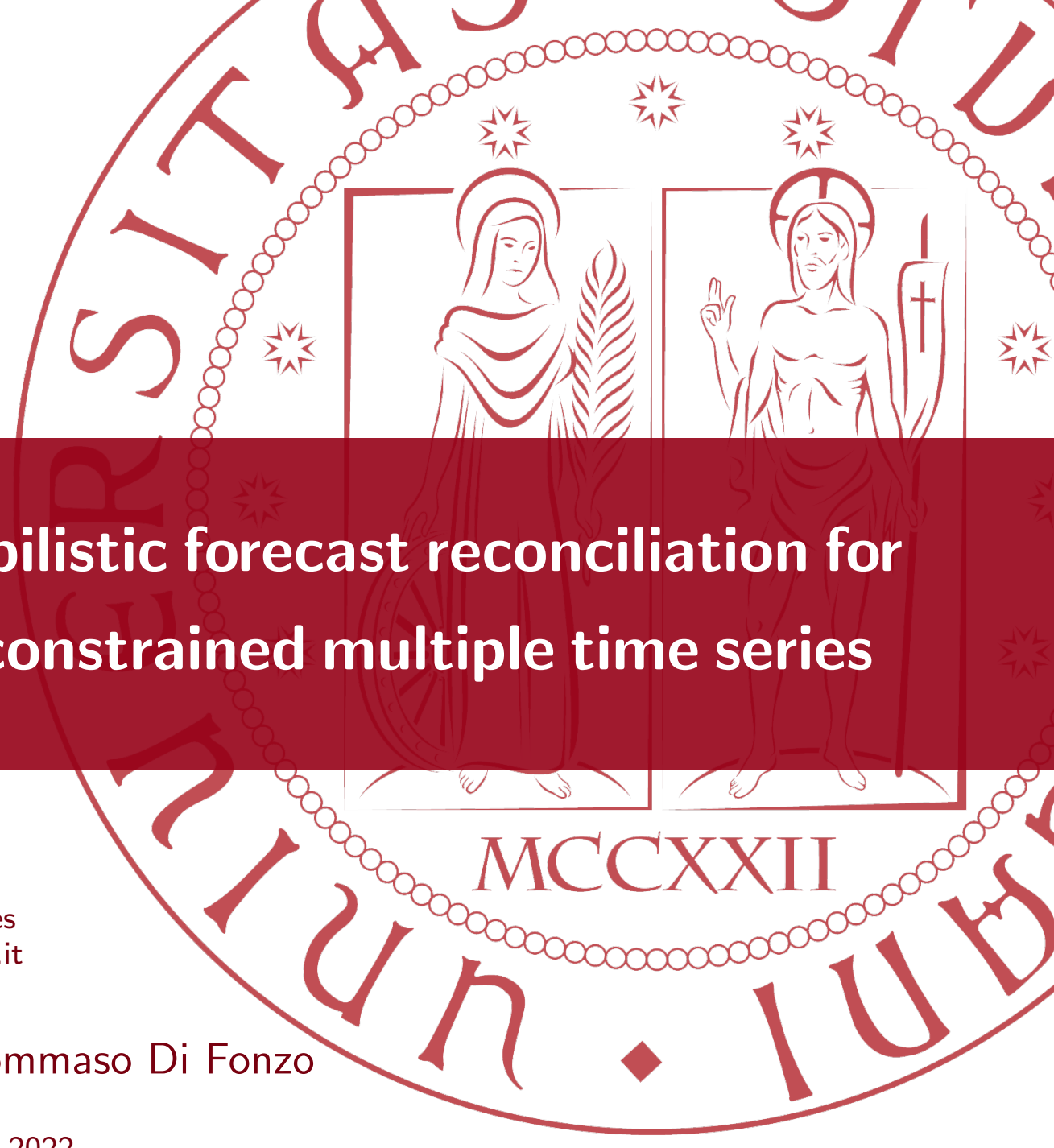
8th International conference on
Time Series and Forecasting

ITISE
2022



Point and probabilistic forecast reconciliation for general linearly constrained multiple time series

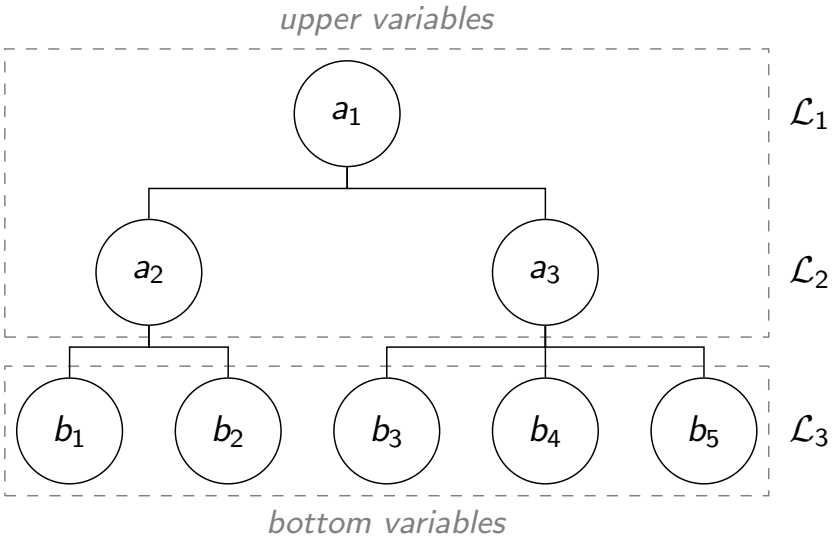
University of Padua (Italy)
Department of Statistical Sciences
@ daniele.girolimetto@phd.unipd.it
github.com/daniGiro

Daniele Girolimetto & Tommaso Di Fonzo

Gran Canaria (Spain) - June 29, 2022

Genuine hierarchical/grouped time series

Hyndman *et al.* (2011)



