Forecast reconciliation: Methodological issues and applications

Chapter 4 - Forecast reconciliation for general linearly constrained multiple time series¹

Online appendix

Candidate: Daniele Girolimetto

Supervisors: Prof. Tommaso Di Fonzo, Prof. George Athanasopoulos

and Prof. Rob J Hyndman

Contents

A	Exa	Examples of derivation of the linear combination matrix A using the Reduced Row					
	Ech	elon Form (rref)	2				
	A.1	Linearly constrained multiple time series, $n = 5$ and $p = 3 \dots \dots \dots$	2				
	A.2	Linearly constrained multiple time series, $n = 4$ and $p = 3 \dots \dots \dots$	3				
	A.3	Linearly constrained multiple time series, $n = 35$ and $p = 15 \dots \dots \dots$	4				
_							
В	Aus	$tralian \ GDP$ from income and expenditure sides	10				
	B.1	Australian National Accounts variables	10				
	B.2	Accuracy indices	13				
C	Euro	opean Area GDP from output, income and expenditure sides	17				
	C.1	European National Accounts variables	17				
	C.2	Accuracy indices	18				

¹Girolimetto, D. and Di Fonzo, T. (2023b) Point and probabilistic forecast reconciliation for general linearly constrained multiple time series. arXiv doi:10.48550/arXiv.2305.05330

A Examples of derivation of the linear combination matrix A using the Reduced Row Echelon Form (rref)

A.1 Linearly constrained multiple time series, n = 5 and p = 3

First, consider a system of n = 5 linearly constrained time series such that for t = 1, ..., T

$$\begin{cases} 2x_{1,t} - 4x_{2,t} - 8x_{3,t} + 6x_{4,t} + 3x_{5,t} = 0 \\ x_{2,t} + 3x_{3,t} + 2x_{4,t} + 3x_{5,t} = 0 \\ 3x_{1,t} - 2x_{2,t} + 8x_{5,t} = 0 \end{cases}$$

or, through a matrix representation,

$$\Gamma_{(3\times5)}x_t=\mathbf{0}_{(3\times1)}, \qquad t=1,\ldots,T,$$

where $x_t = [x_{1,t} \ x_{2,t} \ x_{3,t} \ x_{4,t} \ x_{5,t}]'$ and

$$\Gamma_{(3\times5)} = \left[egin{array}{cccccc} 2 & -4 & -8 & 6 & 3 \ 0 & 1 & 3 & 2 & 3 \ 3 & -2 & 0 & 0 & 8 \end{array}
ight].$$

Reducing the coefficient matrix $\Gamma_{(3\times5)}$ to $\mathbf{Z}_{(3\times5)}$ (rref) yields

We observe that there are $n_c = 3$ constrained variables (red background) in position $\{1,2,4\}$ and $n_u = 2$ free variables (blue background) in position $\{3,5\}$. Therefore, we can build the permutation matrix

$$P_{(5 imes 5)} = \left[egin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \ 0 & 1 & 0 & 0 & 0 \ 0 & 0 & 1 & 0 & 0 \ 0 & 0 & 1 & 0 & 0 \ 0 & 0 & 0 & 0 & 1 \end{array}
ight]$$

and compute

$$C_{(3 imes5)} = \mathbf{Z}P' = \left[egin{array}{c|cccc} 1 & 0 & 0 & 2 & 4 \ 0 & 1 & 0 & 3 & 2 \ 0 & 0 & 1 & 0 & 0.5 \end{array}
ight].$$

The linear combination matrix A and structural-like matrix S are given by respectively,

$$egin{aligned} A_{(3 imes2)} = \left[egin{array}{ccc} -2 & -4 \ -3 & -2 \ 0 & -0.5 \end{array}
ight] & ext{and} & S_{(5 imes2)} = \left[egin{array}{ccc} -2 & -4 \ -3 & -2 \ 0 & -0.5 \ \hline 1 & 0 \ 0 & 1 \end{array}
ight]
ight\} oldsymbol{I}_{2}. \end{aligned}$$

These matrices are such that

$$c_t = A_{(3\times 2)}u_t, y_t = S_{(5\times 2)}u_t, \text{for } t = 1,...,T,$$

where

$$c_t = \left[egin{array}{c} x_{1,t} \ x_{2,t} \ x_{4,t} \end{array}
ight], \quad u_t = \left[egin{array}{c} x_{3,t} \ x_{5,t} \end{array}
ight], \quad ext{and} \quad y_t = \left[egin{array}{c} c_t \ u_t \end{array}
ight].$$

A.2 Linearly constrained multiple time series, n = 4 and p = 3

First, consider a system of n = 4 linearly constrained time series such that for t = 1, ..., T

$$\begin{cases} x_{1,t} - 2x_{2,t} - x_{3,t} + 3x_{4,t} = 0 \\ 2x_{1,t} - 4x_{2,t} - 3x_{3,t} + 2x_{4,t} = 0 \\ 4x_{1,t} - 8x_{2,t} - 6x_{3,t} + 4x_{4,t} = 0 \end{cases}$$

and

$$\Gamma_{(3\times4)}x_t=\mathbf{0}_{(3\times1)}, \qquad t=1,\ldots,T,$$

where $x_t = [x_{1,t} \ x_{2,t} \ x_{3,t} \ x_{4,t}]'$ and

$$\Gamma = \begin{bmatrix} 1 & -2 & -1 & 3 \\ 2 & -4 & -3 & 2 \\ 4 & -8 & -6 & 4 \end{bmatrix} \xrightarrow{rref} \begin{bmatrix} 1 & -2 & 0 & 7 \\ 0 & 0 & 1 & 4 \\ \hline 0 & 0 & 0 & 0 \end{bmatrix}.$$

