciliation

- Problem: minimizing the trace of the covariance matrix $\mathrm{Var}\left(\boldsymbol{y}_h - \tilde{\boldsymbol{y}}_h\right)$.
- Solution: $\boldsymbol{G} = \left(\boldsymbol{S}'\boldsymbol{W}_h^{-1}\boldsymbol{S}\right)^{-1}\boldsymbol{S}'\boldsymbol{W}_h^{-1}$.
- $\boldsymbol{W}_h$ estimators: OLS, WLSs, WLSv, MinT, MinTs.

Intuition behind $\boldsymbol{W}$

The trace minimization problem can be reformulated as a linear equality constrained least squares problem.

Optimization problem

$$\begin{aligned} \min_{\tilde{\boldsymbol{y}}} \quad & \frac{1}{2}(\hat{\boldsymbol{y}} - \tilde{\boldsymbol{y}})'\boldsymbol{W}^{-1}(\hat{\boldsymbol{y}} - \tilde{\boldsymbol{y}}) \\ \text{s.t.} \quad & \tilde{\boldsymbol{y}} = \boldsymbol{S}\tilde{\boldsymbol{b}} \end{aligned}$$

- Generalized Least Squares problem.
- The greater the estimated variance of the base forecast errors, the greater the range of adjustments permitted for reconciliation.
- It's hard to say which estimator of $\boldsymbol{W}$ is better.
- Data of interest & forecast goals.

Some potential issues

- Assume $\boldsymbol{W}_h \approx k_h \boldsymbol{W}_1$, the estimate of $\boldsymbol{G}$ does not change with forecast horizons.
- The long-term reconciled forecasts may perform extremely poorly compared to base forecasts, especially when
  - base forecasts of some series within a hierarchy are of poor quality;
  - model misspecification exists for some series in the hierarchy.

Some potential issues

- Assume $\boldsymbol{W}_h \approx k_h \boldsymbol{W}_1$, the estimate of $\boldsymbol{G}$ does not change with forecast horizons.
- The long-term reconciled forecasts may perform extremely poorly compared to base forecasts, especially when
  - base forecasts of some series within a hierarchy are of poor quality;
  - model misspecification exists for some series in the hierarchy.

Question 2

Can we identify series with poorly-performing forecasts and eliminate their negative effect when implementing reconciliation?

Outline

How to achieve selection?

The purpose

$$\tilde{\boldsymbol{y}}_h = \boldsymbol{S}\boldsymbol{G}\hat{\boldsymbol{y}}_h$$

Eliminate the negative effect of some series on forecast reconciliation.

About $\boldsymbol{G}$: Zero out some columns of $\boldsymbol{G}$.

About $\boldsymbol{S}$: Do not zero out the corresponding rows of $\boldsymbol{S}$.

$$\begin{bmatrix} \tilde{y}_{\text{Total}} \\ \tilde{y}_{\text{A}} \\ \tilde{y}_{\text{B}} \\ \tilde{y}_{\text{AA}} \\ \tilde{y}_{\text{AB}} \\ \tilde{y}_{\text{BA}} \\ \tilde{y}_{\text{BB}} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w_{11} & 0 & w_{13} & w_{14} & w_{15} & w_{16} & w_{17} \\ w_{21} & 0 & w_{23} & w_{24} & w_{25} & w_{26} & w_{27} \\ w_{31} & 0 & w_{33} & w_{34} & w_{35} & w_{36} & w_{37} \\ w_{41} & 0 & w_{43} & w_{44} & w_{45} & w_{46} & w_{47} \end{bmatrix} \begin{bmatrix} \hat{y}_{\text{Total}} \\ \hat{y}_{\text{A}} \\ \hat{y}_{\text{B}} \\ \hat{y}_{\text{AA}} \\ \hat{y}_{\text{AB}} \\ \hat{y}_{\text{BA}} \\ \hat{y}_{\text{BB}} \end{bmatrix}$$

Method I: Regularized best-subset selection

Best-subset selection

$$\min_{\boldsymbol{G}} \quad \frac{1}{2}\left(\hat{\boldsymbol{y}} - \boldsymbol{SG}\hat{\boldsymbol{y}}\right)' \boldsymbol{W}^{-1}\left(\hat{\boldsymbol{y}} - \boldsymbol{SG}\hat{\boldsymbol{y}}\right) + \lambda_0 \sum_{j=1}^{n} \mathbf{1}\left(\boldsymbol{G}_{\cdot j} \neq \mathbf{0}\right)$$

$$\text{s.t.} \quad \boldsymbol{GS} = \boldsymbol{I}_{n_b},$$

- $\mathbf{1}(\cdot)$: the indicator function.
- $\lambda_0 > 0$: controls the number of nonzero columns of $\boldsymbol{G}$ selected.
- $\boldsymbol{SG}\hat{\boldsymbol{y}} = \mathrm{vec}\left(\boldsymbol{SG}\hat{\boldsymbol{y}}\right) = \left(\hat{\boldsymbol{y}}' \otimes \boldsymbol{S}\right) \mathrm{vec}(\boldsymbol{G})$
- Group best-subset selection problem with an additional unbiasedness constraint.

Limitation:

- Computationally infeasible
- In low SNR regimes, the vanilla version of ℓ_0 penalization suffers from overfitting.

Method I: Regularized best-subset selection

Best-subset selection with ridge regularization

$$\begin{aligned} \min_{\boldsymbol{G}} \quad & \frac{1}{2}\left(\hat{\boldsymbol{y}} - \left(\hat{\boldsymbol{y}}' \otimes \boldsymbol{S}\right)\operatorname{vec}(\boldsymbol{G})\right)' \boldsymbol{W}^{-1}\left(\hat{\boldsymbol{y}} - \left(\hat{\boldsymbol{y}}' \otimes \boldsymbol{S}\right)\operatorname{vec}(\boldsymbol{G})\right) \\ & + \lambda_0 \sum_{j=1}^{n} \mathbf{1}\left(\boldsymbol{G}_{\cdot j} \neq \mathbf{0}\right) + \lambda_2 \left\|\operatorname{vec}\left(\boldsymbol{G}\right)\right\|_2^2 \\ \text{s.t.} \quad & \boldsymbol{G}\boldsymbol{S} = \boldsymbol{I}_{n_b}, \end{aligned}$$

- $\lambda_2 \geq 0$: controls the strength of the ridge regularization.
- Sparsity & Shrinkage.
- Motivation: Additional ridge regularization can improve the prediction performance of best-subset selection when SNR is low.