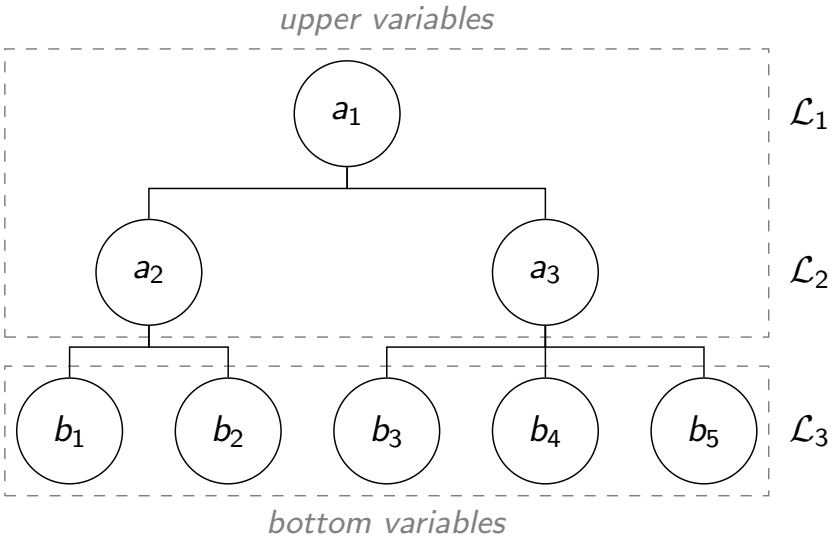
$$\underbrace{\begin{bmatrix} a_{1,t} \\ a_{2,t} \\ a_{3,t} \\ b_{1,t} \\ b_{2,t} \\ b_{3,t} \\ b_{4,t} \\ b_{5,t} \end{bmatrix}}_{\mathbf{y}_t} = \underbrace{\left[\begin{array}{ccccc} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]}_{\mathbf{S}} \underbrace{\begin{bmatrix} b_{1,t} \\ b_{2,t} \\ b_{3,t} \\ b_{4,t} \\ b_{5,t} \end{bmatrix}}_{\mathbf{b}_t}$$

Structural representation: $\mathbf{y}_t = \begin{bmatrix} \mathbf{a}_t \\ \mathbf{b}_t \end{bmatrix} = \mathbf{S}\mathbf{b}_t = \begin{bmatrix} \mathbf{C} \\ \mathbf{I}_{n_b} \end{bmatrix} \mathbf{b}_t, \qquad t = 1, \ldots, T$

- $\mathbf{S} \in \{0,1\}^{(n \times n_b)}$ summing matrix
- $\mathbf{C} \in \{0,1\}^{(n_a \times n_b)}$ aggregation matrix
- $\mathbf{a}_t \in \mathbb{R}^{(n_a \times 1)}$ 'upper' variables
- $\mathbf{b}_t \in \mathbb{R}^{(n_b \times 1)}$ 'bottom' variables

ITISE 2022

Genuine hierarchical/grouped time series



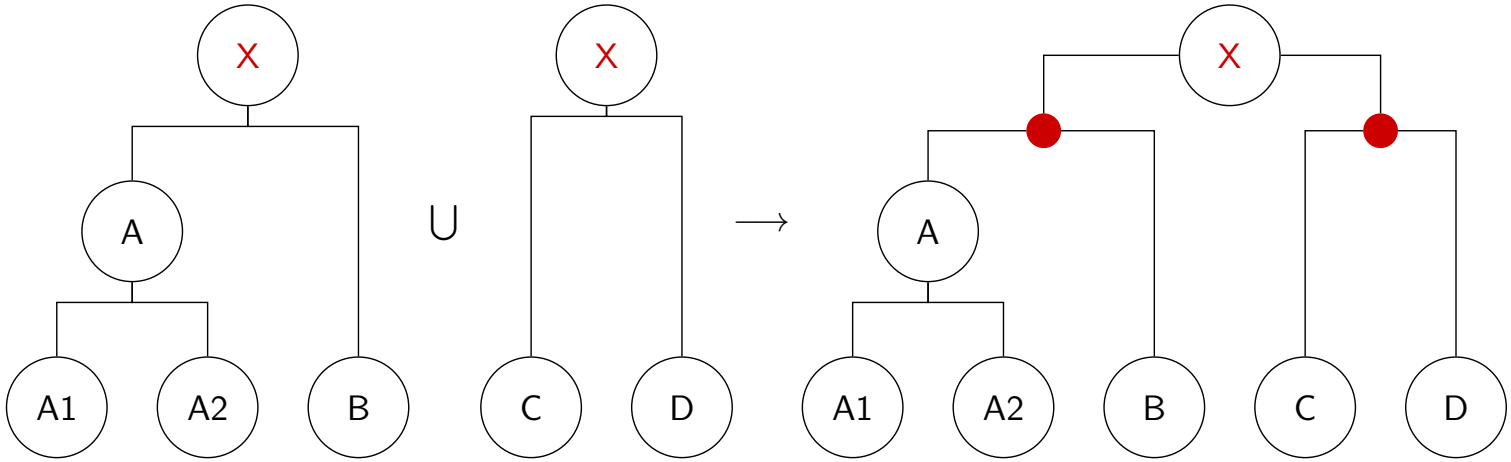
$$\underbrace{\left[\begin{array}{ccc|ccccc} 1 & 0 & 0 & -1 & -1 & -1 & -1 & -1 \\ 0 & 1 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & -1 & -1 \end{array}\right]}_{\mathbf{U}'} \underbrace{\left[\begin{array}{c} a_{1,t} \\ a_{2,t} \\ a_{3,t} \\ b_{1,t} \\ b_{2,t} \\ b_{3,t} \\ b_{4,t} \\ b_{5,t} \end{array}\right]}_{\mathbf{y}_t} = \underbrace{\left[\begin{array}{c} 0 \\ 0 \\ 0 \end{array}\right]}_{\mathbf{0}_{(3\times 1)}}$$

Zero-constrained representation: $\mathbf{U}'\mathbf{y}_t = \mathbf{0}_{(n_a \times 1)}, \qquad t = 1, ..., T$

- $\mathbf{U}' = [\mathbf{I}_{n_a} \quad -\mathbf{C}] \in \{-1, 0, 1\}^{(n_a \times n)}$ (**full rank**) zero constraints kernel matrix (Di Fonzo and Girolimetto, 2021, see also Wickramasuriya *et al.*, 2019)

ITISE 2022

General linearly constrained multiple time series



Zero-constrained representation: $\mathbf{\Gamma}'\mathbf{x}_t = \mathbf{0}_{(r\times 1)}, \qquad t = 1, ..., T$

- $\mathbf{\Gamma}' \in \mathbb{R}^{(r\times n)}$, matrix of known coefficients (not necessarily full rank)

$$\underbrace{\begin{bmatrix} 1 & -1 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 1 & -1 & -1 & 0 & 0 & 0 \end{bmatrix}}_{\mathbf{\Gamma}'} \underbrace{\begin{bmatrix} X_t \\ A_t \\ A1_t \\ A2_t \\ B_t \\ C_t \\ D_t \end{bmatrix}}_{\mathbf{x}_t} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{0}_{(3\times 1)}} \iff \begin{aligned} X &= A + B \\ X &= C + D \\ A &= A1 + A2 \end{aligned}$$