In this case, it clearly appears that there is a redundant relationship (e.g., the third equation is equal to the second one multiplied by 2). This is confirmed by the fact that, reducing the coefficient matrix Γ , we obtain $n_c = 2$ constrained variables (red background) in position $\{1,4\}$, $n_u = 2$ free variables (blue background) in position $\{2,3\}$, and the last row is null. Then,

$$m{P}_{(4 imes4)} = egin{bmatrix} 1 & 0 & 0 & 0 \ 0 & 0 & 1 & 0 \ 0 & 1 & 0 & 0 \ 0 & 0 & 0 & 1 \end{bmatrix}$$
 , $m{Z}_{(2 imes4)} = egin{bmatrix} 1 & -2 & 0 & 7 \ 0 & 0 & 1 & 4 \end{bmatrix}$

and

$$C_{(2 imes4)}=ZP'=\left[egin{array}{c|c}1&0&-2&7\0&1&0&4\end{array}
ight].$$

Finally,

$$m{A}_{(3 imes2)} = \left[egin{array}{ccc} -2 & 7 \ 0 & 4 \end{array}
ight] \quad ext{and} \quad m{S}_{(5 imes2)} = \left[egin{array}{ccc} -2 & 7 \ 0 & 4 \ -1 & 0 \ 0 & 1 \end{array}
ight] m{A}_{(2 imes2)} \ m{I}_2 \ .$$

	Hierarchy				Complete
	1	2	3	4	structure
$\overline{n_b}$	4	7	3	8	22
n_a	3	5	1	6	13
п	7	12	4	14	35

	Hierarchy				Complete
	1	2	3	4	structure
n_u	4	7	3	8	20
n_u n_c	3	5	1	6	15
n	7	12	4	14	35

(a) Bottom and upper variables

(b) Free (unconstrained) and constrained variables

Table A.1: Table A.1a reports the number of bottom (n_b) , upper (n_a) and all (n) time series of the four component hierarchies. Table A.1b reports the number of free (n_u) , constrained (n_c) and all (n) time series.

A.3 Linearly constrained multiple time series, n = 35 and p = 15

Figure A.1 shows a rather complex linearly constrained system, consisting in four hierarchies sharing only the top-level variables. Hierarchies 1 and 2 share the same top-level variable X and hierarchies 3 and 4 share the same top-level variable Y. The total variable Z is the sum of X and Y. Table A.1a reports the number of bottom (n_b) , upper (n_a) and all (n) time series of the four component hierarchies and Table A.1b, the number of free (n_u) , constrained (n_c) and all (n) time series. This type of linear constrained structure is not unusual in National Accounts, for example when GDP from different accounting sides is considered.

Moving from the top of the structure in Figure A.1, the relationships between the 35 variables can be expressed as

- 1. Z = X + Y
- 2. X = A + B
- 3. X = C + D
- 4. Y = E + F + G
- 5. Y = H + I
- 6. A = AB + AAA + AAB
- 7. AA = AAA + AAB
- 8. C = CA + CB + CC
- 9. D = DAA + DAB + DBA + DBB
- 10. DA = DAA + DAB
- 11. DB = DBA + DBB
- 12. H = HA + HB
- 13. HA = HAA + HAB + HAC
- 14. I = IB + IC + IAA + IAB
- 15. IA = IAA + IAB

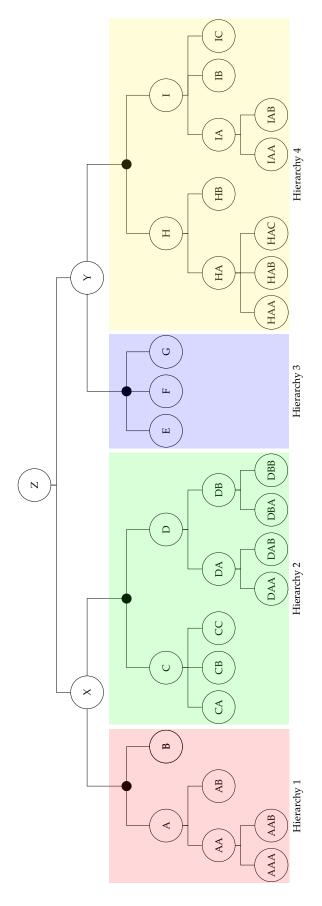


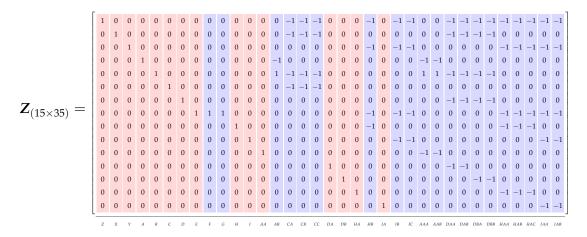
Figure A.1: Representation of a multiple linearly constrained time series formed by four genuine hierarchies in which hierarchies 1 and 2 share the same top-level variable X and hierarchies 3 and 4 share the same top-level variable Y. The total variable Z is the sum of X and Y.