Method I: Regularized best-subset selection

Big-M based MIP formulation

$$\min_{\boldsymbol{G},\boldsymbol{z},\check{\boldsymbol{e}},\boldsymbol{g}^+} \frac{1}{2}\check{\boldsymbol{e}}'\boldsymbol{W}_h^{-1}\check{\boldsymbol{e}} + \lambda_0 \sum_{j=1}^{n} z_j + \lambda_2 \boldsymbol{g}^{+\prime}\boldsymbol{g}^+$$

$$\begin{aligned}
\text{s.t.}\quad & \hat{\boldsymbol{y}}_h - \left(\hat{\boldsymbol{y}}_h' \otimes \boldsymbol{S}\right)\operatorname{vec}(\boldsymbol{G}) = \check{\boldsymbol{e}} \quad \cdots (C1)\\
& \boldsymbol{G}\boldsymbol{S} = \boldsymbol{I}_{n_b} \Leftrightarrow \left(\boldsymbol{S}' \otimes \boldsymbol{I}_{n_b}\right)\operatorname{vec}(\boldsymbol{G}) = \operatorname{vec}\left(\boldsymbol{I}_{n_b}\right) \quad \cdots (C2)\\
& \sum_{i=1}^{n_b} g^+_{i+(j-1)n_b} \leqslant \mathcal{M} z_j, \quad j \in [n] \quad \cdots (C3)\\
& \boldsymbol{g}^+ \geqslant \operatorname{vec}(\boldsymbol{G}) \quad \cdots (C4)\\
& \boldsymbol{g}^+ \geqslant -\operatorname{vec}(\boldsymbol{G}) \quad \cdots (C5)\\
& z_j \in \{0,1\}, \quad j \in [n] \quad \cdots (C6)
\end{aligned}$$

- $\mathcal{M}$: a Big-M parameter (a priori specified).
- z_j: a binary variable.

Method I: Regularized best-subset selection

Hyperparameter

- ℓ_0 regularization parameter
  - $\lambda_{0\,\max} = \frac{1}{2}\left(\hat{\boldsymbol{y}}_h - \tilde{\boldsymbol{y}}_h^{\text{bench}}\right)' \boldsymbol{W}_h^{-1}\left(\hat{\boldsymbol{y}}_h - \tilde{\boldsymbol{y}}_h^{\text{bench}}\right)$
  - $\lambda_{0\,\min} = 0.0001\lambda_{0\,\max}$
  - Generate a grid of k values between $\lambda_{0\,\min}$ and $\lambda_{0\,\max}$, where $\lambda_{0,j} = \lambda_{0\,\max}\left(\lambda_{0\,\min}/\lambda_{0\,\max}\right)^{j/(k-1)}$ for $j = 0, \dots, k-1$.
  - $\lambda_0 = \{0, \lambda_{0,0}, \dots, \lambda_{0,k-1}\}$. we set $k = 20$.
- ℓ_2 regularization parameter
  - $\lambda_2 = \{0, 10^{-2}, 10^{-1}, 10^{0}, 10^{1}, 10^{2}\}$

Method II: Intuitive method

The MinT reconciliation matrix: $\boldsymbol{G} = \left(\boldsymbol{S}'\boldsymbol{W}_h^{-1}\boldsymbol{S}\right)^{-1}\boldsymbol{S}'\boldsymbol{W}_h^{-1}$.

Instead of trying to zero out some columns of $\boldsymbol{G}$ as in regularized best-subset selection, We utilize the MinT solution and assume $\bar{\boldsymbol{G}} = (\boldsymbol{S}'\boldsymbol{A}'\boldsymbol{W}^{-1}\boldsymbol{A}\boldsymbol{S})^{-1}\boldsymbol{S}'\boldsymbol{A}'\boldsymbol{W}^{-1}$.

- $\bar{\boldsymbol{S}} = \boldsymbol{A}\boldsymbol{S}$
- $\boldsymbol{A} = \mathrm{diag}(z_i)$ is a diagonal matrix with $z_i \in \{0, 1\}$.
- Estimate the whole $\boldsymbol{G} \Longrightarrow$ estimate $\boldsymbol{A}$.

Intuitive method

$$
\begin{aligned}
\min_{\boldsymbol{A}} \quad & \frac{1}{2}\left(\hat{\boldsymbol{y}} - \boldsymbol{S}\bar{\boldsymbol{G}}\hat{\boldsymbol{y}}\right)'\boldsymbol{W}^{-1}\left(\hat{\boldsymbol{y}} - \boldsymbol{S}\bar{\boldsymbol{G}}\hat{\boldsymbol{y}}\right) + \lambda_0 \sum_{j=1}^{n} z_i \\
\text{s.t.} \quad & \bar{\boldsymbol{G}} = (\boldsymbol{S}'\boldsymbol{A}'\boldsymbol{W}^{-1}\boldsymbol{A}\boldsymbol{S})^{-1}\boldsymbol{S}'\boldsymbol{A}'\boldsymbol{W}^{-1} \\
& \bar{\boldsymbol{G}}\boldsymbol{S} = \boldsymbol{I}
\end{aligned}
$$

Method II: Intuitive method

Example

```r
S <- rbind(c(1,1,1,1), c(1,1,0,0), c(0,0,1,1), diag(1,4))
W_inv <- diag(c(4,2,2,rep(1,4))) |> solve()
G <- solve(t(S) %*% W_inv %*% S) %*% (t(S) %*% W_inv) |> round(2)

A <- diag(c(1,0,rep(1, 5)))
G_bar <- solve(t(A %*% S) %*% W_inv %*% A %*% S) %*% (t(A %*% S) %*% W_inv) |> round(2)
list(G = G, G_bar = G_bar)
```

```
## $G
##      [,1]  [,2]  [,3]  [,4]  [,5]  [,6]  [,7]
## [1,] 0.08  0.21 -0.04  0.71 -0.29 -0.04 -0.04
## [2,] 0.08  0.21 -0.04 -0.29  0.71 -0.04 -0.04
## [3,] 0.08 -0.04  0.21 -0.04 -0.04  0.71 -0.29
## [4,] 0.08 -0.04  0.21 -0.04 -0.04 -0.29  0.71
##
## $G_bar
##      [,1] [,2]  [,3]  [,4]  [,5]  [,6]  [,7]
## [1,] 0.14    0 -0.07  0.86 -0.14 -0.07 -0.07
## [2,] 0.14    0 -0.07 -0.14  0.86 -0.07 -0.07
## [3,] 0.07    0  0.21 -0.07 -0.07  0.71 -0.29
## [4,] 0.07    0  0.21 -0.07 -0.07 -0.29  0.71
```

Method II: Intuitive method

Problem reformulation for intuitive method

$$\min_{\boldsymbol{A},\bar{\boldsymbol{G}},\boldsymbol{C}} \quad \frac{1}{2}\left(\hat{\boldsymbol{y}} - \left(\hat{\boldsymbol{y}}' \otimes \boldsymbol{S}\right)\mathrm{vec}(\bar{\boldsymbol{G}})\right)' \boldsymbol{W}^{-1}\left(\hat{\boldsymbol{y}} - \left(\hat{\boldsymbol{y}}' \otimes \boldsymbol{S}\right)\mathrm{vec}(\bar{\boldsymbol{G}})\right) + \lambda_0 \sum_{j=1}^{n} z_i$$

$$\begin{aligned} \text{s.t.} \quad & \bar{\boldsymbol{G}}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{I} \\ & \bar{\boldsymbol{G}} = \boldsymbol{C}\boldsymbol{S}'\boldsymbol{A}'\boldsymbol{W}^{-1} \\ & \bar{\boldsymbol{G}}\boldsymbol{S} = \boldsymbol{I} \end{aligned}$$