ITISE 2022

General linearly constrained multiple time series

From zero-constrained to structural-like representation

- Working on $\mathbf{\Gamma}'$, we derive $\overline{\mathbf{U}}' = \begin{bmatrix} \mathbf{I} & -\overline{\mathbf{C}} \end{bmatrix}$ and $\mathbf{y}_t = \mathbf{P}\mathbf{x}_t$
- Structural-like representation through $\overline{\mathbf{S}} = \begin{bmatrix} \overline{\mathbf{C}} \\ \mathbf{I} \end{bmatrix}$ permits to extend the classical point and probabilistic forecast reconciliation results
- For medium/large systems (many constraints and/or variables) manually operating may be time-consuming, challenging and prone to error

Our proposal

$$\mathbf{\Gamma}' \xrightarrow[\text{QR}]{\text{rref}} \overline{\mathbf{U}}' = \begin{bmatrix} \mathbf{I}_{n_v} & -\overline{\mathbf{C}} \end{bmatrix} \quad \Rightarrow \quad \mathbf{y}_t = \mathbf{P}\mathbf{x}_t = \begin{bmatrix} \mathbf{v}_t \\ \mathbf{f}_t \end{bmatrix} = \begin{bmatrix} \overline{\mathbf{C}} \\ \mathbf{I}_{n_f} \end{bmatrix} \mathbf{f}_t = \overline{\mathbf{S}}\mathbf{f}_t$$

- $\overline{\mathbf{U}}' \in \mathbb{R}^{(n_v \times n)}$ (full rank) zero constraints matrix
- $\overline{\mathbf{S}} \in \mathbb{R}^{(n \times n_f)}$ analogous of the summing matrix $\mathbf{S}$
- $\overline{\mathbf{C}} \in \mathbb{R}^{(n_v \times n_f)}$ linear combination matrix such that $\mathbf{v}_t = \overline{\mathbf{C}}\mathbf{f}_t$
- $\mathbf{v}_t \in \mathbb{R}^{(n_v \times 1)}$ 'basic' variables, and $\mathbf{f}_t \in \mathbb{R}^{(n_f \times 1)}$ 'free' variables

ITISE 2022

General linearly constrained multiple time series

From zero-constrained to structural-like representation

- Working on $\mathbf{\Gamma}'$, we derive $\overline{\mathbf{U}}' = \begin{bmatrix} \mathbf{I} & -\overline{\mathbf{C}} \end{bmatrix}$ and $\mathbf{y}_t = \mathbf{P}\mathbf{x}_t$
- Structural-like representation through $\overline{\mathbf{S}} = \begin{bmatrix} \overline{\mathbf{C}} \\ \mathbf{I} \end{bmatrix}$ permits to extend the classical point and probabilistic forecast reconciliation results
- For medium/large systems (many constraints and/or variables) manually operating may be time-consuming, challenging and prone to error

Our proposal

$$\mathbf{\Gamma}' \xrightarrow[\text{QR}]{\text{rref}} \overline{\mathbf{U}}' = \begin{bmatrix} \mathbf{I}_{n_v} & -\overline{\mathbf{C}} \end{bmatrix} \quad \Rightarrow \quad \mathbf{y}_t = \mathbf{P}\mathbf{x}_t = \begin{bmatrix} \mathbf{v}_t \\ \mathbf{f}_t \end{bmatrix} = \begin{bmatrix} \overline{\mathbf{C}} \\ \mathbf{I}_{n_f} \end{bmatrix} \mathbf{f}_t = \overline{\mathbf{S}}\mathbf{f}_t$$

- $\overline{\mathbf{U}}' \in \mathbb{R}^{(n_v \times n)}$ (full rank) zero constraints matrix
- $\overline{\mathbf{S}} \in \mathbb{R}^{(n \times n_f)}$ analogous of the summing matrix $\mathbf{S}$
- $\overline{\mathbf{C}} \in \mathbb{R}^{(n_v \times n_f)}$ linear combination matrix such that $\mathbf{v}_t = \overline{\mathbf{C}}\mathbf{f}_t$
- $\mathbf{v}_t \in \mathbb{R}^{(n_v \times 1)}$ 'basic' variables, and $\mathbf{f}_t \in \mathbb{R}^{(n_f \times 1)}$ 'free' variables

ITISE 2022

Point and probabilistic forecast reconciliation

Extension of Wickramasuriya *et al.* (2019) and Panagiotelis *et al.* (2020, 2021)

Target $\overline{\mathbf{U}}'\mathbf{y}_h = \mathbf{0}$ | Base forecasts $\overline{\mathbf{U}}'\widehat{\mathbf{y}}_h \neq \mathbf{0}$ | $\rightarrow$ | Reconciled forecasts $\overline{\mathbf{U}}'\widetilde{\mathbf{y}}_h = \mathbf{0}$

Two equivalent point forecast reconciliation formulae

Structural reconciliation approach

$$\widetilde{\mathbf{y}}_h = \underbrace{\overline{\mathbf{S}}\left(\overline{\mathbf{S}}'\mathbf{W}_h^{-1}\overline{\mathbf{S}}\right)^{-1}\overline{\mathbf{S}}'\mathbf{W}_h^{-1}}_{\psi}\widehat{\mathbf{y}}_h$$

Projection reconciliation approach

$$\widetilde{\mathbf{y}}_h = \underbrace{\left[\mathbf{I} - \mathbf{W}_h\overline{\mathbf{U}}\left(\overline{\mathbf{U}}'\mathbf{W}_h'\overline{\mathbf{U}}\right)^{-1}\overline{\mathbf{U}}'\right]}_{\psi}\widehat{\mathbf{y}}_h$$

Probabilistic forecast reconciliation

(Reconciled samples thm, Panagiotelis *et al.*, 2020)

$$\left(\widehat{\mathbf{y}}^{[1]},\dots,\widehat{\mathbf{y}}^{[L]}\right) \xrightarrow{\widetilde{\mathbf{y}}^{[\ell]} = \psi\left(\widehat{\mathbf{y}}^{[\ell]}\right)} \left(\widetilde{\mathbf{y}}^{[1]},\dots,\widetilde{\mathbf{y}}^{[L]}\right)$$

sample drawn from an incoherent probability measure $\widehat{\nu}$

sample drawn from the reconciled probability measure $\widetilde{\nu}$

Obtaining $\left(\widehat{\mathbf{y}}^{[1]},\dots,\widehat{\mathbf{y}}^{[L]}\right)$: **Joint bootstrap** (Gamakumara *et al.*, 2018)

