

Getting Back on Track: Forecasting After Extreme Observations

Pål Boug Håvard Hungnes
and Takamitsu Kurita

Presentation at
IAAE Annual Conference,
Turin, June 2025

Outline

- 1 Introduction
- 2 Empirical example
- 3 Theoretical formulations
- 4 Monte Carlo experiments
- 5 Conclusions

- Extreme events (e.g., pandemics) generate transitory outliers in macroeconomic time series without affecting the long-run properties
- We compare the forecasting performance of three outlier treatments – NC, IO and AO corrections – within the CVAR framework
- Contributions:
 - **Empirical example:** Norwegian consumption and COVID-19
 - **Theoretical formulations:** Forecast bias and uncertainty
 - **Monte Carlo experiments:** Simulations with general settings
- AO corrections outperform IO and NC when macroeconomic time series rapidly return to their initial trajectories following periods of transitory extreme observations

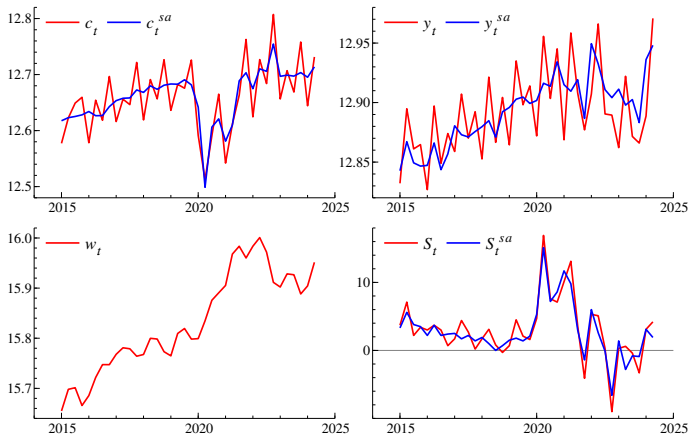
Empirical example

As a motivating empirical example, we consider a data set for the Norwegian consumption function.

- First, we give an overview of the data set with particular focus on time series behaviour during the pandemic.
- Second, we carry out a benchmark cointegration analysis prior to the pandemic based on a CVAR with $Y_t = (c_t, y_t, w_t)'$.
- Third, we estimate CVAR models with AO corrections, IO corrections and no outlier corrections during the pandemic.
- Lastly, we compare the forecasting performance of the three CVAR models after the pandemic.

Empirical example - overview of data

Figure: Consumption (c_t in logs), real disposable income (y_t in logs), real wealth (w_t in logs) and savings ratio (S_t in per cent)



Source: The databank of the model KVARTS at Statistics Norway.

Empirical example - benchmark cointegration analysis

- The benchmark cointegration analysis is based on a CVAR model of order $k = 6$ written as

$$\Delta Y_t = \alpha \beta' Y_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta Y_{t-i} + \mu + \varepsilon_t,$$

where $Y_t = (c_t, y_t, w_t)'$ and $t = 1983q3, \dots, 2019q4$.

Main findings:

- One cointegrating vector ($r = 1$)
- Homogeneity ($\beta_y + \beta_w = -1$) between c_t , y_t and w_t ,

$$\hat{\alpha}' = \begin{pmatrix} -0.238, 0.228, 0.131 \\ (0.085) \quad (0.071) \quad (0.119) \end{pmatrix} \text{ and } \hat{\beta}' = \begin{pmatrix} 1, -0.787, -0.213 \\ (-) \quad (0.027) \quad (0.027) \end{pmatrix}$$

Empirical example - three estimated CVAR models

$$\text{NC: } \Delta Y_t = \alpha \beta' Y_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta Y_{t-i} + \mu + \varepsilon_t$$

$$\text{IO: } \Delta Y_t = \alpha \beta' Y_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta Y_{t-i} + \phi D_t + \mu + \varepsilon_t$$

$$\text{AO: } \Delta(Y_t - \theta D_t) = \alpha \beta' (Y_{t-1} - \theta D_{t-1}) + \sum_{i=1}^{k-1} \Gamma_i \Delta(Y_{t-i} - \theta D_{t-i}) + \mu + \varepsilon_t$$

where

$$D_t = (1_{t=2020q1}, 1_{t=2020q2}, \dots, 1_{t=2022q1})',$$

$$\theta = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \theta_c, \quad \text{i.e., only } c \text{ is subject to AO corrections.}$$