Method II: Intuitive method

Problem reformulation for intuitive method

$$\min_{\boldsymbol{A},\bar{\boldsymbol{G}},\boldsymbol{C}} \quad \frac{1}{2}\left(\hat{\boldsymbol{y}} - \left(\hat{\boldsymbol{y}}' \otimes \boldsymbol{S}\right)\mathrm{vec}(\bar{\boldsymbol{G}})\right)' \boldsymbol{W}^{-1}\left(\hat{\boldsymbol{y}} - \left(\hat{\boldsymbol{y}}' \otimes \boldsymbol{S}\right)\mathrm{vec}(\bar{\boldsymbol{G}})\right) + \lambda_0 \sum_{j=1}^{n} z_i$$

$$\text{s.t.} \quad \bar{\boldsymbol{G}}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{I}$$

$$\bar{\boldsymbol{G}} = \boldsymbol{C}\boldsymbol{S}'\boldsymbol{A}'\boldsymbol{W}^{-1}$$

$$\bar{\boldsymbol{G}}\boldsymbol{S} = \boldsymbol{I}$$

Hyperparameter (ℓ_0 regularization parameter)

- $\lambda_{0\,\max} = \frac{1}{2}\left(\hat{\boldsymbol{y}}_h - \tilde{\boldsymbol{y}}_h^{\text{bench}}\right)' \boldsymbol{W}_h^{-1}\left(\hat{\boldsymbol{y}}_h - \tilde{\boldsymbol{y}}_h^{\text{bench}}\right)$, and $\lambda_{0\,\min} = 0.0001\lambda_{0\,\max}$
- Generate a grid of k values between $\lambda_{0\,\min}$ and $\lambda_{0\,\max}$, where $\lambda_{0,j} = \lambda_{0\,\max}\left(\lambda_{0\,\min}/\lambda_{0\,\max}\right)^{j/(k-1)}$ for $j = 0, \ldots, k-1$.
- $\lambda_0 = \{0, \lambda_{0,0}, \ldots, \lambda_{0,k-1}\}$. we set $k = 20$.

Method III: Group lasso method

Group lasso with the unbiasedness constraint

$$\min_{\boldsymbol{G}} \quad \frac{1}{2}\left(\hat{\boldsymbol{y}} - \left(\hat{\boldsymbol{y}}' \otimes \boldsymbol{S}\right)\mathrm{vec}(\boldsymbol{G})\right)' \boldsymbol{W}^{-1}\left(\hat{\boldsymbol{y}} - \left(\hat{\boldsymbol{y}}' \otimes \boldsymbol{S}\right)\mathrm{vec}(\boldsymbol{G})\right)$$

$$+ \lambda \sum_{j=1}^{n} w_j \left\|\boldsymbol{G}_{\cdot j}\right\|_2$$

$$\text{s.t.} \quad \boldsymbol{G}\boldsymbol{S} = \boldsymbol{I}_{n_b},$$

- $\lambda \geq 0$: tuning parameter.
- w_j: penalty weight in order to make model more flexible.
- The penalty function is intermediate between the ℓ_1-penalty that is used in the lasso and the ℓ_2-penalty that is used in ridge regression.

Method III: Group lasso method

SOCP formulation

$$\min_{\boldsymbol{G},\check{\boldsymbol{e}},\boldsymbol{g}^+} \frac{1}{2}\check{\boldsymbol{e}}'\boldsymbol{W}_h^{-1}\check{\boldsymbol{e}} + \lambda\sum_{j=1}^{n} w_j c_j$$

$$\text{s.t.} \quad \hat{\boldsymbol{y}}_h - \left(\hat{\boldsymbol{y}}_h' \otimes \boldsymbol{S}\right)\mathrm{vec}(\boldsymbol{G}) = \check{\boldsymbol{e}} \quad \cdots (C1)$$

$$c_j = \sqrt{\sum_{i=1}^{n_b} g^2_{i+(j-1)n_b}}, \quad j \in [n] \quad \cdots (C2)$$

$$\boldsymbol{G}\boldsymbol{S} = \boldsymbol{I}_{n_b} \Leftrightarrow \left(\boldsymbol{S}' \otimes \boldsymbol{I}_{n_b}\right)\mathrm{vec}(\boldsymbol{G}) = \mathrm{vec}\left(\boldsymbol{I}_{n_b}\right) \quad \cdots (C3)$$

Method III: Group lasso method

Hyperparameter

- Penalty weights: assign different penalty weights w_j on each group, e.g., $w_j = 1/\left\|\boldsymbol{G}_{\cdot j}^{\text{bench}}\right\|_2$.
- λ sequence.
  - We ignore the unbiasedness constraint,

$$\lambda_{\max} = \max_{j=1,\ldots,n} \left\| -\left(\left(\hat{\boldsymbol{y}}' \otimes \boldsymbol{S}\right)_{\cdot cj}\right)' \boldsymbol{W}^{-1}\hat{\boldsymbol{y}} \right\|_2 / w_j$$

    is the smallest λ value such that all predictors have zero coefficients, i.e., $\boldsymbol{G} = \boldsymbol{O}$.
  - $\lambda_{\min} = 0.0001\lambda_{\max}$.
  - Generate a grid of k values between $\lambda_{\min}$ and $\lambda_{\max}$, $\lambda_j = \lambda_{\max}\left(\lambda_{\min}/\lambda_{\max}\right)^{j/(k-1)}$ for $j = 0, \ldots, k-1$.
  - $\lambda = \{0, \lambda_0, \ldots, \lambda_{k-1}\}$. we set $k = 20$.

Outline

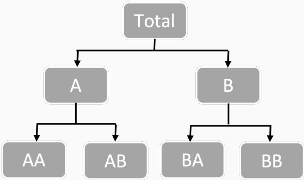
Data generation

The bottom-level series were generated using the basic structural time series model

$$\boldsymbol{b}_t = \boldsymbol{\mu}_t + \boldsymbol{\gamma}_t + \boldsymbol{\eta}_t$$

where $\boldsymbol{\mu}_t, \boldsymbol{\gamma}_t$, and $\boldsymbol{\eta}_t$ are the trend, seasonal, and error components, respectively,

$$\begin{aligned}
\boldsymbol{\mu}_t &= \boldsymbol{\mu}_{t-1} + \boldsymbol{v}_t + \boldsymbol{\varrho}_t, & \boldsymbol{\varrho}_t &\sim \mathcal{N}\left(\mathbf{0}, \sigma_\varrho^2 \boldsymbol{I}_4\right), \\
\boldsymbol{v}_t &= \boldsymbol{v}_{t-1} + \boldsymbol{\zeta}_t, & \boldsymbol{\zeta}_t &\sim \mathcal{N}\left(\mathbf{0}, \sigma_\zeta^2 \boldsymbol{I}_4\right), \\
\boldsymbol{\gamma}_t &= -\sum_{i=1}^{s-1} \boldsymbol{\gamma}_{t-i} + \boldsymbol{\omega}_t, & \boldsymbol{\omega}_t &\sim \mathcal{N}\left(\mathbf{0}, \sigma_\omega^2 \mathbf{I}_4\right),
\end{aligned}$$



and ϱ_t, ζ_t, and ω_t are errors independent of each other and over time.