ITISE 2022

Australian *GDP* from Income and Expenditure sides

95 (*GDP* + 15 Income + 79 Expenditure) quarterly ts at current prices, 1984:Q4-2018:Q1

Forecasting experiment (Athanasopoulos *et al.*, 2020):

- ARIMA base forecasts (expanding windows, 94 forecast origins)
- Forecast horizons $h = 1, \ldots, 4$

Forecast reconciliation approaches:

- **ols**: $\mathbf{W}_h = \sigma^2 \mathsf{I}_n$ (Hyndman *et al.*, 2011)
- **wls**: $\mathbf{W}_h = \widehat{\mathbf{W}}_D = \operatorname{diag} \widehat{\mathbf{W}}_1$ (Hyndman *et al.*, 2016)
- **shr**: $\mathbf{W}_h = \widehat{\mathbf{W}}_{shr} = \lambda \widehat{\mathbf{W}}_D + (1-\lambda)\widehat{\mathbf{W}}_1$ (Wickramasuriya *et al.*, 2019)

$\widehat{\mathbf{W}}_1$ is the $(n \times n)$ covariance matrix of the in-sample one-step-ahead base forecasts errors

The base forecasts are reconciled using the `htsrec` function of the R-package **FoReco** (v.0.2.4, Girolimetto and Di Fonzo, 2022)

ITISE 2022

Australian *GDP* from Income and Expenditure sides

95 (*GDP* + 15 Income + 79 Expenditure) quarterly ts at current prices, 1984:Q4-2018:Q1

Forecasting experiment (Athanasopoulos *et al.*, 2020):

- ARIMA base forecasts (expanding windows, 94 forecast origins)
- Forecast horizons $h = 1, ..., 4$

Forecast reconciliation approaches:

- **ols**: $\mathbf{W}_h = \sigma^2 \mathsf{I}_n$ (Hyndman *et al.*, 2011)
- **wls**: $\mathbf{W}_h = \widehat{\mathbf{W}}_D = \operatorname{diag} \widehat{\mathbf{W}}_1$ (Hyndman *et al.*, 2016)
- **shr**: $\mathbf{W}_h = \widehat{\mathbf{W}}_{shr} = \lambda \widehat{\mathbf{W}}_D + (1 - \lambda)\widehat{\mathbf{W}}_1$ (Wickramasuriya *et al.*, 2019)

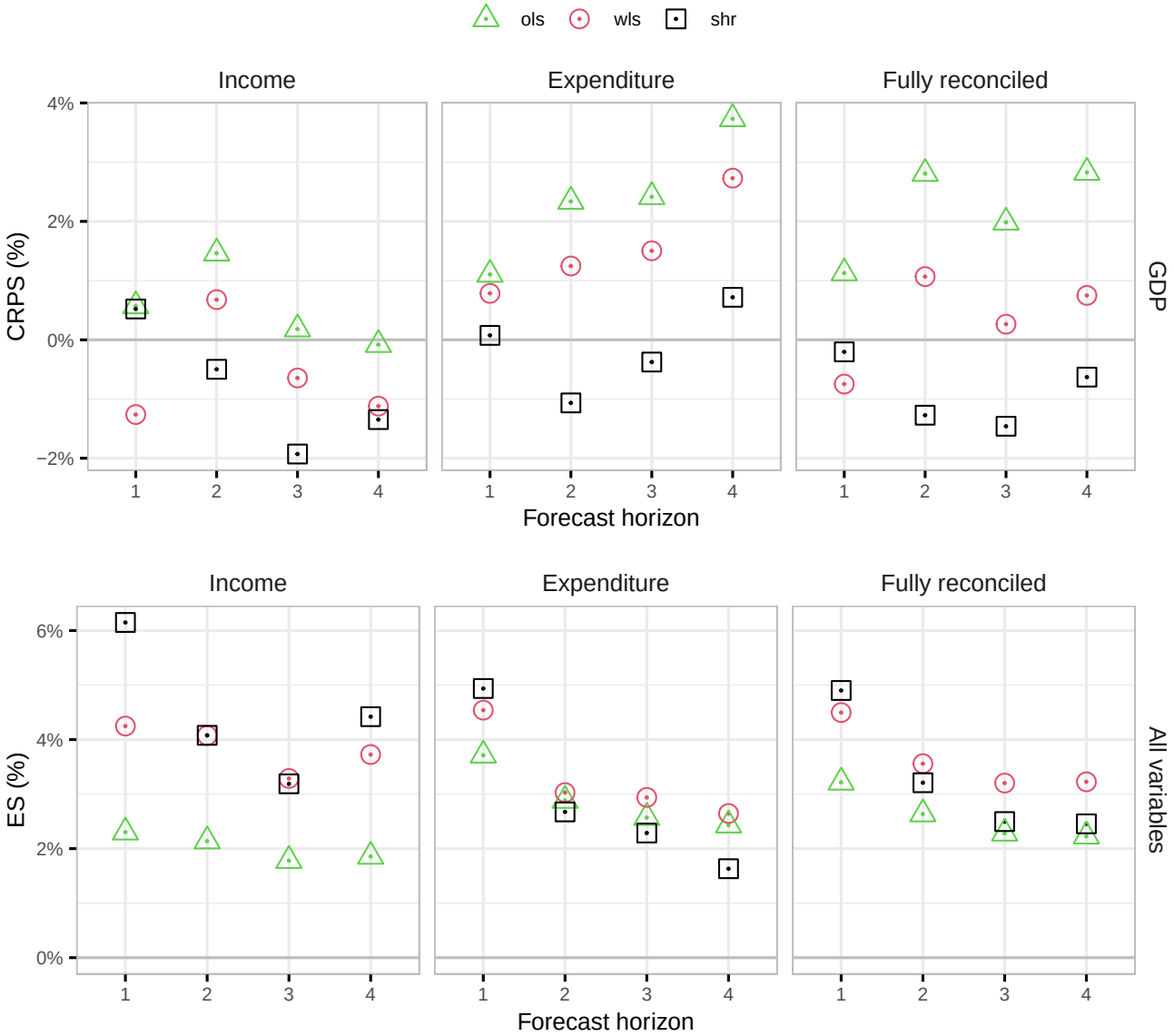
$\widehat{\mathbf{W}}_1$ is the $(n \times n)$ covariance matrix of the in-sample one-step-ahead base forecasts errors

The base forecasts are reconciled using the `htsrec` function of the R-package **FoReco** (v.0.2.4, Girolimetto and Di Fonzo, 2022)

