Statistisk sentralbyrå
Statistics Norway

Getting Back on Track: Forecasting After Extreme Observations

Pål Boug, Håvard Hungnes and Takamitsu Kurita

TALL SOM FORTELLER
DISCUSSION PAPERS
1018

Discussion Papers: comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

The Discussion Papers series presents results from ongoing research projects and other research and analysis by SSB staff. The views and conclusions in this document are those of the authors.

Published: December 2024

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:

https://www.ssb.no/discussion-papers
http://ideas.repec.org/s/ssb/dispap.html

ISSN 1892-753X (electronic)

Abstract

This paper examines the forecast accuracy of cointegrated vector autoregressive models when confronted with extreme observations at the end of the sample period. It focuses on comparing two outlier correction methods, additive outliers and innovational outliers, within a forecasting framework for macroeconomic variables. Drawing on data from the COVID-19 pandemic, the study empirically demonstrates that cointegrated vector autoregressive models incorporating additive outlier corrections outperform both those with innovational outlier corrections and no outlier corrections in forecasting post-pandemic household consumption. Theoretical analysis and Monte Carlo simulations further support these findings, showing that additive outlier adjustments are particularly effective when macroeconomic variables rapidly return to their initial trajectories following short-lived extreme observations, as in the case of pandemics. These results carry important implications for macroeconomic forecasting, emphasising the usefulness of additive outlier corrections in enhancing forecasts after periods of transient extreme observations.

Keywords: Extreme observations, additive outliers, innovational outliers, cointegrated vector autoregressive models, forecasting.

JEL classification: C32, C53, E21, E27.

Acknowledgements: We are grateful to seminar participants at Statistics Norway and the Centre for Applied and Theoretical Econometrics at BI Norwegian Business School, particularly Ragnar Nymoen and Genaro Sucarrat, for their helpful discussions. We also thank Thomas von Brasch and Terje Skjerpen for their valuable comments and suggestions on earlier drafts. Pål Boug and Håvard Hungnes gratefully acknowledge financial support from the Research Council of Norway (grant no. 326419). Takamitsu Kurita gratefully acknowledges financial support from JSPS KAKENHI 18K01600 and 24K04877. The usual disclaimer applies.

Address: Pål Boug, Statistics Norway, Research Department. E-mail: pal.boug@ssb.no;
Håvard Hungnes, Statistics Norway, Research Department. E-mail: havard.hungnes@ssb.no;
Takamitsu Kurita, Kyoto Sangyo University, Kyoto, Japan. E-mail: tkurita@cc.kyoto-su.ac.jp.

Sammendrag

Denne studien sammenlikner prognoseegenskaper til kointegrerende vektor autoregressive modeller (CVAR-modeller) som håndterer ekstreme observasjoner gjennom enten additive korreksjoner for uteliggere (AO-korreksjoner) eller innovative korreksjoner for uteliggere (IO-korreksjoner). Sammenlikningen er gjennomført empirisk, teoretisk og ved hjelp av Monte Carlo-simuleringer.

Som et motiverende empirisk eksempel anvender vi et oppdatert datasett for den norske konsumfunksjonen, som viser et betydelig strukturelt brudd under COVID-19-pandemien. Basert på målene RMSFE («root mean squared forecast error») og MAPE («mean absolute percentage error») viser vi at en CVAR-modell for husholdningenes konsum, inntekt og formue med AO-korreksjoner for de ekstreme observasjonene under pandemien gir bedre prognoser enn CVAR-modeller med IO-korreksjoner og uten korreksjoner. Den prognostiske overlegenheten til modellen med AO-korreksjoner er særlig fremtredende når en «normal» observasjon er tilgjengelig i slutten av estimeringsperioden.

