Co-Err with Thy Neighbor: Understanding the Covariance Matrix in Hierarchical Forecasting

Kandrika Pritularga, Ivan Svetunkov, Nikolaos Kourentzes

Lancaster Universty

Thessaloniki, 18 June 2019

Marketing Analytics & Forecasting

Lancaster University Management School

- Introduction
- Hierarchical Forecasting
- Empirical Experimentations
- Remarks

Hierarchical Forecasting in Organisation

- Organisations are mostly structured in hierarchies
  - Based on geography, products, distribution
- Need to align operational decisions from the bottom to the top of the hierarchy
  - Reducing unnecessary costs, e.g. miscoordination, stock-out
  - Optimising the flow of information in a supply chain
- Forecasting becomes essential in order to plan in well-coordinated manner
  - Heuristic approach: Bottom-Up, Top-Down, Middle Out
  - Current approach: MinT Reconciliation (**?????**)

Current Method: MinT Reconciliation

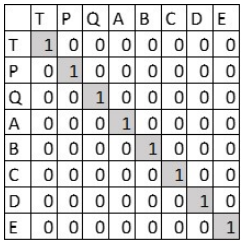
- MinT Reconciliation in a nutshell,

$$\begin{aligned} \tilde{\boldsymbol{y}}_{t+h} &= \boldsymbol{S}\boldsymbol{P}\hat{\boldsymbol{y}}_{t+h} \\ \boldsymbol{P} &= (\boldsymbol{S}'\boldsymbol{W}^{-1}\boldsymbol{S})^{-1}\boldsymbol{S}'\boldsymbol{W}^{-1} \end{aligned}$$

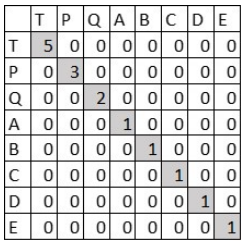
- $\boldsymbol{S}$ is a summation matrix, $\boldsymbol{P}$ is a reconciliation matrix, $\hat{\boldsymbol{y}}_{t+h}$ is unreconciled forecasts, and $\tilde{\boldsymbol{y}}_{t+h}$ is reconciled forecasts
- $\boldsymbol{W}$ is an sample covariance matrix from h-step ahead forecast errors but becomes difficult to estimate when $h > 1$
- Then $\boldsymbol{W}$ is approximated by one-step ahead in-sample forecast errors

MinT: What covariance matrix should we use?

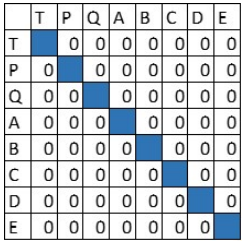
$$\boldsymbol{P} = (\boldsymbol{S}'\boldsymbol{W}^{-1}\boldsymbol{S})^{-1}\boldsymbol{S}'\boldsymbol{W}^{-1}$$



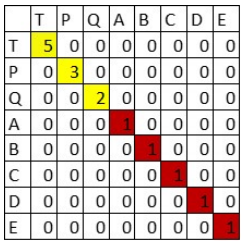
OLS



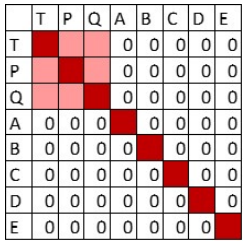
SS



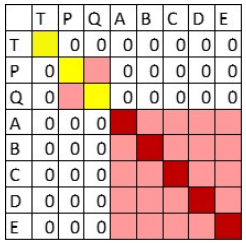
DiagW



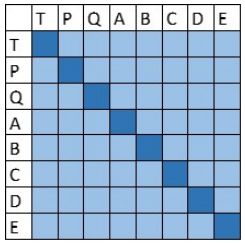
DiagW-SS



ccUB-DiagW



ccPS



MinT-Shrink

Structures only

Decent house

Fancy house

MinT: What covariance matrix should we use?