ITISE 2022

Results: skill scores

Cumulative Rank Probability Score and Energy Score (`scoringRules`, Jordan *et al.*, 2019)



GDP

- ols always outperforms both wls and shr

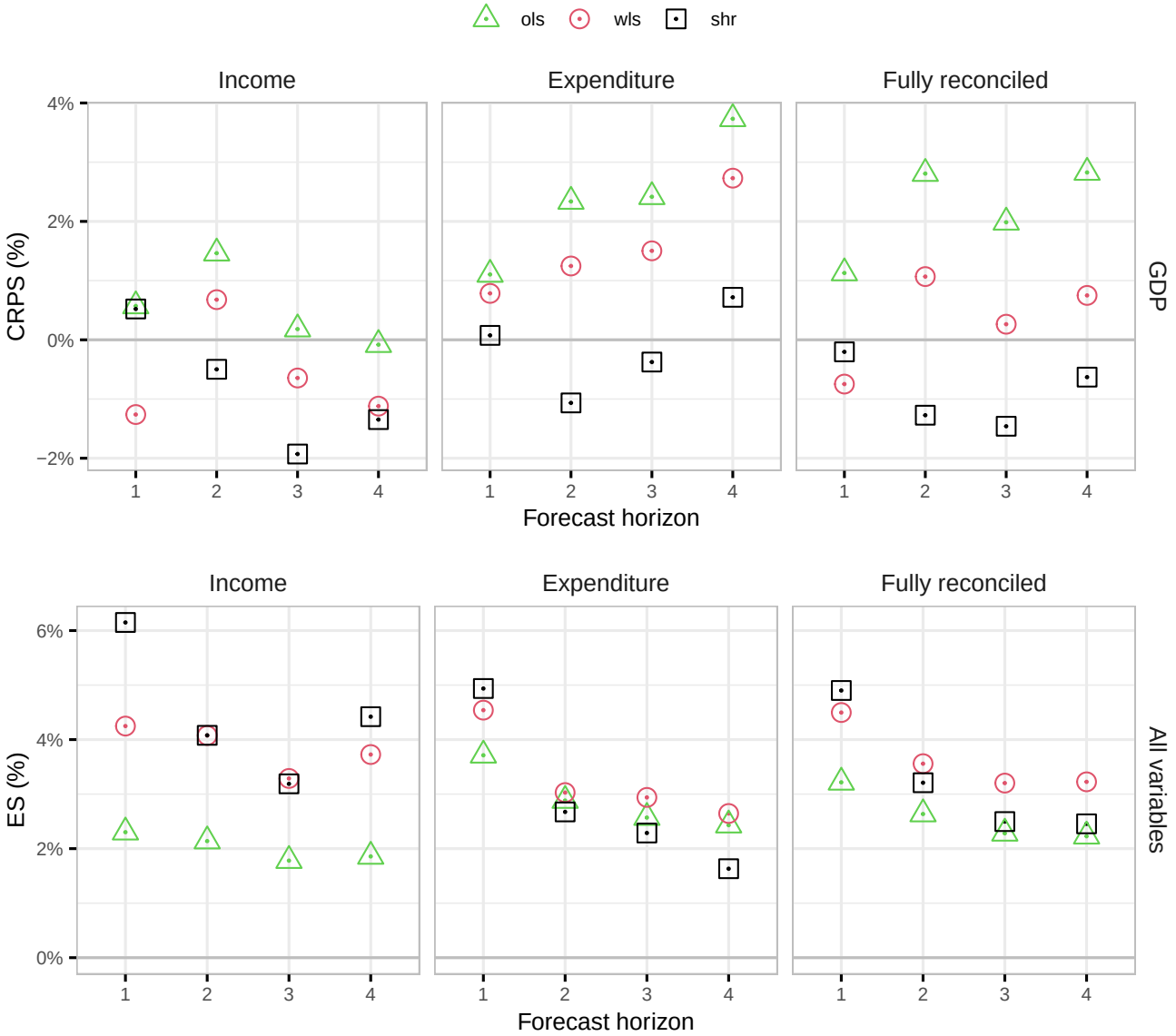
All variables

- reconciliation improves on the base forecasts' accuracy
- shr and wls are almost always the best

ITISE 2022

Results: skill scores

Cumulative Rank Probability Score and Energy Score (`scoringRules`, Jordan *et al.*, 2019)