These relationships can be grouped into a system of 15 equations:

$$x_{Z,t} - x_{X,t} - x_{Y,t} = 0,$$
 $x_{DAA,t} - x_{DAB,t} - x_{DBA,t} - x_{DBB,t} = 0,$ $x_{X,t} - x_{A,t} - x_{B,t} = 0,$ $x_{DA,t} - x_{DAA,t} - x_{DAB,t} - x_{DBB,t} = 0,$ $x_{DA,t} - x_{DAA,t} - x_{DAB,t} - x_{DAB,t} = 0,$ $x_{DB,t} - x_{DBA,t} - x_{DBB,t} = 0,$ $x_{DB,t} - x_{DBB,t} - x_{DBB,t} = 0,$ $x_{DB,t} - x_{DB,t} - x_{DBB,t} = 0,$ $x_{DB,t} - x_{DBB,t} - x_{DBB,t} = 0,$ $x_{DB,t} - x_{DBB,t} - x_{DBB,t} = 0,$ $x_{DB,t} - x_{DB,t} - x_{DBB,t} = 0,$ $x_{DB,t} - x_{DB,t} - x_{DB,t} - x_{DB,t} = 0,$ $x_{DB,t} - x_{DB,t} - x_{DB,t} - x_{DB,t} = 0,$ $x_{DB,t} - x_{DB,t} - x_{DB,t} - x_{DB,t} = 0,$ $x_{DB,t} - x_{DB,t} - x_{DB,t} - x_{DB,t} = 0,$ $x_{DB,t} - x_{DB,t} - x_{DB,t} - x_{DB,t} = 0,$ $x_{DB,t} - x_{DB,t} - x_{DB,t} - x_{DB,t} = 0,$ $x_{DB,t} - x_{DB,t} - x_{DB,t} - x_{DB,t} = 0,$ $x_{DB,t} - x_{DB,t} - x_{DB,t} - x_{DB,t} - x_{DB,t} = 0,$ $x_{DB,t} - x_{DB,t} - x_{DB,t} - x_{DB,t} = 0,$ $x_{DB,t} - x_{DB,t} - x_{DB,t} - x_{DB,t} = 0,$ $x_{DB,t} - x_{DB,t} - x_{DB,t} - x_{DB$

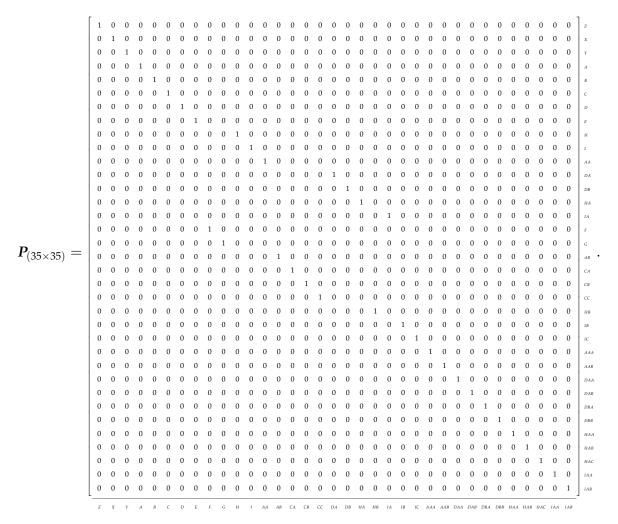
with coefficient matrix

Reducing this matrix through rref yields



and

where P is a (35 × 35) permutation matrix moving the pivot columns (red background) in Z to the left side of matrix C:



Finally, the linear combination matrix *A* is given by:

It's worth noting that *B* and *E*, originally considered as bottom variables in hierarchies 1 and 3, respectively, in the general linear constrained representation are treated as constrained variables.

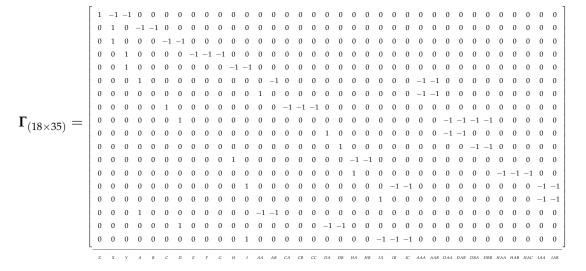
Consider now the situation in which the user works with a number of relationships larger than the minimum needed to define the set of linearly independent constraints. For example, the following 3 (redundant) relationships are added to previous ones:

$$6.1 A = AA + AB$$

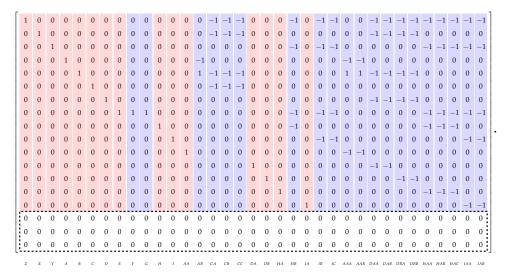
$$9.1 D = DA + DB$$

14.1
$$I = IA + IB + IC$$

In this case, matrix Γ has 18 rows instead of 15, and



↓ rref



In the end, the redundant relationships have the effect of adding three null rows to the rref of matrix Γ , but the same free (20) and constrained (15) variables are found.