	T	P	Q	A	B	C	D	E
T	1	0	0	0	0	0	0	0
P	0	1	0	0	0	0	0	0
Q	0	0	1	0	0	0	0	0
A	0	0	0	1	0	0	0	0
B	0	0	0	0	1	0	0	0
C	0	0	0	0	0	1	0	0
D	0	0	0	0	0	0	1	0
E	0	0	0	0	0	0	0	1

OLS

	T	P	Q	A	B	C	D	E
T	5	0	0	0	0	0	0	0
P	0	3	0	0	0	0	0	0
Q	0	0	2	0	0	0	0	0
A	0	0	0	1	0	0	0	0
B	0	0	0	0	1	0	0	0
C	0	0	0	0	0	1	0	0
D	0	0	0	0	0	0	1	0
E	0	0	0	0	0	0	0	1

SS

	T	P	Q	A	B	C	D	E
T		0	0	0	0	0	0	0
P	0		0	0	0	0	0	0
Q	0	0		0	0	0	0	0
A	0	0	0		0	0	0	0
B	0	0	0	0		0	0	0
C	0	0	0	0	0		0	0
D	0	0	0	0	0	0		0
E	0	0	0	0	0	0	0	

DiagW

	T	P	Q	A	B	C	D	E
T	5	0	0	0	0	0	0	0
P	0	3	0	0	0	0	0	0
Q	0	0	2	0	0	0	0	0
A	0	0	0	1	0	0	0	0
B	0	0	0	0	1	0	0	0
C	0	0	0	0	0	1	0	0
D	0	0	0	0	0	0	1	0
E	0	0	0	0	0	0	0	1

DiagW-SS

	T	P	Q	A	B	C	D	E
T				0	0	0	0	0
P				0	0	0	0	0
Q				0	0	0	0	0
A	0	0	0		0	0	0	0
B	0	0	0	0		0	0	0
C	0	0	0	0	0		0	0
D	0	0	0	0	0	0		0
E	0	0	0	0	0	0	0	

ccUB-DiagW

	T	P	Q	A	B	C	D	E
T		0	0	0	0	0	0	0
P	0		0	0	0	0	0	0
Q	0	0		0	0	0	0	0
A	0	0	0		0	0	0	0
B	0	0	0	0		0	0	0
C	0	0	0	0	0		0	0
D	0	0	0	0	0	0		0
E	0	0	0	0	0	0	0	

ccPS

	T	P	Q	A	B	C	D	E
T								
P								
Q								
A								
B								
C								
D								
E								

MinT-Shrink

Structures only

Decent house

Fancy house

Little is known how to choose the covariance matrix!

Motivation

Research Question

- How can we choose the **covariance matrix** in MinT Reconciliation to **improve forecast accuracy**?

Contribution

- To propose **a selection in reconciliation weight estimation**
- To provide alternatives to MinT-Shrink with fewer estimations in the covariance matrix

MinT: How can we choose the covariance matrix?

- Measure the gain from MinT Reconciliation ($RelRMSE$)
  - $AvgRelRMSE = (\prod_{i=1}^{n} \frac{RMSE_{recon,i}}{RMSE_{base,i}})^{\frac{1}{n}}$
- Measure the covariance matrix (Q_2)
  - Measure the distance between the estimated covariance matrix and the 'true' one (**?**)
  - 'True' covariance matrix is unknown; we estimate a diagonal covariance matrix from STL residuals
    - Time series decomposition as proxies of DGP; it might be the 'true' residuals
    - A diagonal covariance matrix: 'perfect' models will yield white noise errors and those will not be correlated
  - Capture the complexities of the covariance matrix

MinT: How can we choose the covariance matrix?

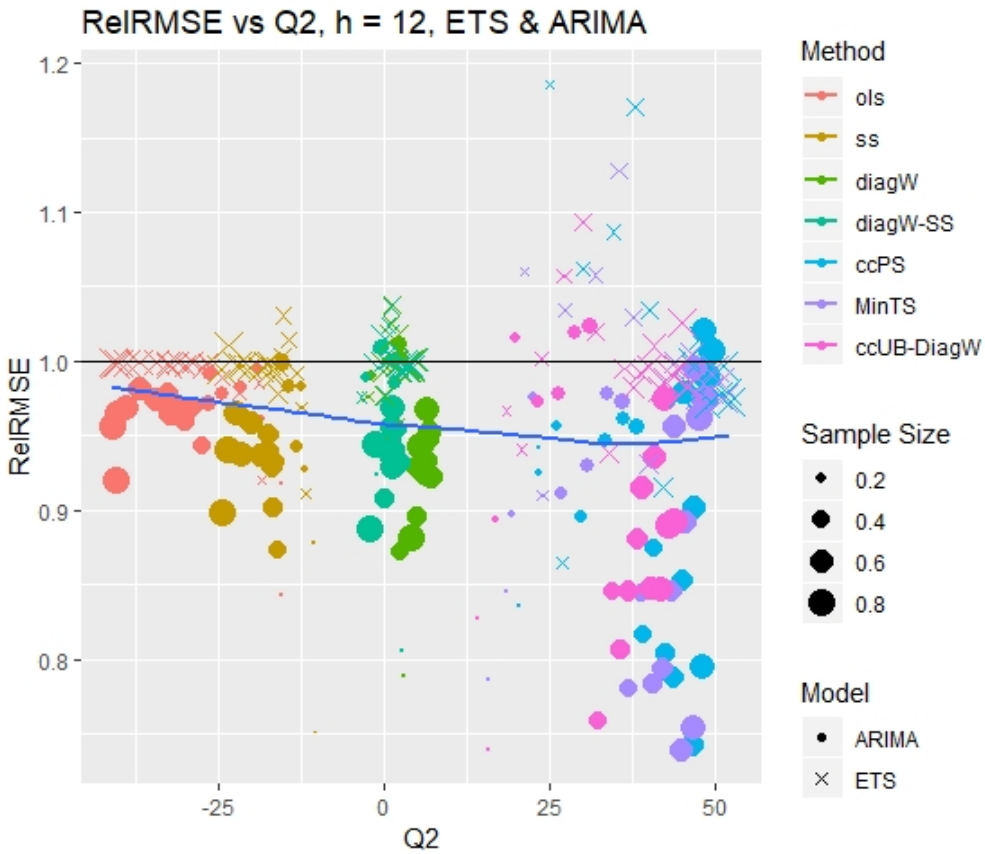
Consider other factors that may affect forecast reconciliation