GDP

- ols always outperforms both wls and shr

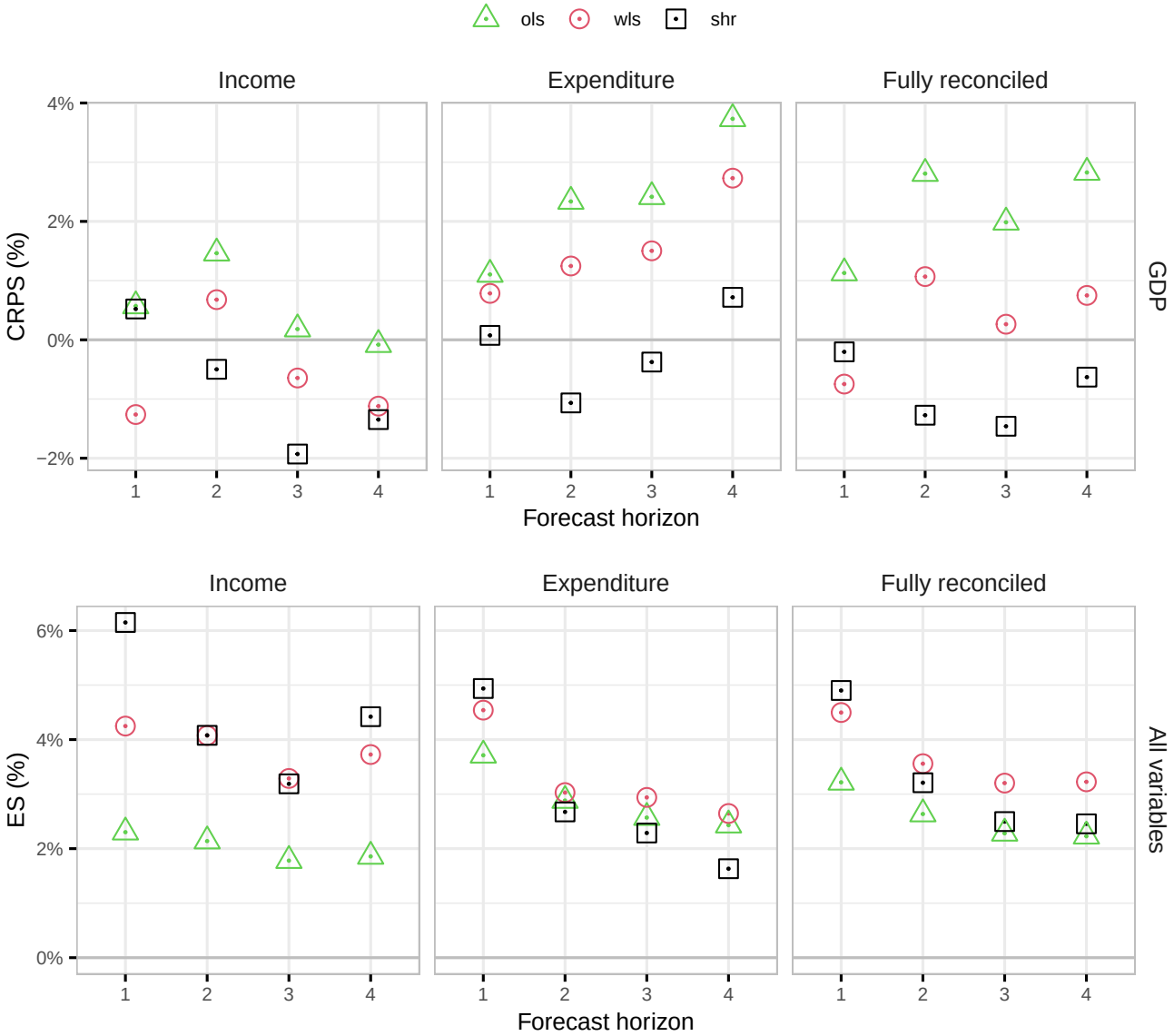
All variables

- reconciliation improves on the base forecasts' accuracy
- shr and wls are almost always the best

ITISE 2022

Results: skill scores

Cumulative Rank Probability Score and Energy Score (`scoringRules`, Jordan *et al.*, 2019)



GDP

- ols always outperforms both wls and shr

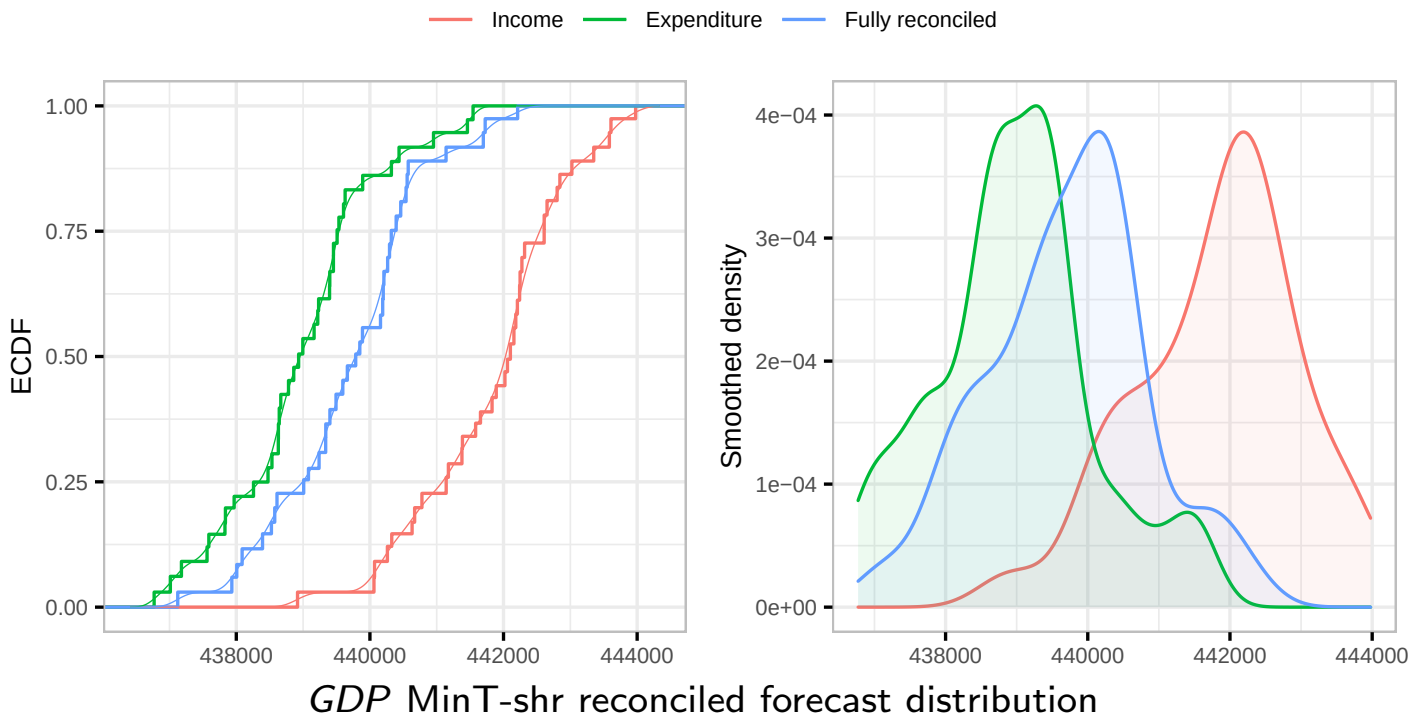
All variables

- reconciliation improves on the base forecasts' accuracy
- shr and wls are almost always the best

ITISE 2022

Conclusions

- Extension of the classical point and probabilistic forecast reconciliation results
- In the forecasting experiment, the simultaneously reconciled probabilistic forecasts give results as good as those of Athanasopoulos *et al.* (2020)
- The newly proposed approach produces forecasts that are **fully coherent** with all economic constraints coming from National Accounts relationships



ITISE 2022

800 1222·2022 ANNI

UNIVERSITÀ DEGLI STUDI DI PADOVA

ITISE 2022

Daniele Girolimetto

@ daniele.girolimetto@phd.unipd.it

github.com/daniGiro

Tommaso Di Fonzo

@ difonzo@stat.unipd.it

References I

[1] Athanasopoulos, G., Gamakumara, P., Panagiotelis, A., Hyndman, R.J., Affan, M. (2020), Hierarchical Forecasting, in Fuleky, P. (ed.), *Macroeconomic Forecasting in the Era of Big Data*, Cham, Springer, 689–719.

[2] Bisaglia, L., Di Fonzo, T., Girolimetto, D. (2020), Fully reconciled GDP forecasts from Income and Expenditure sides, in Pollice, A., Salvati, N., Schirripa Spagnolo, F. (eds.), *Book of short papers SIS 2020*, 951–956, Pearson.