B Australian GDP from income and expenditure sides

B.1 Australian National Accounts variables

Variable	Series ID	Description
Gdp Sde Exp Imp Gne	A2302467A A2302566J A2302564C A2302565F A2302563A	GDP - Gross Domestic Product Statistical Discrepancy(E) Exports of goods and services Imports of goods and services Gross national exp.
GneDfdFceGvtNatDef GneDfdFceGvtNatNdf GneDfdFceGvtNat GneDfdFceGvtSnl GneDfdFceGvt	A2302523J A2302524K A2302525L A2302526R A2302527T	Gen. gov National; Final consumption exp Defence Gen. gov National; Final consumption exp Non-defence Gen. gov National; Final consumption exp. Gen. gov State and local; Final consumption exp, Gen. gov.; Final consumption exp.
GneDfdFce GneDfdGfcPvtTdwNnu GneDfdGfcPvtTdwAna GneDfdGfcPvtTdw GneDfdGfcPvtOtc	A2302529W A2302543T A2302544V A2302545W A2302546X	All sectors; Final consumption exp. Pvt.; Gross fixed capital formation (GFCF) Pvt.; GFCF - Dwellings - Alterations and additions Pvt.; GFCF - Dwellings - Total Pvt.; GFCF - Ownership transfer costs
GneDfdGfcPvtPbiNdcNbd GneDfdGfcPvtPbiNdcNec	A2302533L A2302534R	Pvt. GFCF - Non-dwelling construction - New building Pvt.; GFCF - Non-dwelling construction - New engineering
GneDfdGfcPvtPbiNdcSha	A2302535T	construction Pvt.; GFCF - Non-dwelling construction - Net purchase of second hand assets
GneDfdGfcPvtPbiNdc GneDfdGfcPvtPbiNdmNew GneDfdGfcPvtPbiNdmSha	A2302536V A2302530F A2302531J	Pvt.; GFCF - Non-dwelling construction - Total Pvt.; GFCF - Machinery and equipment - New Pvt.; GFCF - Machinery and equipment - Net purchase of second hand assets
GneDfdGfcPvtPbiNdm	A2302532K	Pvt.; GFCF - Machinery and equipment - Total
GneDfdGfcPvtPbiCbr GneDfdGfcPvtPbiIprRnd	A2716219R A2716221A	Pvt.; GFCF - Cultivated biological resources Pvt.; GFCF - Intellectual property products - Research and de-
GneDfdGfcPvtPbiIprMnp	A2302539A	velopment Pvt.; GFCF - Intellectual property products - Mineral and petroleum exploration
GneDfdGfcPvtPbiIprCom GneDfdGfcPvtPbiIprArt GneDfdGfcPvtPbiIpr GneDfdGfcPvtPbi GneDfdGfcPvt	A2302538X A2302540K A2716220X A2302542R A2302547A	Pvt.; GFCF - Intellectual property products - Computer software Pvt.; GFCF - Intellectual property products - Artistic originals Pvt.; GFCF - Intellectual property products Total Pvt.; GFCF - Total private business investment Pvt.; GFCF
GneDfdGfcPubPcpCmw GneDfdGfcPubPcpSnl GneDfdGfcPubPcp GneDfdGfcPubGvtNatDef GneDfdGfcPubGvtNatNdf	A2302548C A2302549F A2302550R A2302551T A2302552V	Plc. corporations - Commonwealth; GFCF Plc. corporations - State and local; GFCF Plc. corporations; GFCF Total Gen. gov National; GFCF - Defence Gen. gov National; GFCF - Non-defence
GneDfdGfcPubGvtNat GneDfdGfcPubGvtSnl GneDfdGfcPubGvt GneDfdGfcPub GneDfdGfc	A2302553W A2302554X A2302555A A2302556C A2302557F	Gen. gov National ; GFCF Total Gen. gov State and local; GFCF Gen. gov.; GFCF Plc.; GFCF All sectors; GFCF

Table B.2: *Variables, series IDs and their descriptions for the expenditure approach. Source: Athanasopoulos et al.* (2020)