Simulation setup

Other details

- $s = 4$ for quarterly data, $n = 180$, $h = 16$.
- $\sigma_{\varrho}^2 = 2, \sigma_{\zeta}^2 = 0.007$, and $\sigma_{\omega}^2 = 7$.
- The initial values for $\boldsymbol{\mu}_0, \boldsymbol{v}_0, \gamma_0, \gamma_1, \gamma_2$ were generated independently from a multivariate normal distribution with mean zero and covariance matrix, $\Sigma_0 = I_4$.
- Each component of $\boldsymbol{\eta}_t$ was generated from an $\mathrm{ARIMA}(p, 0, q)$ process with p and q taking values of 0 and 1 with equal probability.
- The bottom-level series were then appropriately summed to obtain the data for higher levels.
- This process was repeated 500 times.

Results: Base forecasts are generated by ETS

Out-of-sample forecast performance (average RMSE).

	Top				Middle				Bottom				Average			
Method	h=1	1-4	1-8	1-16	h=1	1-4	1-8	1-16	h=1	1-4	1-8	1-16	h=1	1-4	1-8	1-16
Base	9.6	10.7	12.6	15.6	6.3	7.3	8.6	10.8	4.2	4.9	5.9	7.5	5.6	6.4	7.6	9.6
BU	-1.0	0.4	0.6	0.7	-0.3	0.0	0.1	0.0	0.0	0.0	0.0	0.0	-0.3	0.1	0.2	0.2
OLS	-0.7	**-0.2**	**0.0**	**0.0**	-0.1	**-0.3**	**-0.2**	**-0.3**	0.1	**-0.2**	**-0.2**	**-0.1**	-0.2	**-0.2**	**-0.2**	**-0.1**
OLS-subset	**-0.8**	0.2	0.3	0.4	**-0.2**	-0.1	0.0	-0.1	0.1	0.1	0.0	0.1	-0.2	0.0	0.1	0.1
OLS-intuitive	**-0.9**	-0.1	0.1	0.2	**-0.2**	**-0.3**	-0.1	-0.1	0.2	0.0	0.0	0.0	-0.2	-0.1	0.0	0.0
OLS-lasso	**-1.3**	-0.1	0.3	0.4	**-0.5**	**-0.3**	-0.1	-0.1	**-0.1**	-0.1	-0.1	**-0.1**	**-0.5**	**-0.2**	0.0	0.0
WLSs	-0.9	-0.1	**0.0**	0.2	-0.3	**-0.3**	**-0.2**	-0.2	0.0	**-0.2**	**-0.2**	**-0.1**	-0.3	**-0.2**	-0.1	**-0.1**
WLSs-subset	**-1.0**	0.1	0.3	0.3	-0.2	-0.2	0.0	-0.1	0.1	0.0	-0.1	0.0	-0.3	-0.1	0.0	0.0
WLSs-intuitive	**-1.0**	-0.1	0.1	0.3	-0.3	**-0.3**	-0.1	-0.1	0.1	-0.1	-0.1	0.0	-0.3	**-0.2**	-0.1	0.0
WLSs-lasso	**-1.3**	0.0	0.3	0.5	**-0.5**	-0.2	0.0	-0.1	**-0.1**	-0.1	-0.1	0.0	**-0.5**	-0.1	0.0	0.1
WLSv	-0.9	-0.1	0.1	0.2	-0.3	**-0.3**	**-0.2**	-0.2	0.0	**-0.2**	**-0.2**	**-0.1**	-0.3	**-0.2**	-0.1	**-0.1**
WLSv-subset	-0.9	0.2	0.4	0.5	-0.3	-0.1	0.1	0.0	0.0	0.0	0.0	0.1	-0.3	0.0	0.1	0.2
WLSv-intuitive	**-1.0**	0.0	0.2	0.3	-0.3	-0.2	-0.1	-0.1	0.0	0.0	0.0	0.0	**-0.4**	-0.1	0.0	0.0
WLSv-lasso	**-1.3**	0.0	0.3	0.5	**-0.5**	-0.2	0.0	-0.1	**-0.1**	-0.1	-0.1	0.0	**-0.5**	-0.1	0.0	0.1
MinT	-0.7	0.1	0.2	0.2	-0.3	-0.1	0.0	-0.1	0.4	0.1	0.0	**-0.1**	-0.1	0.1	0.1	0.0
MinT-subset	-0.7	0.3	0.5	0.6	-0.2	0.1	0.2	0.1	**0.3**	0.2	0.1	0.1	-0.1	0.2	0.2	0.2
MinT-intuitive	-0.7	0.1	0.2	0.2	-0.3	-0.1	0.0	-0.1	0.4	0.1	0.0	**-0.1**	-0.1	0.1	0.1	0.0
MinT-lasso	**-1.3**	**-0.1**	0.2	0.3	**-0.6**	**-0.2**	0.0	-0.1	**0.3**	**0.0**	0.0	0.0	**-0.4**	**-0.1**	**0.0**	0.0
MinTs	-0.9	-0.1	0.1	0.1	-0.4	**-0.3**	**-0.2**	**-0.3**	0.1	-0.1	**-0.2**	**-0.1**	-0.3	**-0.2**	-0.1	**-0.1**
MinTs-subset	**-1.0**	0.1	0.2	0.4	-0.4	-0.2	-0.1	-0.1	**0.0**	0.0	0.0	0.0	**-0.4**	-0.1	0.0	0.1
MinTs-intuitive	-0.9	-0.1	0.1	0.1	-0.4	**-0.3**	**-0.2**	**-0.3**	0.1	-0.1	**-0.2**	**-0.1**	-0.3	**-0.2**	-0.1	**-0.1**
MinTs-lasso	**-1.4**	-0.1	0.2	0.4	**-0.6**	**-0.3**	-0.1	-0.1	**-0.1**	-0.1	-0.1	**-0.1**	**-0.6**	**-0.2**	0.0	0.0

Results: Base forecasts are generated by ETS

Ratio of each series being retained after subset selection in 500 instances.

	Top	A	B	AA	AB	BA	BB	Summary
OLS-subset	0.52	0.54	0.58	0.87	0.89	0.90	0.83	
OLS-intuitive	0.68	0.57	0.61	0.82	0.86	0.84	0.81	
OLS-lasso	0.62	0.52	0.53	1.00	1.00	1.00	1.00	
WLSs-subset	0.53	0.59	0.64	0.89	0.91	0.87	0.89	
WLSs-intuitive	0.65	0.58	0.61	0.86	0.92	0.87	0.88	
WLSs-lasso	0.60	0.58	0.59	1.00	1.00	1.00	1.00	
WLSv-subset	0.52	0.62	0.64	0.88	0.89	0.87	0.89	
WLSv-intuitive	0.64	0.57	0.55	0.87	0.93	0.87	0.92	
WLSv-lasso	0.60	0.60	0.61	1.00	1.00	1.00	1.00	
MinT-subset	0.55	0.56	0.57	0.91	0.92	0.89	0.90	
MinT-intuitive	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
MinT-lasso	0.76	0.81	0.80	0.97	0.97	0.97	0.97	
MinTs-subset	0.47	0.46	0.52	0.91	0.92	0.91	0.90	
MinTs-intuitive	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
MinTs-lasso	0.63	0.64	0.67	1.00	1.00	1.00	1.00	

Results: Scenario I - AA

Out-of-sample forecast performance (average RMSE).