- Sample sizes
  $Size = \frac{\#\text{ in-sample}}{\#\text{ total observation}}$
- Comparative performance of forecasting models
  $RelSE = \left[\frac{\text{standard error of estimate from forecasting models}}{\text{standard error of estimate from random walk}}\right]^{-1}$, for series i

Data and Forecasting Task

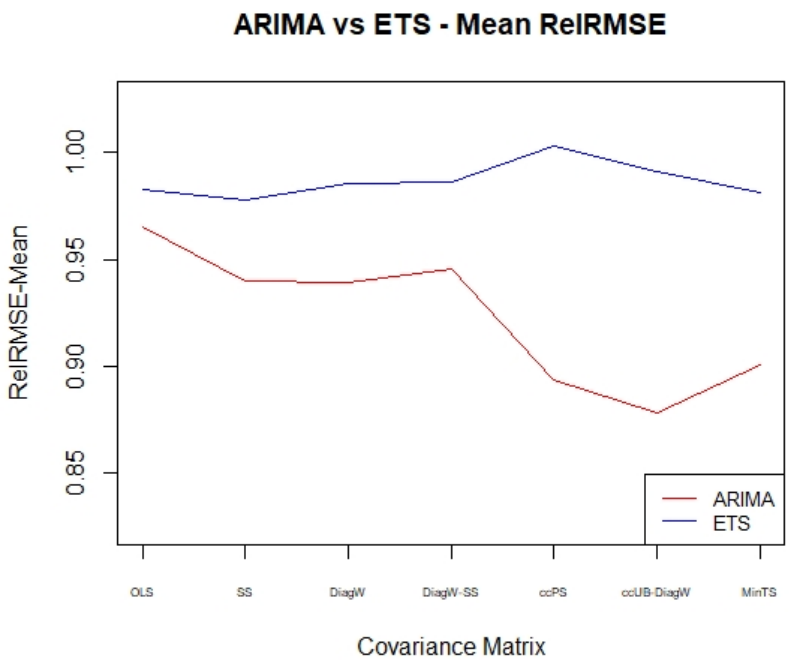
- Scandinavian unemployment data, consisting of Norway, Sweden, Denmark, and Finland (FRED Database)
- Grouped time series, can be structured from age, gender, and countries with 45 series and 312 points in time, monthly (Jan 1989 - Dec 2014)
- To forecast individually using automatic modelling process, both ARIMA and ETS
- Fixed in-sample rolling origin for 12-step ahead forecasts with 6 origins
- 19 sets of in-sample, ranging from 36, 48, to 252