[3] Di Fonzo, T., and Girolimetto, D. (2021), Cross-temporal forecast reconciliation: Optimal combination method and heuristic alternatives, *International Journal of Forecasting*, in press.

[4] Gamakumara, P., Panagiotelis, A., Athanasopoulos, G., Hyndman, R.J. (2018), *Probabilistic Forecasts in Hierarchical Time Series*, Department of Econometrics and Business Statistics, Monash University, Working Paper 11/18.

[5] Girolimetto, D., Di Fonzo, T. (2022), *FoReco: Point Forecast Reconciliation*, R package version 0.2.4, https://CRAN.R-project.org/package=FoReco.

[6] Gneiting, T. and Katzfuss, M. (2014), Probabilistic forecasts, *Annual Review of Statistics and Its Application*, 1, 125-151.

[7] 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.

[8] Hyndman, R. J., Lee, A.J., and Wang E., (2016), Fast computation of reconciled forecasts for hierarchical and grouped time series, *Computational Statistics & Data Analysis*, 97, 16–32.

ITISE 2022

References II

[9] Jordan A, Krüger F, Lerch S (2019). Evaluating Probabilistic Forecasts with scoringRules. *Journal of Statistical Software*, 90(12), 1–37.

[10] Panagiotelis, A., Gamakumara, P., Athanasopoulos, G., Hyndman, R.J. (2020), *Probabilistic Forecast Reconciliation: Properties, Evaluation and Score Optimisation*, Monash University, Department of Econometrics and Business Statistics, Working Paper 26/20.

[11] Panagiotelis, A., Gamakumara, P., Athanasopoulos, G., Hyndman, R.J. (2021), Forecast reconciliation: A geometric view with new insights on bias correction, *International Journal of Forecasting*, 37, 1, 343–359.

[12] Strang, G., (2019), *Linear algebra and learning from data*, Wellesley, Cambridge Press.

[13] 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.

ITISE 2022

Joint bootstrap on the base forecasts

Gamakumara *et al.* (2018)

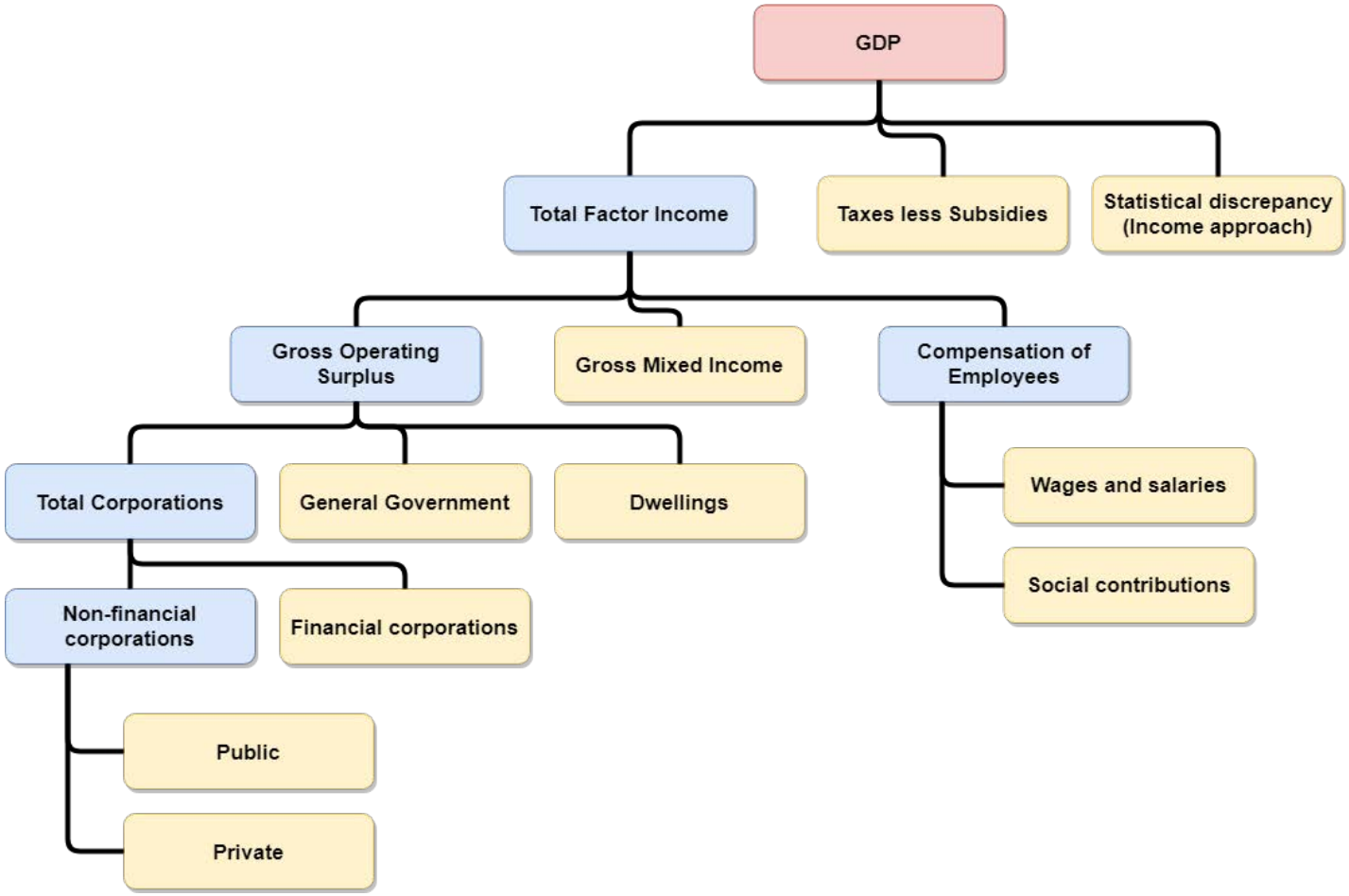
1. appropriate univariate models M_i for each series in the system are fitted based on the training data $\{y_{i,t}\}_{t=1}^{T}$, $i = 1, ..., n$, and the one-step-ahead in-sample forecast errors are stacked in an $(n \times T)$ matrix, $\widehat{\mathbf{E}} = \{\widehat{e}_{i,t}\}$;

2. $\widehat{\mathbf{y}}_{i,h}^{[l]} = f_i\left(M_i, \widehat{e}_{i,h}^{[l]}\right)$ is computed for $h = 1, ..., H$ and $l = 1, ..., L$, where $f(\cdot)$ is a function of the fitted univariate model and associated error, $\widehat{\mathbf{y}}_{i,h}^{[l]}$ is a sample path simulated for the $i-$th series, and $\widehat{e}_{i,h}^{[l]}$ is the $(i,h)-$th element of an $(n \times H)$ block bootstrap matrix containing H consecutive columns randomly drawn from $\widehat{\mathbf{E}}$.