	Top				Middle				Bottom				Average			
Method	h=1	1-4	1-8	1-16	h=1	1-4	1-8	1-16	h=1	1-4	1-8	1-16	h=1	1-4	1-8	1-16
Base	9.6	10.7	12.6	15.6	6.3	7.3	8.6	10.8	6.4	7.5	8.3	9.8	6.8	7.9	9.0	10.9
BU	57.8	68.5	53.7	38.9	58.2	61.8	48.1	34.4	0.0	0.0	0.0	0.0	27.0	29.6	23.8	17.7
OLS	0.6	2.2	1.8	1.4	7.1	6.4	4.6	3.1	-7.6	-8.6	-8.2	-7.3	-2.1	-2.5	-2.7	-2.6
OLS-subset	0.6	**1.8**	**1.5**	**1.3**	7.2	**5.2**	**3.8**	**2.6**	**-8.3**	**-12.9**	**-11.6**	**-9.9**	**-2.4**	**-5.2**	**-4.8**	**-4.1**
OLS-intuitive	0.8	2.6	2.1	1.8	7.5	**6.1**	**4.4**	**3.0**	**-9.0**	**-12.8**	**-11.6**	**-9.9**	**-2.7**	**-4.8**	**-4.5**	**-3.8**
OLS-lasso	0.6	2.2	1.8	1.6	7.4	6.7	4.8	3.2	-7.6	-8.5	-8.1	-7.2	-2.0	-2.4	-2.6	-2.5
WLSs	7.3	10.6	8.1	5.9	15.6	16.0	11.8	8.0	-6.9	-7.8	-7.4	-6.4	1.9	2.0	1.0	0.2
WLSs-subset	**5.0**	**5.7**	**4.6**	**3.6**	**12.3**	**10.0**	**7.5**	**5.2**	**-7.6**	**-10.5**	**-9.6**	**-8.2**	**0.2**	**-2.0**	**-2.1**	**-2.0**
WLSs-intuitive	**7.1**	**9.2**	**7.1**	**5.2**	16.5	**15.5**	**11.5**	**7.9**	-6.8	**-9.2**	**-8.4**	**-7.3**	2.1	**0.9**	**0.1**	**-0.4**
WLSs-lasso	7.3	**10.3**	**8.0**	5.9	15.7	16.1	11.8	8.1	**-7.0**	-7.8	-7.3	-6.4	1.9	2.0	1.0	0.2
WLSv	1.0	2.9	2.3	1.9	4.5	4.3	3.2	2.1	-25.8	-26.4	-22.7	-18.3	-12.4	-12.6	-10.7	-8.4
WLSv-subset	**-1.0**	**0.3**	**0.4**	**0.5**	**0.6**	**0.6**	**0.5**	**0.3**	**-32.3**	**-32.2**	**-27.3**	**-21.7**	**-17.3**	**-17.3**	**-14.2**	**-10.9**
WLSv-intuitive	**-0.5**	**0.2**	**0.3**	**0.5**	**0.9**	**0.7**	**0.5**	**0.3**	**-32.3**	**-32.3**	**-27.4**	**-21.7**	**-17.1**	**-17.3**	**-14.2**	**-10.9**
WLSv-lasso	**0.4**	**1.5**	**1.5**	**1.4**	**3.0**	**2.5**	**2.0**	**1.3**	**-28.5**	**-29.2**	**-24.9**	**-19.9**	**-14.4**	**-14.9**	**-12.3**	**-9.5**
MinT	-0.4	0.7	0.9	0.6	0.7	0.7	0.6	0.3	-32.9	-33.4	-28.3	-22.5	-17.5	-17.8	-14.6	-11.3
MinT-subset	**-0.6**	0.7	**0.8**	0.7	**0.6**	0.8	0.6	0.3	**-33.0**	-33.1	-28.0	-22.3	**-17.6**	-17.6	-14.5	-11.2
MinT-intuitive	-0.4	0.7	0.9	0.6	0.7	0.7	0.6	0.3	-32.9	-33.4	-28.3	-22.5	-17.5	-17.8	-14.6	-11.3
MinT-lasso	**-0.7**	**0.3**	**0.6**	**0.4**	**0.3**	**0.4**	**0.4**	**0.1**	**-33.2**	**-33.7**	**-28.5**	**-22.6**	**-17.8**	**-18.1**	**-14.8**	**-11.4**
MinTs	-0.9	0.6	0.7	0.5	0.6	0.6	0.5	0.2	-32.9	-33.5	-28.3	-22.5	-17.6	-17.9	-14.6	-11.3
MinTs-subset	-0.7	0.9	1.1	1.0	0.7	0.8	0.7	0.4	**-33.0**	-33.1	-27.9	-22.2	-17.6	-17.5	-14.3	-11.0
MinTs-intuitive	-0.9	0.6	0.7	0.5	0.6	0.6	0.5	0.2	-32.9	-33.5	-28.3	-22.5	-17.6	-17.9	-14.6	-11.3
MinTs-lasso	-0.9	**0.4**	**0.6**	0.5	0.6	**0.4**	**0.4**	**0.1**	**-33.2**	**-33.6**	**-28.4**	**-22.6**	**-17.7**	**-18.0**	**-14.8**	**-11.4**

Results: Scenario I - AA

Ratio of each series being retained after subset selection in 500 instances.

	Top	A	B	AA	AB	BA	BB	Summary
OLS-subset	0.52	0.79	0.57	0.79	1	0.91	0.85	
OLS-intuitive	0.80	0.90	0.81	0.80	1	0.85	0.86	
OLS-lasso	0.90	1.00	0.68	1.00	1	1.00	1.00	
WLSs-subset	0.85	0.91	0.86	0.90	1	0.97	0.97	
WLSs-intuitive	0.92	0.95	0.67	0.92	1	0.92	0.95	
WLSs-lasso	0.72	1.00	0.72	1.00	1	1.00	1.00	
WLSv-subset	0.50	0.62	0.42	0.19	1	0.81	0.87	
WLSv-intuitive	0.59	0.55	0.49	0.17	1	0.76	0.86	
WLSv-lasso	0.40	1.00	0.41	0.77	1	1.00	1.00	
MinT-subset	0.66	0.90	0.61	0.72	1	0.91	0.93	
MinT-intuitive	1.00	1.00	1.00	1.00	1	1.00	1.00	
MinT-lasso	0.80	0.96	0.84	0.72	1	0.98	0.97	
MinTs-subset	0.57	0.88	0.52	0.67	1	0.89	0.92	
MinTs-intuitive	1.00	1.00	1.00	1.00	1	1.00	1.00	
MinTs-lasso	0.68	1.00	0.66	0.74	1	1.00	1.00	

Results: Scenario II - A

Out-of-sample forecast performance (average RMSE).