Variable	Series ID	Description
GneDfdHfc GneDfdFceHfcFud GneDfdFceHfcAbt GneDfdFceHfcAbtCig GneDfdFceHfcAbtAlc	A2302254W A2302237V A3605816F A2302238W A2302239X	Household Final Consumption expenditure Food Alcoholic beverages and tobacco Cigarettes and tobacco Alcoholic beverages
GneDfdFceHfcCnf GneDfdFceHfcHwe GneDfdFceHfcHweRnt GneDfdFceHfcHweWsc GneDfdFceHfcHweEgf	A2302240J A3605680F A3605681J A3605682K A2302242L	Clothing and footwear Housing, water, electricity, gas and other fuels Actual and imputed rent for housing Water and sewerage charges Electricity, gas and other fuel
GneDfdFceHfcFhe GneDfdFceHfcFheFnt GneDfdFceHfcFheApp GneDfdFceHfcFheTls GneDfdFceHfcHlt	A2302243R A3605683L A3605684R A3605685T A2302244T	Furnishings and household equipment Furniture, floor coverings and household goods Household appliances Household tools Health
GneDfdFceHfcHltMed GneDfdFceHfcHltHsv GneDfdFceHfcTpt GneDfdFceHfcTptPvh GneDfdFceHfcTptOvh	A3605686V A3605687W A3605688X A2302245V A2302246W	Medicines, medical aids and therapeutic appliances Total health services Transport Purchase of vehicles Operation of vehicles
GneDfdFceHfcTptTsv GneDfdFceHfcCom GneDfdFceHfcRnc GneDfdFceHfcEdc GneDfdFceHfcHcr	A2302247X A2302248A A2302249C A2302250L A2302251R	Transport services Communications Recreation and culture Education services Hotels, cafes and restaurants
GneDfdFceHfcHcrCsv GneDfdFceHfcHcrAsv GneDfdFceHfcMis GneDfdFceHfcMisOgd GneDfdFceHfcMisIfs GneDfdFceHfcMisOsv	A3605694V A3605695W A3605696X A3605697A A2302252T A3606485T	Catering services Accommodation services Miscellaneous goods and services Other goods Insurance and other financial services Other services

Table B.3: *Variables, series IDs and their descriptions for Household Final Consumption - expenditure approach. Source: Athanasopoulos et al.* (2020)

Variable	Series ID	Description
GneCii	A2302562X	Changes in Inventories
GneCiiPfm	A2302560V	Farm
GneCiiPba	A2302561W	Public authorities
GneCiiPnf	A2302559K	Private; Non-farm Total
GneCiiPnfMin	A83722619L	Private; Mining (B)
GneCiiPnfMan	A3348511X	Private; Manufacturing (C)
GneCiiPnfWht	A3348512A	Private; Wholesale trade (F)
GneCiiPnfRet	A3348513C	Private; Retail trade (G)
GneCiiPnfOnf	A2302273C	Private; Non-farm; Other non-farm industries

Table B.4: *Variables, series IDs and their descriptions for Changes in Inventories - expenditure approach. Source: Athanasopoulos et al.* (2020)

Variable	Series ID	Description
Sdi Tsi TfiCoeWns TfiCoeEsc	A2302413V A2302412T A2302399K A2302400J	Statistical discrepancy (I) Taxes less subsidies (I) Compensation of employees; Wages and salaries Compensation of employees; Employers' social contributions
TfiCoe TfiGosCopNfnPvt TfiGosCopNfnPub TfiGosCopNfn TfiGosCopFin	A2302401K A2323369L A2302403R A2302404T A2302405V	Compensation of employees Private non-financial corporations; Gross operating surplus Public non-financial corporations; Gross operating surplus Non-financial corporations; Gross operating surplus Financial corporations; Gross operating surplus
TfiGosCop TfiGosGvt TfiGosDwl TfiGos TfiGmi Tfi	A2302406W A2298711F A2302408A A2302409C A2302410L A2302411R	Total corporations; Gross operating surplus General government; Gross operating surplus Dwellings owned by persons; Gross operating surplus All sectors; Gross operating surplus Gross mixed income Total factor income

Table B.5: Variables, series IDs and their descriptions for the income approach. Source: Athanasopoulos et al. (2020)

B.2 Accuracy indices

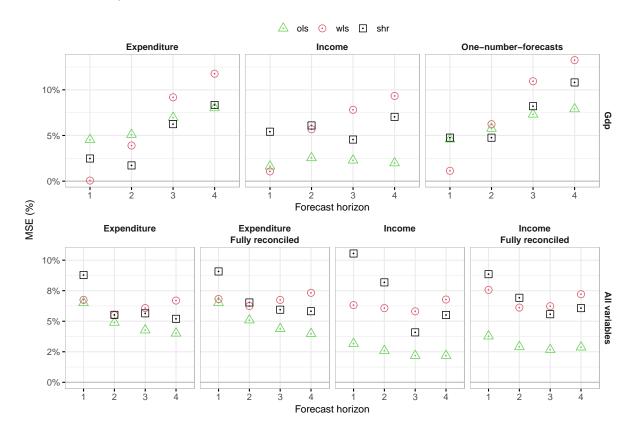


Figure B.2: MSE-skill scores (relative to base forecasts) for point forecasts from alternative reconciliation approaches (Australian QNA variables).

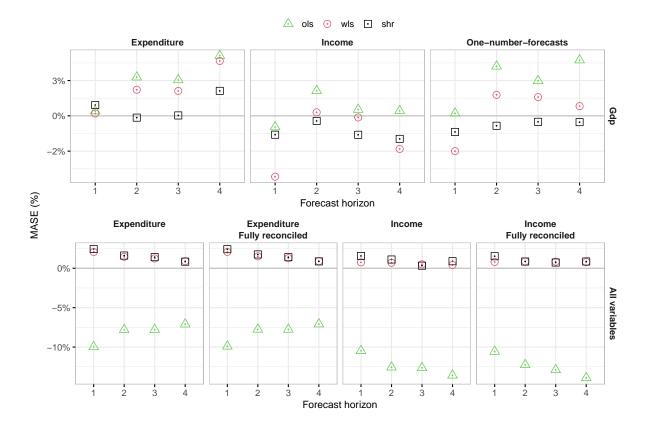


Figure B.3: *MASE-skill scores (relative to base forecasts) for point forecasts from alternative reconciliation approaches (Australian QNA variables).*