Med utgangspunkt i forenklede antakelser, som en kort tidsforsinkelse, et begrenset antall ekstreme observasjoner og fravær av strukturelle brudd i prognoseperioden, utvikler vi et teoretisk rammeverk som forklarer hvorfor CVAR-modellen med AO-korreksjoner gir bedre prognoser når makroøkonomiske variabler raskt vender tilbake til sine normale forløp etter midlertidige ekstreme observasjoner. Vi viser spesielt at CVAR-modeller med IO-korreksjoner og uten korreksjoner ofte gir skjeve prognoser, mens modellen med AO-korreksjoner leverer forventningsrette prognoser med potensielt lavere usikkerhet. Videre viser vi at prognoseusikkerheten kan reduseres ytterligere når modellen tar høyde for en «normal» observasjon som følger rett etter de ekstreme observasjonene. Våre Monte Carlo-simuleringer, utført under mer generelle antakelser, inkludert en lengre tidsforsinkelse, et større antall ekstreme observasjoner og tilstedeværelse av strukturelle brudd i prognoseperioden, støtter den teoretiske rangeringen av prognoseegenskapene til de tre CVAR-modellene.

Basert på våre funn argumenterer vi for at CVAR-modellen med AO-korreksjoner er særlig godt egnet til å håndtere spesielle hendelser som involverer ekstreme observasjoner. Dette gjør modellen til et foretrukket verktøy for å prognostisere sentrale makroøkonomiske variabler, som husholdningskonsum, ved gjennomføring av økonomisk politikk.

1 Introduction

Many macroeconomic time series exhibit extreme observations during events such as financial crises, wars, natural disasters or pandemics. These observations typically manifest as structural breaks in empirical models and may lead to poor forecasting performance depending on the origins and characteristics of the breaks. Therefore, a careful analysis of the forecasting implications of alternative methods for handling such extreme observations is crucial in time series modelling, see *e.g.* Bauwens *et al.* (2015) and Castle *et al.* (2015, 2016, 2024). This paper provides new insights into this issue through an empirical example from the COVID-19 pandemic, alongside theoretical formulations and Monte Carlo simulations of forecasts within the multivariate context of the cointegrated vector autoregressive (CVAR) model. Our primary focus is to compare the forecasting performance of CVAR models that incorporate additive outlier (AO) and innovational outlier (IO) corrections as alternative methods for handling extreme observations in-sample. The main contribution of our study is that the CVAR model with AO corrections delivers the most accurate forecasts when macroeconomic variables rapidly return to their initial trajectories following transient extreme observations.

More than fifty years ago, Fox (1972) distinguished between IOs and AOs in analysing the effects of extreme observations on time series behaviour. An IO corresponds to a situation where a single extreme shock affects the innovation term of a time series, whereas an AO corresponds to a situation where a single extreme event affects only a specific observation. As pointed out by Fox (1972), an IO impacts not only the observation at the time of the shock but also influences subsequent observations through the dynamic structure of the time series. In contrast, an AO removes the extreme observation independently of the series' dynamics. Whereas IOs may be more suitable in cases such as financial crises, wars or natural disasters, where the economy may take considerable time to recover, AOs may be more appropriate in instances of pandemics, where the economy may quickly return to its pre-pandemic trajectory. However, the question of which type of outlier correction is most effective in a particular forecasting exercise remains open.

The last four decades of research on time series forecasting, including the roles of IOs and AOs, are reviewed in De Gooijer and Hyndman (2006) and Petropoulos *et al.* (2022). Numerous studies have explored the detection and estimation of different types of outliers and their influence on forecasting performance within the context of ARIMA-models (see *e.g.* Ljung, 1993; Ledolter, 1989; Findley *et al.*, 1998;