Empirical example - three estimated CVAR models

We provide two sets of estimates of the three CVAR models where

- the last observation of the pandemic, 2022q1, represents the end of the estimation period, i.e., $t = 1983q3, \dots, 2022q1$ (first set)
- a "normal" observation after the pandemic, 2022q2, represents the end of the estimation period, i.e., $t = 1983q3, \dots, 2022q2$ (second set)

Empirical example - three estimated CVAR models

Main findings:

- the two sets of estimates of α and β are, overall, quite similar across both the CVAR models and the estimation periods
- the homogeneity restriction ($\beta_y + \beta_w = -1$) is not rejected in all cases
- not surprisingly, the first set of estimates of the IO model *is* exactly identical to those of the benchmark analysis
- the first set of estimates of the AO model is *not* exactly identical to those of the IO model (as only c_t is subject to additive outlier corrections)

Empirical example - forecasting performance

- Forecasting periods:
 - Without a post-extreme "normal" observation: 2022q2 – 2024q2
 - With a "normal" observation: 2022q3 – 2024q2
- RMSFEs of the three CVAR models

Model	RMSFE (without "normal")	RMSFE (with "normal")
NC	0.057	0.083
IO	0.069	0.096
AO	0.055	0.049

Empirical example - forecasting performance

The different forecasting properties are due to the varying influences of extreme observations on the parameter estimates and the conditioning values of c_t on the forecasts.

- Estimates of the AO model are effectively based on observations of c_t prior to the pandemic and observations of y_t and w_t including the pandemic.
- Estimates of the IO model and the NC model are, respectively, based on observations of c_t , y_t and w_t prior to the pandemic and including the pandemic.
- Forecasts from the AO model are conditioned on the predicted (counterfactual) values of c_t during the pandemic
- Forecasts from the IO model and the NC model are conditioned on the actual values of c_t during the pandemic.

Theoretical formulations

We consider one-step-ahead forecasts since, as pointed out by Clements and Hendry (1993) and Hendry (1986), such forecasts are usually sufficient for ranking forecasting performance of different models.

- First, we set up a general model for the data generation process (DGP).
- Then, we consider forecast bias and uncertainty of the AO, IO and NC models when the extreme observations are at the end of the estimation period.
- Lastly, we consider forecast bias and uncertainty of the three models when there is a "normal" observation prior to forecasting.

Theoretical formulations - DGP

We assume n additive outliers close to the end of the sample period:

$$Y_t = X_t + \theta D_t,$$

where $D_t = (1_{t=T-n+1}, \dots, 1_{t=T})'$ is a vector of indicator variables and $\theta = (\theta_{T-n+1}, \dots, \theta_T)$ is a matrix of corresponding coefficients.

The observations net of extreme observations follow a CVAR of order k :

$$\Delta X_t = \alpha \beta' X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu + \varepsilon_t$$

With $k = 1$:

$$Y_{T+1} = \mu + A_1 (Y_T - \theta_T) + \varepsilon_{T+1},$$

where $A_1 = I_p + \alpha \beta'$.

Theoretical formulations - comparison

The expected forecast errors are (when disregarding the estimation uncertainty of μ and A_1 (and Ω) to isolate the effects of the extreme observations):

$$\mathcal{E} \left[\hat{\varepsilon}_{T+1|T}^{AO}; \hat{\mu} = \mu, \hat{A}_1 = A_1 \right] = 0$$

$$\mathcal{E} \left[\hat{\varepsilon}_{T+1|T}^{IO}; \hat{\mu} = \mu, \hat{A}_1 = A_1 \right] = -A_1 \theta_T$$

$$\mathcal{E} \left[\hat{\varepsilon}_{T+1|T}^{NC}; \hat{\mu} = \mu, \hat{A}_1 = A_1 \right] = -A_1 \theta_T$$

- The forecasts with the AO model are unbiased
- The forecasts with the IO and NC models are biased depending on
 - The size of the extreme observations, θ_T
 - The persistence of the dynamics, A_1