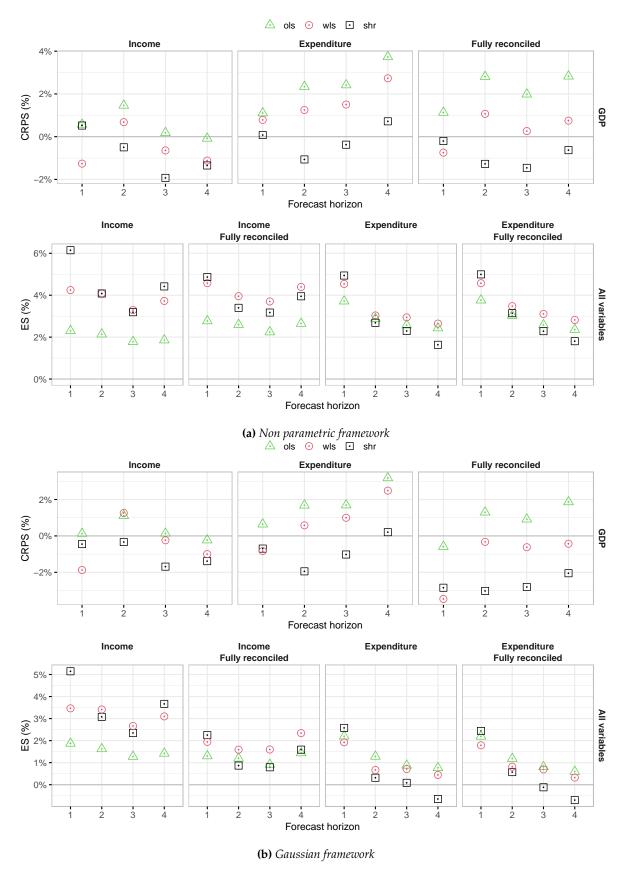


Figure B.4: *CRPS and ES-skill scores (relative to base forecasts) for the probabilistic forecasts from alternative reconciliation approaches (Australian QNA variables).*

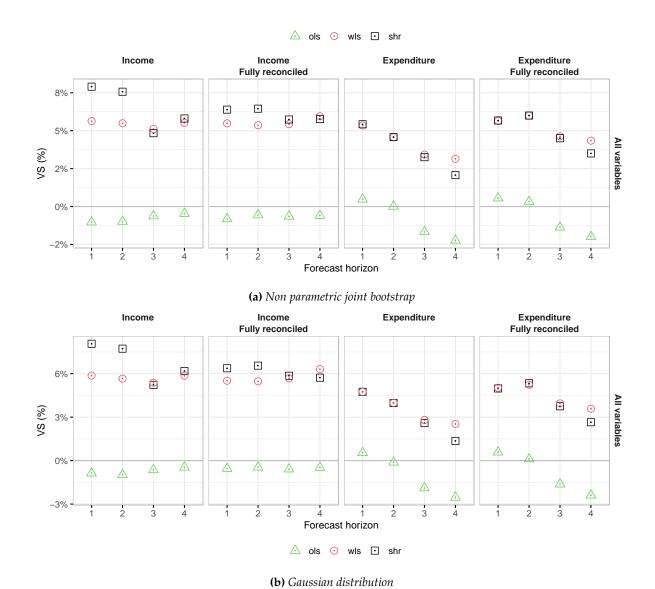


Figure B.5: VS-skill scores (relative to base forecasts) for the probabilistic forecasts from alternative reconciliation approaches (Australian QNA variables).

C European Area GDP from output, income and expenditure sides

C.1 European National Accounts variables

ID	Description
GDP	Gross domestic product
	expenditure side
P3	Final consumption expenditure
P3_S13	Final consumption expenditure of general government
P31_S141_S15	Household and NPISH final consumption expenditure
P31_S13	Individual consumption expenditure of general government
P5G	Gross capital formation
P521_P53	Changes in inventories and acquisitions less disposals of valuables
P6	Exports of goods and services
P7	Imports of goods and services
B11	External balance of goods and services
B111	External balance - Goods
B112	External balance - Services
P3_P5	Final consumption expenditure and gross capital formation
P3_P6	Final consumption expenditure, gross capital formation and exports of
	goods and services
P32_S13	Collective consumption expenditure of general government
P31_S14	Final consumption expenditure of households
P31_S15	Final consumption expenditure of NPISH
P51G	Gross fixed capital formation
P52	Changes in inventories
P53	Acquisitions less disposals of valuables
P61	Exports of goods
P62	Exports of services
P71	Imports of goods
P72	Imports of services
YA0	Statistical discrepancy (expenditure approach)
	income side
D1	Compensation of employees
D2X3	Taxes on production and imports less subsidies
D11	Wages and salaries
D12	Employers' social contributions
B2A3G	Operating surplus and mixed income, gross
D2	Taxes on production and imports
D3	Subsidies
YA2	Statistical discrepancy (income approach)
	output side
D21X31	Taxes less subsidies on products
B1G	Total gross value added
YA1	Statistical discrepancy (output/production approach)

Table C.6: List of the European National Accounts variables (ESA 2010 classification). Source: https://ec.europa.eu/eurostat/web/national-accounts/data/database