Battaglia *et al.*, 2020; Laome *et al.*, 2021) and GARCH-models (see *e.g.* Franses and Ghijsels, 1999; Catalán and Trívez, 2007; Hotta and Trucíos, 2018; Akpan *et al.*, 2019; Patrocinio *et al.*, 2024). Several studies have also examined the effects of IOs and AOs on inference and estimation within the context of CVAR models (see *e.g.* Franses and Haldrup, 1994; Franses and Lucas, 1998; Johansen *et al.*, 2000; Lütkepohl *et al.*, 2004; Nielsen, 2004, 2008; Hungnes, 2010; Johansen and Nielsen, 2018; Kurita and Nielsen, 2019). Furthermore, a number of studies have explored the impacts of various methods for addressing outliers and structural breaks on forecasting performance in the context of CVAR models (see *e.g.* Clements and Hendry, 1996, 1998, 1999; Castle *et al.*, 2015, 2016, 2024).

Surprisingly, however, alternative AO corrections for handling extreme observations in forecasting studies have rarely been applied in CVAR models. To the best of our knowledge, the only study that addresses AO corrections in the context of forecasting with a CVAR model is that of Ahumada and Garegnani (2012), who conduct an empirical forecasting competition between several models of money demand in Argentina. To date, no studies seem to have developed theoretical frameworks or conducted Monte Carlo simulations to assess the forecasting ability of CVAR models that incorporate AO corrections for extreme observations.

Our study is the first to compare the forecasting performance of CVAR models that incorporate both AOs and IOs as alternative correction methods for extreme observations, using empirical analysis, theoretical examination and Monte Carlo simulations. As a motivating empirical example, we build on Boug *et al.* (2021) and consider a refreshed data set for the Norwegian consumption function, which exhibits a substantial structural break during the COVID-19 pandemic. We demonstrate, using the root mean squared forecast error (RMSFE) and mean absolute percentage error (MAPE) metrics, that a CVAR model for household consumption, income and wealth with AO corrections for the extreme observations during the pandemic outperforms both a CVAR model with IO corrections and a CVAR model without any corrections when forecasting after the pandemic. The forecasting superiority of the model with AO corrections is particularly pronounced when a "normal" observation is available for estimation at the end of the in-sample period. Using simplifying assumptions of a short lag length, a limited number of extreme observations and no out-of-sample structural breaks, we develop a theoretical framework that explains the superior forecast accuracy of the CVAR model with AO corrections when macroeconomic variables rapidly return to their initial trajectories following transient extreme observations. Specifically, we show that

while CVAR models with IO and no outlier corrections produce biased forecasts, the model with AO corrections provides unbiased forecasts with potentially lower forecast uncertainty. Moreover, we demonstrate that forecast uncertainty can be further reduced when the model conditions on a "normal" observation following the extreme observations. Our Monte Carlo simulations, conducted under more general settings, including longer lag lengths, a higher number of extreme observations and structural breaks during the forecasting period, reinforce the theoretical ranking of the forecast accuracy of the three CVAR models.

Based on our findings, we argue that the CVAR model with AO corrections for special events involving extreme observations may be preferable when forecasting important macroeconomic variables such as household consumption. The insights gained from this model can significantly enhance the precision of fiscal and monetary policy decisions, enabling policymakers to better manage economic growth and business cycle fluctuations.

The remainder of the paper is organised as follows: Section 2 presents the findings from the empirical example. Section 3 develops the theoretical framework motivated by these findings. Section 4 undertakes the Monte Carlo simulations with generalised settings to support the theoretical findings. Section 5 offers some conclusions.

2 A motivating empirical example

As a motivating empirical example, we build on Boug *et al.* (2021) and consider a refreshed data set, including the extreme observations during the COVID-19 pandemic, for the Norwegian consumption function. First, we give an overview of the data set with a particular focus on time series behaviour during the COVID-19 pandemic. Second, we carry out a benchmark cointegration analysis based on a trivariate CVAR model for household consumption, income and wealth. Third, we estimate CVAR models with additive outlier corrections, innovational outlier corrections and no outlier corrections for the extreme observations during the pandemic. Finally, we compare the forecasting performance of the three estimated CVAR models using available observations after the pandemic.[1]

[1]The econometric modelling in this section was carried out with PcGive 15 and Ox 8, see Doornik and Hendry (2018) and Doornik (2013). The data set applied and the Ox code written to estimate the CVAR model with additive outlier corrections are available from the authors upon request.