Theoretical formulations - comparison

The forecast uncertainty are (measured by the squared forecast errors):

$$\mathcal{E} \left[\hat{\varepsilon}_{T+1|T}^{AO} \hat{\varepsilon}_{T+1|T}^{AO'}; \hat{\mu} = \mu, \hat{A}_1 = A_1, \hat{\Omega} = \Omega \right] = \Omega + \sum_{i=1}^n A_1^i \Omega A_1^{i'}$$

$$\mathcal{E} \left[\hat{\varepsilon}_{T+1|T}^{IO} \hat{\varepsilon}_{T+1|T}^{IO'}; \hat{\mu} = \mu, \hat{A}_1 = A_1, \hat{\Omega} = \Omega \right] = \Omega + A_1 \theta_T \theta_T' A_1'$$

$$\mathcal{E} \left[\hat{\varepsilon}_{T+1|T}^{NC} \hat{\varepsilon}_{T+1|T}^{NC'}; \hat{\mu} = \mu, \hat{A}_1 = A_1, \hat{\Omega} = \Omega \right] = \Omega + A_1 \theta_T \theta_T' A_1'$$

- The latter term of the forecast uncertainty of the IO and NC models is due to the non-estimation of θ_T .
- If $n = 1$, the forecast uncertainty of the AO model is the smallest, provided that the extreme observation is substantial, i.e.
 $A_1 \theta_T \theta_T' A_1' - A_1 \Omega A_1' = A_1 (\theta_T \theta_T' - \Omega) A_1' \succ 0 \Leftrightarrow \theta_T \theta_T' - \Omega \succ 0$ (when A_1 is non-singular).
- If n is large, the forecast uncertainty of the AO model can be the largest (when $\sum_{i=1}^n A_1^i \Omega A_1^{i'} - A_1 \theta_T \theta_T' A_1' \succ 0$) even though this is the correctly specified model.

Theoretical formulations - AO forecast with one "normal" obs

- We now show that the forecasting uncertainty of the AO model is reduced with a "normal" observation prior to forecasting assuming only one extreme observation ($n = 1$).
- When $k = 2$, the DGP implies that

$$Y_{T+2} = \mu + A_1 Y_{T+1} + A_2 (Y_T - \theta_T) + \varepsilon_{T+2},$$

where $A_1 = (I_p + \Gamma_1) + \alpha\beta'$ and $A_2 = -\Gamma_1$.

- For the AO model, we have

$$\begin{aligned}\hat{\varepsilon}_{T+2|T+1}^{AO} &= [\mu - \hat{\mu}] + [A_1 - \hat{A}_1] Y_{T+1} + [A_2 - \hat{A}_2] (Y_T - \theta_T) \\ &\quad - \hat{A}_2 [\theta_T - \hat{\theta}_T] + \varepsilon_{T+2},\end{aligned}$$

where the estimate of θ_T must be considered.

Theoretical formulations - AO forecast with one "normal" obs

First, we assume θ_T is estimated with observations including period T :

$$\mathcal{E} \left[\hat{\varepsilon}_{T+2|T+1}^{AO} \left(\hat{\theta}_{T|T} \right); \hat{\mu} = \mu, \hat{A}_1 = A_1, \hat{A}_2 = A_2 \right] = 0,$$

$$\mathcal{E} \left[\hat{\varepsilon}_{T+2|T+1}^{AO} \left(\hat{\theta}_{T|T} \right) \hat{\varepsilon}_{T+2|T+1}^{AO} \left(\hat{\theta}_{T|T} \right)' ; \dots, \hat{\Omega} = \Omega \right] = \Omega + A_2 \Omega A_2'.$$

Second, we assume θ_T is estimated with observations including period $T + 1$:

$$\mathcal{E} \left[\hat{\varepsilon}_{T+2|T+1}^{AO} \left(\hat{\theta}_{T|T+1} \right); \hat{\mu} = \mu, \hat{A}_1 = A_1, \hat{A}_2 = A_2 \right] = 0,$$

$$\mathcal{E} \left[\hat{\varepsilon}_{T+2|T+1}^{AO} \left(\hat{\theta}_{T|T+1} \right) \hat{\varepsilon}_{T+2|T+1}^{AO} \left(\hat{\theta}_{T|T+1} \right)' ; \dots \right] = \Omega + A_2 \left[\Omega^{-1} + A_1' \Omega^{-1} A_1 \right]^{-1} A_2'.$$