	Top				Middle				Bottom				Average			
Method	h=1	1-4	1-8	1-16	h=1	1-4	1-8	1-16	h=1	1-4	1-8	1-16	h=1	1-4	1-8	1-16
Base	9.6	10.7	12.6	15.6	12.1	14.4	15.3	17.0	4.2	4.9	5.9	7.5	7.2	8.5	9.6	11.4
BU	**-1.0**	0.4	0.6	0.7	**-47.7**	**-49.6**	**-43.6**	**-36.2**	**0.0**	**0.0**	**0.0**	**0.0**	**-23.0**	**-24.0**	**-19.8**	**-15.3**
OLS	8.5	13.9	10.4	7.6	-28.2	-29.4	-26.7	-23.1	22.9	23.9	17.0	11.3	-4.2	-3.8	-4.2	-4.1
OLS-subset	**-0.5**	**0.5**	**0.6**	**0.7**	**-46.3**	**-49.0**	**-43.2**	**-35.9**	**2.2**	**1.0**	**0.7**	**0.5**	**-21.5**	**-23.4**	**-19.4**	**-15.0**
OLS-intuitive	**-0.5**	**0.5**	**0.6**	**0.6**	**-46.5**	**-49.0**	**-43.2**	**-36.0**	**2.2**	**1.2**	**0.7**	**0.5**	**-21.6**	**-23.4**	**-19.4**	**-15.0**
OLS-lasso	**-0.2**	**1.5**	**1.4**	**1.3**	**-46.9**	**-48.9**	**-43.1**	**-35.8**	**0.9**	**0.8**	**0.5**	**0.3**	**-22.1**	**-23.3**	**-19.3**	**-14.9**
WLSs	12.1	18.6	14.0	10.2	-34.4	-35.1	-31.7	-26.9	15.6	17.0	12.0	8.0	-9.0	-8.0	-7.6	-6.5
WLSs-subset	**-0.1**	**1.2**	**1.1**	**1.1**	**-46.7**	**-48.8**	**-43.1**	**-35.8**	**1.5**	**1.1**	**0.8**	**0.6**	**-21.8**	**-23.2**	**-19.2**	**-14.8**
WLSs-intuitive	**0.0**	**1.2**	**1.0**	**0.9**	**-46.5**	**-48.8**	**-43.1**	**-35.9**	**1.7**	**1.3**	**0.9**	**0.6**	**-21.6**	**-23.1**	**-19.2**	**-14.9**
WLSs-lasso	**-0.1**	**1.5**	**1.5**	**1.3**	**-46.7**	**-48.9**	**-43.1**	**-35.8**	**0.9**	**0.8**	**0.5**	**0.3**	**-22.0**	**-23.2**	**-19.3**	**-14.9**
WLSv	-0.8	2.3	1.8	1.6	-46.3	-47.9	-42.3	-35.2	1.6	1.9	1.2	0.8	-21.7	-22.2	-18.6	-14.4
WLSv-subset	-0.7	**1.3**	**1.4**	**1.4**	**-46.9**	**-48.7**	**-42.9**	**-35.6**	**1.0**	**1.0**	**0.8**	**0.6**	**-22.2**	**-23.1**	**-19.1**	**-14.7**
WLSv-intuitive	-0.4	**1.5**	**1.4**	**1.2**	**-46.9**	**-48.6**	**-42.8**	**-35.6**	**0.9**	**1.2**	**0.9**	**0.7**	**-22.2**	**-23.0**	**-19.0**	**-14.7**
WLSv-lasso	-0.6	**1.3**	**1.3**	**1.3**	**-47.2**	**-48.9**	**-43.0**	**-35.7**	**0.6**	**0.8**	**0.5**	**0.4**	**-22.4**	**-23.3**	**-19.2**	**-14.8**
MinT	0.2	0.5	0.6	0.5	-47.5	-49.4	-43.5	-36.1	1.1	0.5	0.3	0.1	-22.3	-23.7	-19.6	**-15.3**
MinT-subset	**-0.1**	0.8	0.9	0.9	-46.9	-49.1	-43.3	-36.0	1.7	0.9	0.5	0.3	-21.9	-23.4	-19.4	-15.1
MinT-intuitive	0.2	0.5	0.6	0.5	-47.5	-49.4	-43.5	-36.1	1.1	0.5	0.3	0.1	-22.3	-23.7	-19.6	**-15.3**
MinT-lasso	**-0.3**	**0.3**	0.6	0.5	**-47.6**	-49.4	-43.5	-36.1	**0.8**	**0.3**	**0.2**	0.1	**-22.5**	**-23.9**	**-19.7**	**-15.3**
MinTs	-0.3	0.3	**0.4**	**0.4**	-47.6	-49.5	**-43.6**	**-36.2**	0.7	0.2	0.1	**0.0**	-22.6	-23.9	**-19.8**	**-15.3**
MinTs-subset	**-0.8**	0.5	0.8	0.8	-47.2	-49.2	-43.4	-36.0	1.0	0.7	0.4	0.3	-22.3	-23.6	-19.5	-15.1
MinTs-intuitive	-0.3	0.3	**0.4**	**0.4**	-47.6	-49.5	**-43.6**	**-36.2**	0.7	0.2	0.1	**0.0**	-22.6	-23.9	**-19.8**	**-15.3**
MinTs-lasso	**-0.9**	**0.2**	0.5	0.5	**-47.7**	-49.5	**-43.6**	**-36.2**	**0.5**	0.2	0.1	0.1	**-22.8**	**-24.0**	**-19.8**	**-15.3**

Results: Scenario II - A

Ratio of each series being retained after subset selection in 500 instances.

	Top	A	B	AA	AB	BA	BB	Summary
OLS-subset	0.55	0.04	0.41	0.74	0.78	0.79	0.83	
OLS-intuitive	0.61	0.04	0.52	0.75	0.69	0.69	0.83	
OLS-lasso	0.04	0.35	0.02	1.00	1.00	1.00	1.00	
WLSs-subset	0.45	0.06	0.36	0.81	0.84	0.81	0.87	
WLSs-intuitive	0.61	0.06	0.48	0.75	0.71	0.73	0.84	
WLSs-lasso	0.02	0.33	0.02	1.00	1.00	1.00	1.00	
WLSv-subset	0.54	0.29	0.46	0.91	0.94	0.86	0.89	
WLSv-intuitive	0.59	0.32	0.53	0.82	0.86	0.77	0.86	
WLSv-lasso	0.27	0.42	0.26	1.00	1.00	1.00	1.00	
MinT-subset	0.69	0.64	0.66	0.95	0.96	0.90	0.90	
MinT-intuitive	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
MinT-lasso	0.82	0.74	0.83	1.00	0.99	0.97	0.97	
MinTs-subset	0.62	0.63	0.58	0.95	0.96	0.90	0.86	
MinTs-intuitive	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
MinTs-lasso	0.68	0.75	0.68	1.00	1.00	1.00	1.00	

Results: Scenario III - Total

Out-of-sample forecast performance (average RMSE).