2.1 Overview of data

Throughout the analysis, following Eitrheim *et al.* (2002), Jansen (2013) and Boug *et al.* (2021), we consider the data set $Y_t = (c_t, y_t, w_t)'$, where c_t is the log of household real consumption excluding housing, y_t is the log of household real disposable income excluding equity dividends and w_t is the log of household real net wealth.[2] All three time series are quarterly, seasonally unadjusted, covering the effective sample period $1983q3 - 2024q2$ for estimation and forecasting purposes.

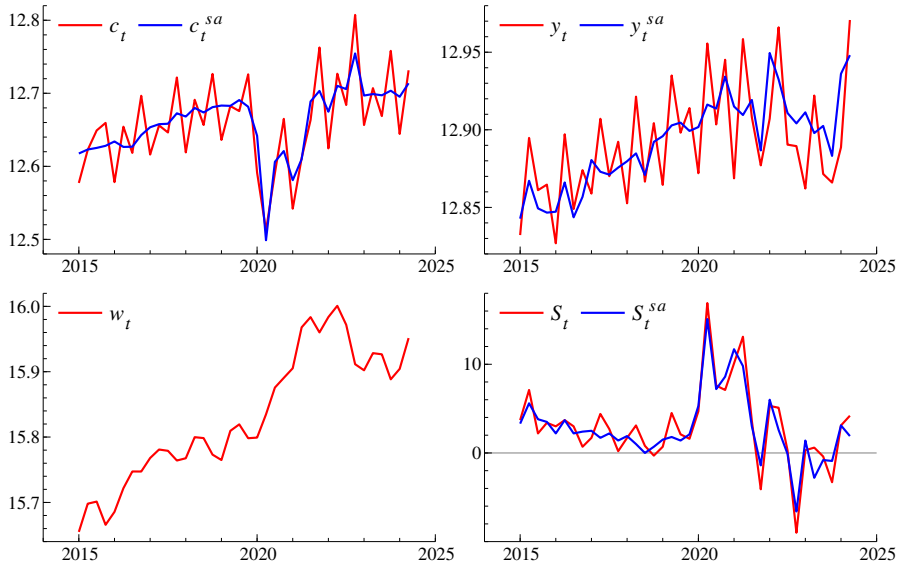
Whereas Boug *et al.* (2021) present a detailed data analysis of consumption, income and wealth for the period up to 2016 with a significant structural break identified around the time of the global financial crisis, we pay attention to influential observations related to major events during the COVID-19 pandemic. Figure 1 shows the seasonally unadjusted time series c_t, y_t and w_t alongside the seasonally unadjusted time series S_t, the household savings ratio excluding equity dividends, for the sub-sample period $2015q1 - 2024q2$. For later reference, Figure 1 also shows the seasonally adjusted time series c_t^{sa}, y_t^{sa} and S_t^{sa}, see Appendix A for details about data definitions and sources.

The pandemic in Norway, which began early in 2020, involved several waves of infections and various phases of infection control measures and restrictions. Although the strictest lockdowns and measures were gradually eased during 2021, the pandemic continued to affect Norway at the beginning of 2022, with the emergence of the Omicron variant necessitating further changes in infection control measures. In Norway, most of the measures were lifted in February 2022, marking a transition to a more normal everyday life.