- If both A_1 and A_2 are non-singular, the latter forecast uncertainty is smaller than the former since $A_2 \Omega A_2' - A_2 \left[\Omega^{-1} + A_1' \Omega^{-1} A_1 \right]^{-1} A_2' \succ 0$. (proof in paper)
- For comparison: $\mathcal{E} \left[\hat{\varepsilon}_{T+2|T+1}^{IO}; \dots \right] = \mathcal{E} \left[\hat{\varepsilon}_{T+2|T+1}^{NC}; \dots \right] = -A_2 \theta_T$.

$$\text{Var} \left[\hat{\varepsilon}_{T+2|T+1}^{IO}; \dots \right] = \text{Var} \left[\hat{\varepsilon}_{T+2|T+1}^{NC}; \dots \right] = \Omega + A_2 \theta_T \theta_T' A_2'.$$

Theoretical formulations - summary of findings

- The AO model provides unbiased forecasts.
- The forecast uncertainty of the AO model increases with the number of extreme observations.
- The forecast uncertainty of the AO model decreases with a "normal" observation at the end of the estimation period.

Monte Carlo experiments

We conduct Monte Carlo experiments under general settings to support the theoretical arguments.

- First, we describe a DGP and a class of three CVAR models.
- Then, we conduct a benchmark comparative study with regard to forecasting performance of the three models.
- Lastly, we compare forecasting performance in the presence of unmodelled level shifts (shown in the paper).

Monte Carlo experiments - DGP

The DGP is formulated as a CVAR process $Y_t = (Y_{1,t}, Y_{2,t}, Y_{3,t})'$ for $t = 1, \dots, T + j + h$, which comprises a combination of

- ① a dynamic process formulated as $X_t = (X_{1,t}, X_{2,t}, X_{3,t})'$ s.t. $r = 1$
- ② a cluster of 9 additive outliers given as $D_t = (1_{t=T-8}, \dots, 1_{t=T})'$

The values for $T + j$ and h vary in experiments, where

- $T + j$ represents the total number of observations for estimation with the default $T = 100$ and $j = 0$
- $h = 8$ and $h = 20$ (shown in the paper) correspond to forecasting horizons

The DGP is a simplified version of the empirical example and the triplet $(Y_{1,t}, Y_{2,t}, Y_{3,t})'$ corresponds to $(c_t, y_t, w_t)'$.

Monte Carlo experiments - DGP

The specification of the DGP is thus

$$Y_t = X_t + \theta D_t$$

and

$$\Delta X_t = \alpha \beta' X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu + \varepsilon_t,$$

where $\theta = (1, 0, 0)' \theta_1$ with

$$\theta_1 = (-0.10 \quad -0.12 \quad -0.12 \quad -0.12 \quad -0.10 \quad -0.10 \quad -\vartheta \quad -\vartheta \quad -\vartheta)$$

for $\vartheta = 0.06$ or $\vartheta = 0.08$.

Monte Carlo experiments - DGP

$$\alpha = \begin{pmatrix} -0.2 \\ 0.2 \\ 0.0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1.0 \\ -0.8 \\ -0.2 \end{pmatrix}, \quad \mu = \begin{pmatrix} -0.136 \\ 0.150 \\ 0.006 \end{pmatrix},$$

$$(\Gamma_1, \Gamma_2, \Gamma_3) = \begin{pmatrix} 0.2 & -0.2 & 0.0 & 0.1 & 0.1 & 0.1 & 0.1 & 0.0 & 0.0 \\ 0.0 & -0.4 & 0.0 & 0.0 & 0.1 & 0.0 & 0.0 & 0.2 & 0.0 \\ 0.0 & 0.0 & 0.3 & 0.0 & -0.2 & 0.3 & 0.0 & 0.0 & -0.1 \end{pmatrix}$$

for $\max k = 4$.

Monte Carlo experiments - DGP

The lag length k is assigned 2 or 4, so that only the corresponding parameter matrices among $(\Gamma_1, \Gamma_2, \Gamma_3)$ are effective in the DGP.