C.2 Accuracy indices

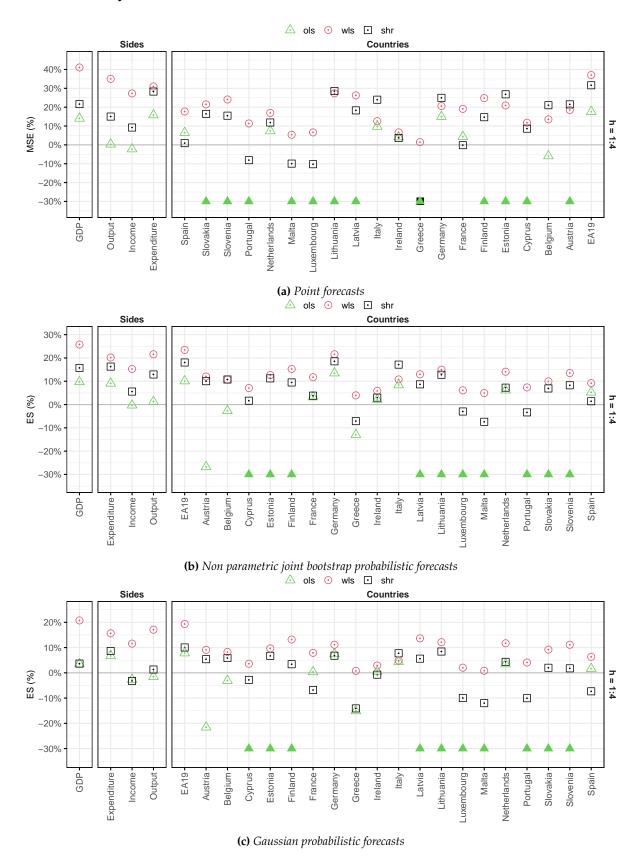


Figure C.6: *MSE* and *ES*-skill scores (relative to base forecasts) for the point and probabilistic forecasts from alternative reconciliation approaches (European Area QNA). To make the figure more readable, the filled symbols indicate that the skill score is less than -30%.

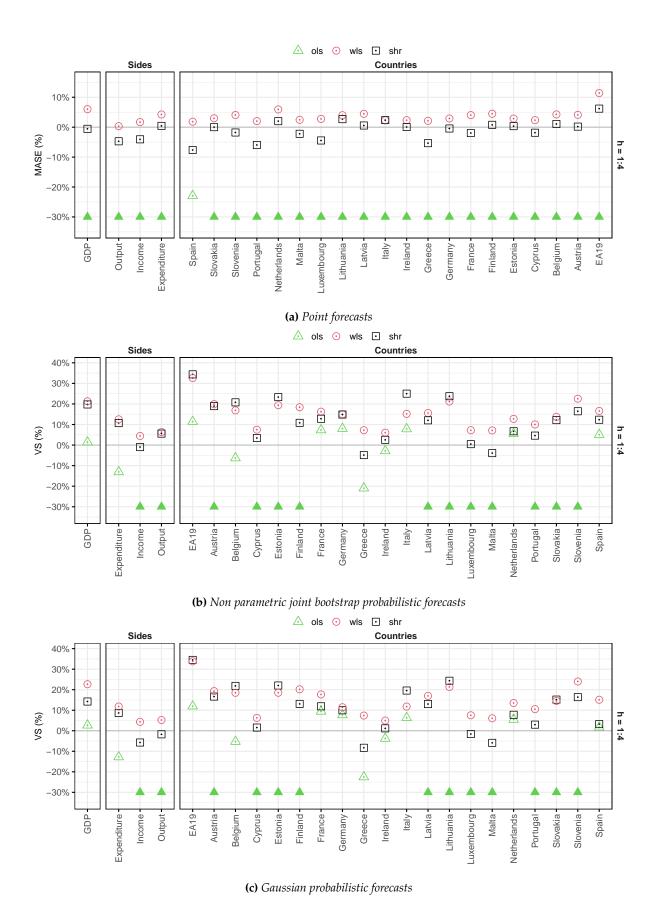


Figure C.7: *MASE* and *VS*-skill scores (relative to base forecasts) for the point and probabilistic forecasts from alternative reconciliation approaches (European Area QNA). To make the figure more readable, the filled symbols indicate that the skill score is less than -30%.

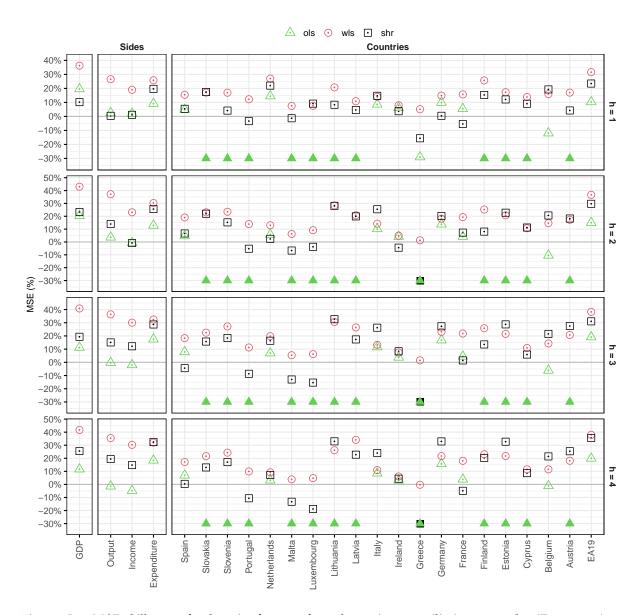


Figure C.8: MSE-skill scores for the point forecasts from alternative reconciliation approaches (European Area QNA) for different forecast horizons. To make the figure more readable, the filled symbols indicate that the skill score is less than -30%.