ITISE 2022

Australian *GDP* Income side

15 variables + *GDP*

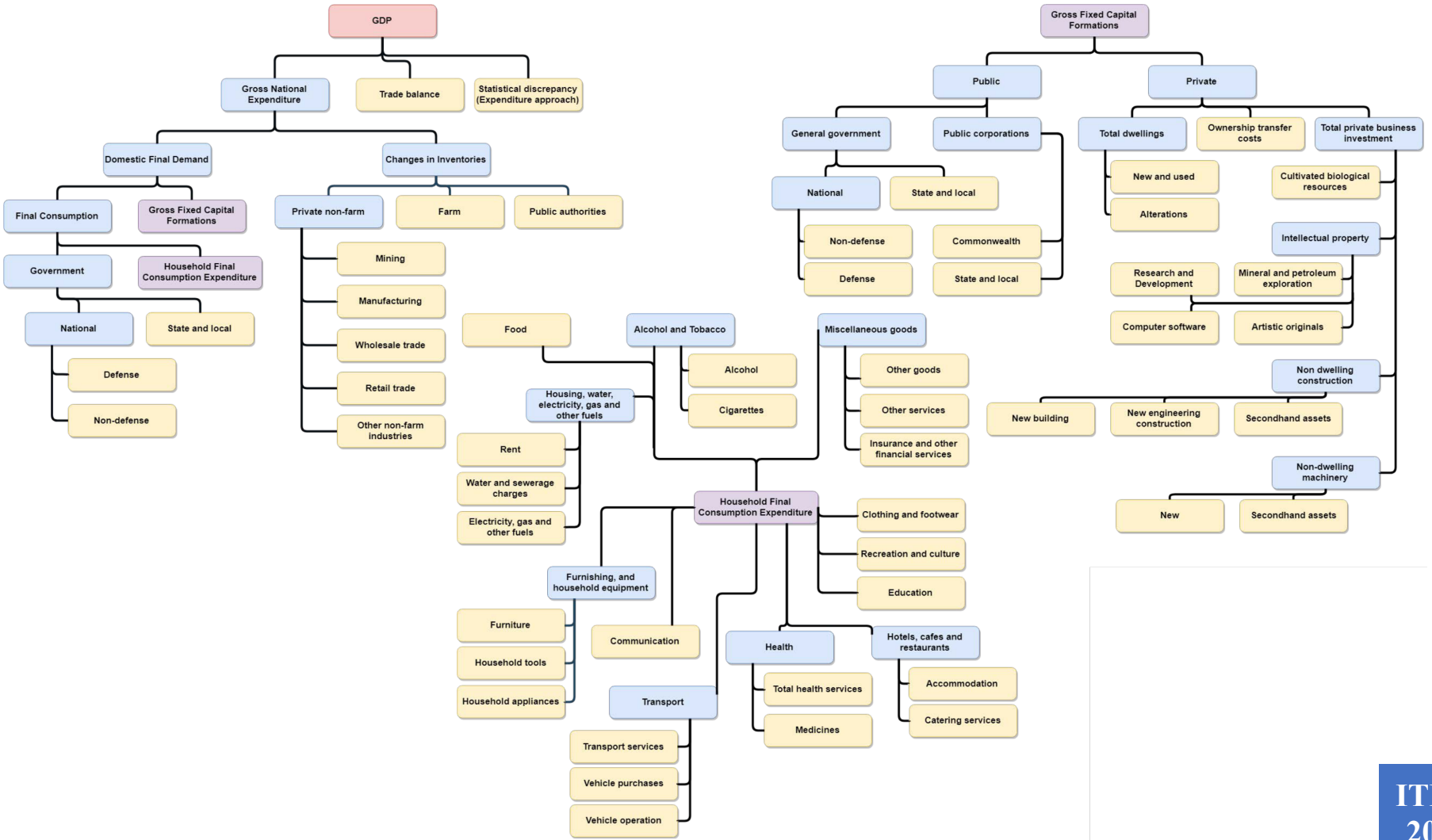


Source: Athanasopoulos et al., 2020, p. 702.

ITISE 2022

Australian *GDP* Expenditure side

79 variables + *GDP*



Source: Athanasopoulos et al., 2020, pp. 703 – 705.

ITISE 2022

FoReco v. 0.2.4

An **R** package for point forecast reconciliation

The **FoReco** (**Fo**recast **Reco**nciliation) package is designed for point forecast reconciliation, a **post-forecasting process** aimed to improve the accuracy of the base forecasts for a system of linearly constrained (e.g. hierarchical/grouped) time series.

It offers classical (bottom-up and top-down), and modern (optimal and heuristic combination) forecast reconciliation procedures for cross-sectional, temporal, and cross-temporal **linearly constrained time series**.

Links:

- **R CRAN**: `https://cran.r-project.org/package=FoReco`
- Code (**GitHub**): `https://github.com/daniGiro/FoReco`
- Site documentation: `https://danigiro.github.io/FoReco`

ITISE 2022

Tommaso Di Fonzo

Full Professor

@ difonzo@stat.unipd.it

Department of Statistical Sciences,
University of Padua (Italy)
Via Cesare Battisti, 241/243, 35121 Padua

Accuracy indices for probabilistic forecasts

Cumulative Rank Probability Score

$$\text{CRPS}(\hat{P}_i, z_i) = \frac{1}{L}\sum_{l=1}^{L}|x_{i,l} - z_i| - \frac{1}{2L^2}\sum_{l=1}^{L}\sum_{j=1}^{L}|x_{i,l} - x_{i,j}|$$

Energy Score

$$\text{ES}(\hat{P}, \mathbf{z}) = \frac{1}{L}\sum_{l=1}^{L}\|\mathbf{x}_l - \mathbf{z}\|_2 - \frac{1}{2(L-1)}\sum_{i=1}^{L-1}\|\mathbf{x}_l - \mathbf{x}_{l+1}\|_2$$

- $\hat{P}_i(\omega) = \frac{1}{L}\sum_{l=1}^{L} 1\left(x_{i,l} \leq \omega\right)$
- $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_L$ is a collection of L random draws taken from the predictive distribution
- $\mathbf{z} \in \mathbb{R}^n$ is the observation vector
- $i = 1, ..., n$ denotes a single variable

ITISE 2022