Factors Affecting $RelRMSE$

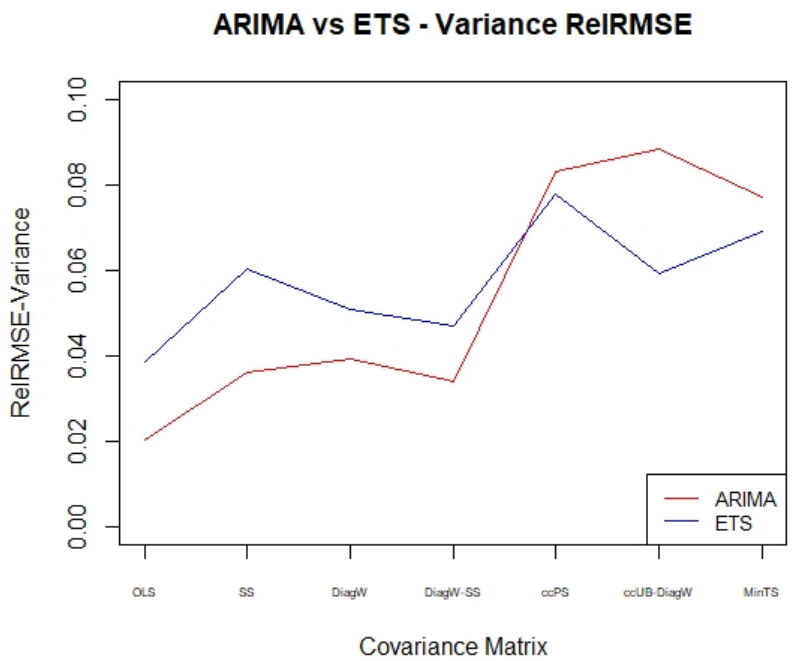


NB: Model in the legend presents the difference between ARIMA and ETS for this purpose only. In the analysis, it is denotes as RelSE and a continous one

$RelRMSE$ in Mean and Variance



RelRMSE-Mean



RelRMSE-Variance

Regression Models: Preliminary Findings

We hypothesise:

- More complex covariance matrices generate more accurate forecasts on average but more variable
- Comparative performance of forecasting models may be a moderating variable between between $RelRMSE$ and Q_2
- The effect of sample sizes to the performance seems small as it affects the performance on more complex covariance matrices only

Regression Models: Preliminary Findings

GAMLSS		
Variables	Mean	Variance
Intercept	1.2812	-66.3984
Q_2	0.0032	0.2497
Q_2^2		-0.0041
$RelSE$	-0.2853	133.3940
$RelSE^2$		0.0030
$Size$	0.0033	-5.9819**
$Size^2$		4.1529*
$Q_2 \times RelSE$	-0.0032	-0.1880
$(Q_2 \times RelSE)^2$		0.0030

NB: ** 0.01, * 0.05, . 0.1, $\star$ 1, otherwise are significant; the remaining results are in the appendix; SEP: skewed power exponential distribution. Generalised Additive Models with Location, Scale, and Shape (GAMLSS) is proposed by **?** and **?**

Revisiting Motivation

Caution: Results are valid for Scandinavian unemployment data only!

- How can we choose the covariance matrix in MinT Reconciliation to improve forecast accuracy?
  - When comparative performances show random walk performs better than any forecasting model, use simple covariance matrices (OLS or SS)
  - Otherwise, use MinT-Shrink, ccPS, or ccUB-DiagW
- Can we 'select' reconciliation weight estimation prior to the reconciliation?
  - Not necessarily with these findings. From these findings, we are trying to devise **a selection method in MinT Reconciliation** [Hopefully!]

Revisiting Motivation

- Does MinT-Shrink provide robust performances?
  - It is still the best choice and performs well but variable
- Do we have any alternative to estimate covariance matrices beside MinT-Shrink?
  - Yes, we do. ccPS and ccUB-DiagW provide similar performances to MinT-Shrink with fewer estimation

Thank you
Questions and feedback are welcome!

k.pritularga@lancaster.ac.uk

Marketing Analytics & Forecasting

Lancaster University
Management School

References I

A1. How is the forecast reconciliation done?

1. Prepare all series across hierarchy
2. Forecast all series according to the specific task, store the standard of errors
3. Store the unreconciled forecasts and the one-step ahead in-sample forecast errors, calculate RMSE
4. Use the in-sample forecast errors to estimate the covariance matrix i
5. Along with the unreconciled forecasts, the covariance matrix enters $\boldsymbol{P}$
6. Generate reconciled forecasts, calculate RMSE and take relative measures