	Top				Middle				Bottom				Average			
Method	h=1	1-4	1-8	1-16	h=1	1-4	1-8	1-16	h=1	1-4	1-8	1-16	h=1	1-4	1-8	1-16
Base	25.0	30.3	30.9	32.3	6.3	7.3	8.6	10.8	4.2	4.9	5.9	7.5	7.8	9.2	10.3	12.0
BU	-62.0	**-64.4**	**-59.0**	-51.5	**-0.3**	**0.0**	**0.1**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**-28.5**	**-30.2**	**-25.3**	**-19.8**
OLS	-34.8	-35.5	-33.5	-30.1	45.3	50.6	37.7	25.1	27.7	29.9	21.2	13.7	3.1	3.8	1.6	-0.2
OLS-subset	**-35.3**	**-41.9**	**-39.2**	**-35.0**	**43.9**	**39.5**	**29.5**	**19.6**	**27.1**	**23.6**	**16.8**	**10.9**	**2.4**	**-3.5**	**-4.2**	**-4.5**
OLS-intuitive	**-41.2**	**-49.2**	**-45.5**	**-40.0**	**35.1**	**26.8**	**20.3**	**13.7**	**21.9**	**15.9**	**11.5**	**7.6**	**-4.0**	**-12.2**	**-10.9**	**-9.1**
OLS-lasso	**-61.8**	**-63.6**	**-58.1**	**-50.9**	**0.4**	**1.3**	**1.3**	**0.7**	**0.3**	**0.8**	**0.6**	**0.4**	**-28.2**	**-29.3**	**-24.5**	**-19.2**
WLSs	-50.9	-52.4	-48.7	-43.3	17.6	20.0	14.5	9.3	9.6	11.3	7.7	4.9	-16.3	-16.7	-14.9	-12.5
WLSs-subset	**-61.8**	**-63.6**	**-58.1**	**-50.7**	**0.3**	**1.4**	**1.4**	**0.9**	**0.3**	**0.9**	**0.7**	**0.6**	**-28.2**	**-29.3**	**-24.4**	**-19.0**
WLSs-intuitive	**-61.8**	**-63.8**	**-58.3**	**-50.9**	**0.0**	**1.0**	**1.0**	**0.7**	**0.3**	**0.7**	**0.6**	**0.5**	**-28.3**	**-29.5**	**-24.6**	**-19.2**
WLSs-lasso	**-61.7**	**-63.5**	**-58.0**	**-50.7**	**0.5**	**1.5**	**1.4**	**0.9**	**0.3**	**0.9**	**0.7**	**0.5**	**-28.1**	**-29.2**	**-24.4**	**-19.1**
WLSv	-61.1	-63.4	-58.1	-50.8	1.0	1.7	1.3	0.8	0.7	1.0	0.6	0.4	-27.6	-29.1	-24.5	-19.2
WLSv-subset	**-61.9**	**-63.6**	**-58.2**	**-50.9**	**0.2**	**1.3**	**1.2**	0.8	**0.1**	**0.8**	0.6	0.5	**-28.3**	**-29.3**	-24.5	-19.2
WLSv-intuitive	**-61.8**	**-63.8**	**-58.3**	**-51.0**	**0.0**	**1.1**	**1.1**	**0.6**	**0.1**	**0.6**	**0.5**	0.4	**-28.4**	**-29.5**	**-24.7**	**-19.3**
WLSv-lasso	**-61.8**	**-63.9**	**-58.4**	**-51.1**	**0.2**	**0.9**	**0.9**	**0.5**	**0.1**	**0.5**	**0.4**	**0.3**	**-28.3**	**-29.6**	**-24.8**	**-19.4**
MinT	-62.1	-64.3	-58.9	**-51.6**	-0.2	0.6	0.5	0.2	0.8	0.5	0.3	0.1	-28.3	-29.9	-25.1	**-19.8**
MinT-subset	-61.8	-63.7	-58.2	-50.9	0.4	1.2	1.3	0.8	0.8	1.0	0.7	0.5	-28.0	-29.3	-24.5	-19.2
MinT-intuitive	-62.1	-64.3	-58.9	**-51.6**	-0.2	0.6	0.5	0.2	0.8	0.5	0.3	0.1	-28.3	-29.9	-25.1	**-19.8**
MinT-lasso	-62.1	**-64.4**	-58.9	-51.5	**-0.3**	**0.3**	**0.4**	**0.1**	**0.6**	**0.3**	**0.1**	0.1	**-28.4**	**-30.1**	**-25.2**	**-19.8**
MinTs	**-62.2**	**-64.4**	**-59.0**	**-51.6**	**-0.3**	0.3	0.4	0.1	0.4	0.3	0.1	**0.0**	**-28.5**	-30.1	-25.2	**-19.8**
MinTs-subset	-62.0	-63.8	-58.4	-51.1	0.4	1.1	1.2	0.7	0.5	0.9	0.7	0.5	-28.2	-29.5	-24.6	-19.3
MinTs-intuitive	**-62.2**	**-64.4**	**-59.0**	**-51.6**	**-0.3**	0.3	0.4	0.1	0.4	0.3	0.1	**0.0**	**-28.5**	-30.1	-25.2	**-19.8**
MinTs-lasso	**-62.2**	**-64.4**	-58.9	-51.5	-0.2	0.3	0.4	0.1	**0.2**	**0.2**	0.1	**0.0**	**-28.5**	-30.1	-25.2	**-19.8**

Results: Scenario III - Total

Ratio of each series being retained after subset selection in 500 instances.

	Top	A	B	AA	AB	BA	BB	Summary
OLS-subset	0.75	0.45	0.44	0.82	0.79	0.83	0.80	
OLS-intuitive	0.47	0.70	0.69	0.86	0.92	0.90	0.89	
OLS-lasso	0.38	0.01	0.01	1.00	1.00	1.00	1.00	
WLSs-subset	0.08	0.42	0.41	0.87	0.85	0.84	0.89	
WLSs-intuitive	0.06	0.55	0.50	0.66	0.87	0.69	0.88	
WLSs-lasso	0.35	0.03	0.03	1.00	1.00	1.00	1.00	
WLSv-subset	0.31	0.67	0.65	0.88	0.90	0.91	0.90	
WLSv-intuitive	0.34	0.63	0.60	0.80	0.89	0.84	0.87	
WLSv-lasso	0.45	0.35	0.36	1.00	1.00	1.00	1.00	
MinT-subset	0.69	0.78	0.80	0.91	0.91	0.91	0.91	
MinT-intuitive	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
MinT-lasso	0.75	0.89	0.86	0.97	0.97	0.97	0.97	
MinTs-subset	0.67	0.74	0.76	0.90	0.89	0.88	0.91	
MinTs-intuitive	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
MinTs-lasso	0.77	0.72	0.73	1.00	1.00	1.00	1.00	

Outline

Data description

Australian domestic tourism

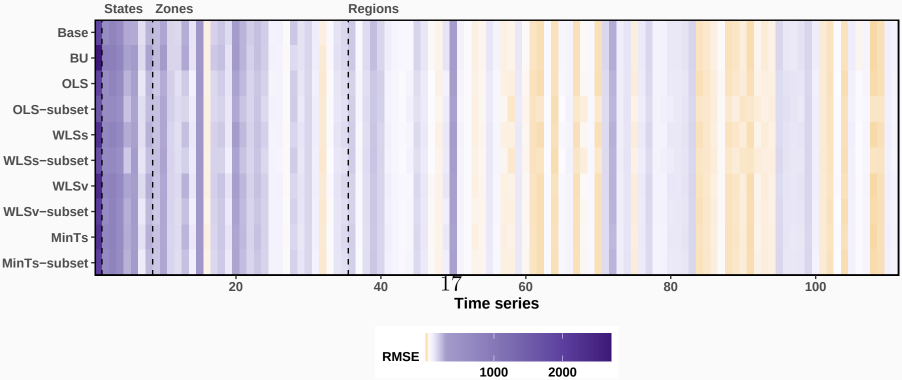
- Monthly series from 1998 Jan to 2017 Dec (20 years).
- Hierarchy structure:
  - Total/State/Zone/Region, 4 levels
  - $n_b = 76$ series at the bottom-level, $n = 111$ series in total.
- Training set: 1998 Jan-2016 Dec.
- Test set: 2017 Jan-2017 Dec.