The error process ε_t is a trivariate *i.i.d.* normal process, $N(0, \Omega)$.

The starting values of X_t mimic rounded-off values of conceivable real-life data for the logs of macroeconomic time series, $(c_t, y_t, w_t)'$

The number of Monte Carlo replications in each experiment is 10,000.

Monte Carlo experiments - DGP

We conduct a range of Monte Carlo experiments to compare forecasting performance (Monte Carlo averages of RMSFEs) of three CVAR models.

The three CVAR models are:

- 1 AO model corrected with 9 additive outlier (dummy) variables
- 2 IO model corrected with 9 innovative outlier (dummy) variables
- 3 NC model not corrected with any outlier (dummy) variables

The AO model is judged to be the correctly specified formulation in that it encompasses the underlying DGP.

Monte Carlo experiments - benchmark study

Table: Forecasting performance of three CVAR models — without a “normal” observation ($j = 0$), short forecasting horizon ($h = 8$) and small additive outliers at the end of the sample period ($\vartheta = 0.06$)

(i) short lag length ($k = 2$)				(ii) long lag length ($k = 4$)			
	NC	IO	AO		NC	IO	AO
$Y_{1,t}$	0.0661	0.0405	0.0335	$Y_{1,t}$	0.0673	0.0472	0.0405
$Y_{2,t}$	0.0225	0.0289	0.0227	$Y_{2,t}$	0.0267	0.0356	0.0270
$Y_{3,t}$	0.0378	0.0424	0.0372	$Y_{3,t}$	0.0469	0.0598	0.0456

Notes: Figures are root mean squared forecast errors (RMSFEs). $T = 100$, $j = 0$, $h = 8$ and $\vartheta = 0.06$. Based on 10,000 replications.

Monte Carlo experiments - benchmark study

Table: Forecasting performance of three CVAR models — without a “normal” observation ($j = 0$), short forecasting horizon ($h = 8$) and large additive outlier at the end of the sample period ($\vartheta = 0.08$)

(i) short lag length ($k = 2$)				(ii) long lag length ($k = 4$)			
	NC	IO	AO		NC	IO	AO
$Y_{1,t}$	0.0813	0.0477	0.0335	$Y_{1,t}$	0.0822	0.0549	0.0406
$Y_{2,t}$	0.0228	0.0328	0.0227	$Y_{2,t}$	0.0269	0.0402	0.0270
$Y_{3,t}$	0.0380	0.0449	0.0372	$Y_{3,t}$	0.0470	0.0643	0.0456

Notes: Figures are root mean squared forecast errors (RMSFEs). $T = 100$, $j = 0$, $h = 8$ and $\vartheta = 0.08$. Based on 10,000 replications.

Monte Carlo experiments - benchmark study

Table: Forecasting performance of three CVAR models — with a “normal” observation ($j = 1$), short forecasting horizon ($h = 8$) and large additive outlier at the end of the sample period ($\vartheta = 0.08$)

(i) short lag length ($k = 2$)				(ii) long lag length ($k = 4$)			
	NC	IO	AO		NC	IO	AO
$Y_{1,t}$	0.0344	0.0361	0.0268	$Y_{1,t}$	0.0413	0.0571	0.0324
$Y_{2,t}$	0.0259	0.0243	0.0225	$Y_{2,t}$	0.0325	0.0372	0.0265
$Y_{3,t}$	0.0403	0.0401	0.0376	$Y_{3,t}$	0.0502	0.0551	0.0456

Notes: Figures are root mean squared forecast errors (RMSFEs). $T = 100$, $j = 1$, $h = 8$ and $\vartheta = 0.08$. Based on 10,000 replications.

- We have illustrated the usefulness of AO corrections in forecasting Norwegian consumption after COVID-19
- We have shown theoretically the forecast biases using IO and NC
- We have applied Monte Carlo simulations with general settings supporting the theoretical findings
- AO corrections outperform IO and NC when macroeconomic time series rapidly return to their initial trajectories following periods of transitory extreme observations

- Clements, M. P. and Hendry, D. F. (1993). On the limitations of comparing mean square forecast errors. *Journal of Forecasting*, 12(8):617–637.
- Hendry, D. F. (1986). The role of prediction in evaluating econometric models. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 407(1832):25–34.

Thank you!!