A2. Residual Diagnostics: GAMLSS-SEP

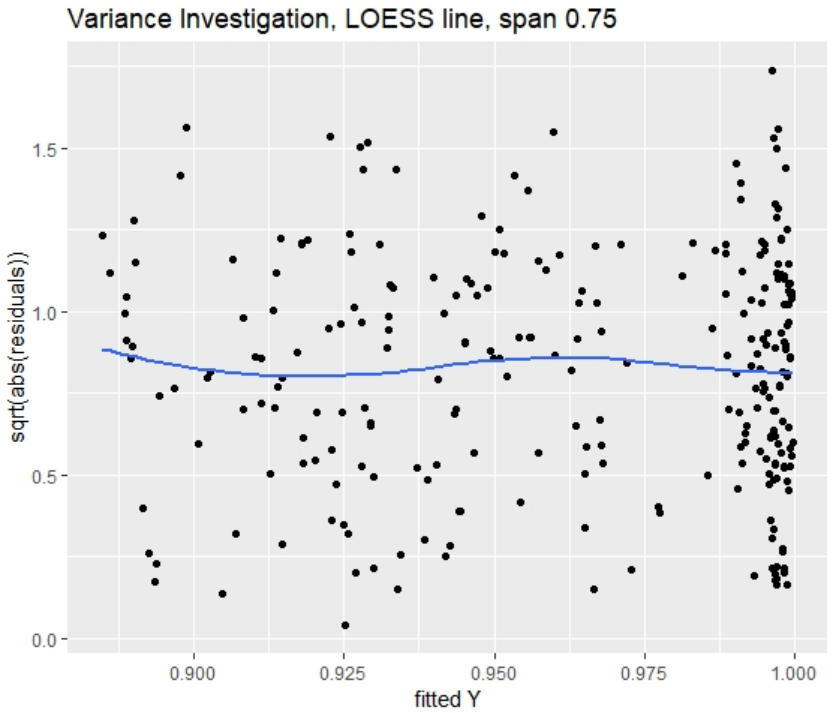


Figure: $\sqrt{|\varepsilon|}$ v.s. fitted values

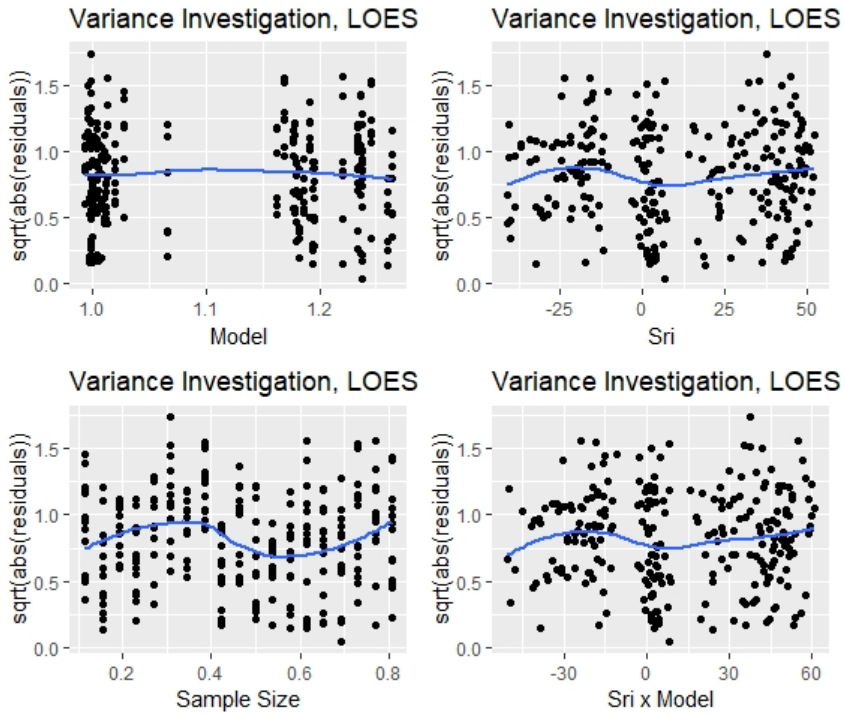


Figure: $\sqrt{|\varepsilon|}$ v.s. independent variables

A3. RelRMSE in Mean and Variance

	Mean		Standard Deviation	
Method	ARIMA	ETS	ARIMA	ETS
OLS	0.9655	0.9830	0.0203	0.0385
SS	0.9402	0.9779	0.0362	0.0602
DiagW	0.9391	0.9855	0.0391	0.0508
DiagW-SS	0.9456	0.9865	0.0339	0.0469
ccPS	0.8935	1.0031	0.0832	0.0777
ccUB-DiagW	0.9006	0.9814	0.0772	0.0691
MinTS	0.8785	0.9912	0.0882	0.0594

Table: Summary statistics of $RelRMSE$ based on the methods