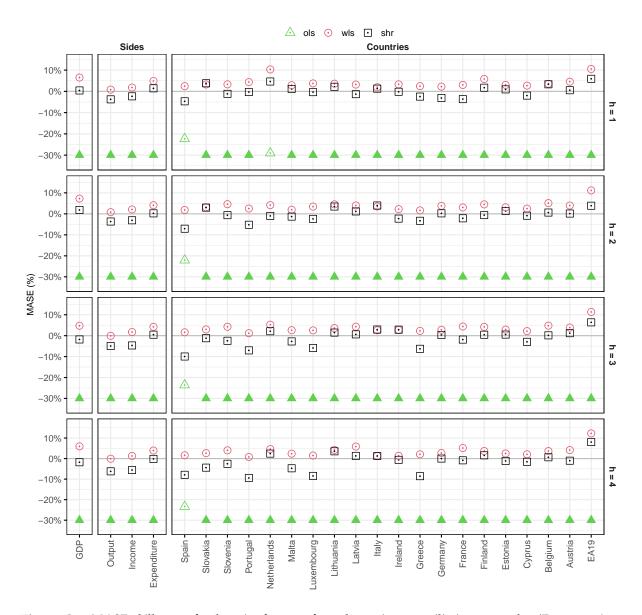


Figure C.9: *MASE-skill scores for the point forecasts from alternative reconciliation approaches* (European Area QNA) for different forecast horizons. To make the figure more readable, the filled symbols indicate that the skill score is less than -30%.

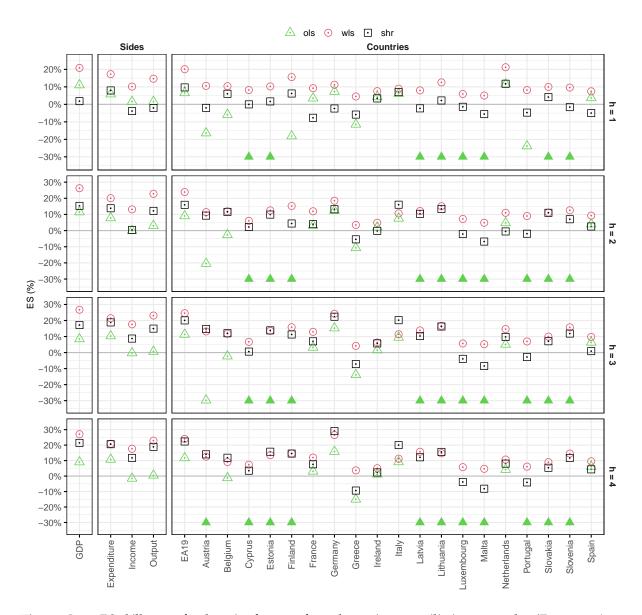


Figure C.10: ES-skill scores for the point forecasts from alternative reconciliation approaches (European Area QNA) for different forecast horizons. To make the figure more readable, the filled symbols indicate that the skill score is less than -30%.

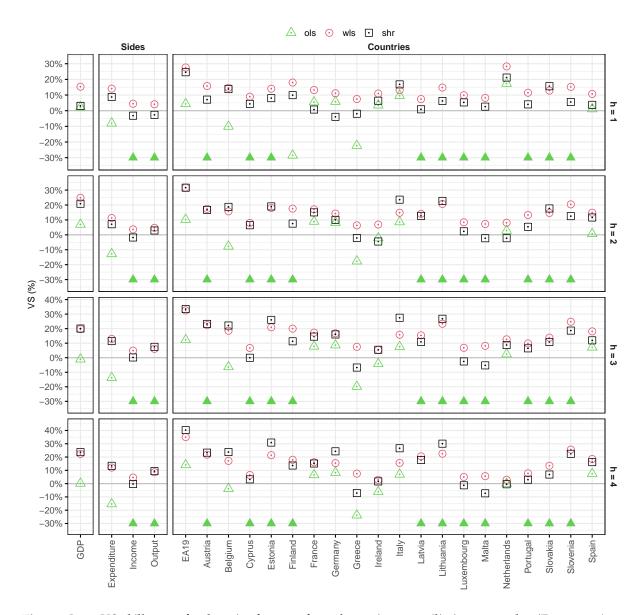


Figure C.11: VS-skill scores for the point forecasts from alternative reconciliation approaches (European Area QNA) for different forecast horizons. To make the figure more readable, the filled symbols indicate that the skill score is less than -30%.

References

Athanasopoulos, G., Gamakumara, P., Panagiotelis, A., Hyndman, R. J. & Affan, M. (2020), Hierarchical Forecasting, *in* P. Fuleky, ed., 'Macroeconomic Forecasting in the Era of Big Data', Vol. 52, Springer International Publishing, Cham, pp. 689–719.