Out-of-sample forecast performance (average RMSE)

	Top				State				Zone				Region				Average			
Method	h=1	1-4	1-8	1-12	h=1	1-4	1-8	1-12	h=1	1-4	1-8	1-12	h=1	1-4	1-8	1-12	h=1	1-4	1-8	1-12
Base	1158.2	716.6	1279.5	1907.6	452.7	323.3	349.9	424.8	165.5	163.6	160.7	179.7	100.8	89.4	88.2	94.1	148.3	127.9	133.1	152.1
BU	89.1	132.8	53.4	42.0	-4.6	10.3	17.0	19.7	1.1	-2.4	0.4	1.0	0.0	0.0	0.0	0.0	5.7	7.6	7.6	8.5
OLS	**-4.7**	**-0.4**	0.5	1.4	-3.0	-3.9	-1.6	-1.5	-2.1	-4.2	-5.6	-7.5	1.0	-0.4	-1.9	-3.2	-1.0	-2.1	-2.7	-3.6
OLS-subset	**-4.7**	8.0	**-1.4**	**-14.1**	-3.0	5.5	0.3	**-7.9**	-2.1	-1.5	-3.7	**-8.7**	1.0	1.7	-0.1	-2.3	-1.0	1.7	-1.2	**-6.5**
OLS-intuitive	**-4.7**	**-0.4**	0.5	1.4	-3.0	-3.9	-1.6	-1.5	-2.1	-4.2	-5.6	-7.5	1.0	-0.4	-1.9	-3.2	-1.0	-2.1	-2.7	-3.6
OLS-lasso	**-4.7**	**-0.4**	0.5	1.4	-3.0	-3.9	-1.6	-1.5	-2.1	-4.2	-5.6	-7.5	1.0	-0.4	-1.9	-3.2	-1.0	-2.1	-2.7	-3.6
WLSs	25.1	55.2	20.8	19.1	-15.8	-5.0	3.5	6.2	-5.9	-5.4	-4.7	-5.0	-0.2	-0.8	-1.6	-2.2	-3.0	-0.1	0.3	0.9
WLSs-subset	25.1	**18.7**	**0.8**	**-7.8**	-15.8	-2.7	**-2.1**	**-6.2**	-5.9	-4.1	**-4.8**	**-8.5**	-0.2	0.3	-1.0	**-2.5**	-3.0	**-0.6**	**-2.1**	**-5.5**
WLSs-intuitive	25.1	55.2	20.8	19.1	-15.8	-5.0	3.5	6.2	-5.9	-5.4	-4.7	-5.0	-0.2	-0.8	-1.6	-2.2	-3.0	-0.1	0.3	0.9
WLSs-lasso	25.1	55.2	20.8	19.1	-15.8	-5.0	3.5	6.2	-5.9	-5.4	-4.7	-5.0	-0.2	-0.8	-1.6	-2.2	-3.0	-0.1	0.3	0.9
WLSv	38.2	76.2	29.6	25.6	-17.4	-3.1	7.0	9.9	-5.0	-4.3	-3.1	-3.2	-4.2	-1.6	-1.8	-2.1	-3.9	1.3	2.0	2.8
WLSv-subset	38.2	**34.5**	**10.7**	**8.5**	-17.4	**-8.8**	**-0.8**	**1.4**	-5.0	**-5.5**	**-5.3**	**-6.7**	-4.1	**-2.0**	**-2.6**	**-3.4**	-3.9	**-2.3**	**-2.0**	**-2.2**
WLSv-intuitive	38.2	76.2	29.6	25.6	-17.4	-3.1	7.0	9.9	-5.0	-4.3	-3.1	-3.2	-4.2	-1.6	-1.8	-2.1	-3.9	1.3	2.0	2.8
WLSv-lasso	38.2	76.2	29.6	25.6	-17.4	-3.1	7.0	9.9	-5.0	-4.3	-3.1	-3.2	-4.2	-1.6	-1.8	-2.1	-3.9	1.3	2.0	2.8
MinTs	20.6	53.6	21.6	19.0	**-22.2**	-7.2	3.5	6.3	**-12.1**	-6.6	-5.1	-5.3	**-5.3**	-2.6	-2.8	-3.1	**-8.6**	-1.8	-0.3	0.4
MinTs-subset	20.6	**20.0**	**6.4**	**5.6**	**-22.2**	**-11.3**	**-2.5**	**-0.1**	**-12.1**	**-7.5**	**-6.4**	**-7.8**	**-5.3**	**-2.9**	**-3.2**	**-3.9**	**-8.6**	**-4.5**	**-3.2**	**-3.3**
MinTs-intuitive	20.6	53.6	21.6	19.0	**-22.2**	-7.2	3.5	6.3	**-12.1**	-6.6	-5.1	-5.3	**-5.3**	-2.6	-2.8	-3.1	**-8.6**	-1.8	-0.3	0.4
MinTs-lasso	20.6	53.6	21.6	19.0	**-22.2**	-7.2	3.5	6.3	**-12.1**	-6.6	-5.1	-5.3	**-5.3**	-2.6	-2.8	-3.1	**-8.6**	-1.8	-0.3	0.4

Further analysis

	Number of time series retained					Optimal parameters	
	Top	State	Zone	Region	Total	λ_0	λ_2
None	1	7	27	76	111	0.00	0.00
OLS-subset	1	2	13	76	92	27.98	10.00
WLSs-subset	1	1	15	76	93	18.73	10.00
WLSv-subset	1	7	27	76	111	0.03	0.01
MinTs-subset	1	7	27	76	111	0.05	0.01



17

Outline

Conclusions

- Three methods to achieve subset selection in forecast reconciliation.
  - **Regularized best-subset selection**
  - Intuitive method
  - Group lasso method
- Regularized best-subset selection method performs well and generally seems to outperform existing methods.
  - Especially effective when dealing with model misspecification issues within the hierarchy.
  - Reduce differences arising from the choice of $\boldsymbol{W}$ estimators.
  - Perform particularly in the context of long-term forecast reconciliation.

References

- Athanasopoulos, G., Hyndman, R. J., Kourentzes, N., & Panagiotelis, A. (2023). Forecast reconciliation: A review.
- Ben Taieb, S., & Koo, B. (2019). Regularized regression for hierarchical forecasting without unbiasedness conditions. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 1337-1347).
- Hazimeh, H., & Mazumder, R. (2020). Fast best subset selection: Coordinate descent and local combinatorial optimization algorithms. Operations Research, 68(5), 1517-1537.
- Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G., & Shang, H. L. (2011). Optimal combination forecasts for hierarchical time series. Computational statistics & data analysis, 55(9), 2579-2589.
- Mazumder, R., Radchenko, P., & Dedieu, A. (2023). Subset selection with shrinkage: Sparse linear modeling when the SNR is low. Operations Research, 71(1), 129-147.
- 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.

THANK YOU

Find me at ...

xqnwang.rbind.io
@Xia0qianWang
@xqnwang
xiaoqian.wang@monash.edu