During the pandemic in Norway, household consumption and savings underwent notable changes. For example, at the onset of the pandemic, when society shut down and spending opportunities were limited, savings rose significantly. Since consumption declined by almost 20 per cent during the first half of 2020, while income increased steadily due to various factors, *inter alia*, social security schemes, the savings ratio reached a record-high level of slightly more than 15 per cent in the same period. The consumption level more than recovered to its pre-pandemic level in $2021q4$, but dropped sharply again in $2022q1$ alongside the infection control measures associated with the Omicron variant. Then, in $2022q2$, when most of the measures were no longer in effect, the consumption level rebounded more or less

[2] Unlike Jansen (2013) and Boug *et al.* (2021), we disregard the real after-tax interest rates faced by households from Y_t to simplify matters without loss of establishing a well-specified CVAR model.

Figure 1: Real consumption (c_t and c_t^{sa}), real disposable income (y_t and y_t^{sa}), real net wealth (w_t) and savings ratio (S_t and S_t^{sa})



Notes: Sub-sample period: $2015q1-2024q2$. The time series c_t, c_t^{sa}, y_t, y_t^{sa} and w_t are measured in logs, the time series S_t and S_t^{sa} are measured in per cent and the superscript sa denotes seasonally adjusted time series. Sources: The databank of the KVARTS model at Statistics Norway and the quarterly national and non-financial sector accounts in the Statbank of Statistics Norway.

to its level in $2019q4$, whereas income increased during the first half of 2022. As a result, the savings ratio went down from its record-high level by more than 10 percentage points during the period of the pandemic. Household wealth, meanwhile, increased substantially during the years of the pandemic, driven by historically low interest rates and steady income growth. The time series for consumption, income and wealth also reveal that their growth paths before the pandemic were higher than those after the pandemic.

Against this backdrop, we define the period from $2020q1$ to $2022q1$ as a series of special events during the pandemic that influenced the time series behaviour of Norwegian consumption in an extreme way. These nine quarters of influential observations form the basis for the estimation of the CVAR models with either additive outlier or innovational outlier corrections.

2.2 Benchmark cointegration analysis

Our benchmark cointegration analysis, conducted over the effective sample period preceding the pandemic $t = 1983q3, \ldots, 2019q4$, follows Boug *et al.* (2021) and is

based on a p-dimensional CVAR model of order $k = 6$ for the $p = 3$ variables c_t, y_t and w_t, augmented with indicators capturing the structural break in the long-run relationship around the time of the global financial crisis. The model after the global financial crisis is specified as

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

where the parameter μ is a p-dimensional vector of constants, the parameters α and β are $p \times r$ matrices for $r \leq p$, thus forming the composite parameter $\Pi = \alpha\beta'$, the autoregressive parameters Γ_i for $i = 1, \ldots, k-1$ are $p \times p$ matrices and the independent Gaussian innovations ε_t have expectation zero and a $p \times p$ positive definite variance-covariance matrix Ω.[3] For later reference, we assume that the following assumption is fulfilled, see Johansen (1996, Theorem 4.2):

Assumption 1

1. *The characteristic polynomial associated with (1),* $A(z) = I_p - \sum_{i=1}^{k} A_k z^k = (1-z)I_p - \Pi z - \sum_{i=1}^{k-1} \Gamma_i (1-z) z^i$, *where* I_p *denotes the* $p \times p$ *identity matrix, satisfies the condition that if* $det[A(z)] = 0$, *then either* $|z| > 1$ *or* $z = 1$.

2. $\text{Rank}\,(\Pi) = r \leq p$.

3. *If* $r < p$, *then* $\det(\alpha'_\perp \Psi \beta_\perp) \neq 0$, *where* $\Psi = I_p - \sum_{i=1}^{k-1} \Gamma_j$ *and* $\alpha_\perp$ *and* $\beta_\perp$ *are the orthogonal complements to* α *and* β, *respectively.*

Assumption 1-1 ensures that Y_t is not an explosive process. Combined with Assumption 1-2, it ensures that there exist $p \times r$ matrices α and β of rank r such that $\Pi = \alpha\beta'$, where we here consider $p = 3$. If also Assumption 1-3 is fulfilled, it ensures that $\Delta Y_t - \mathcal{E}[\Delta Y_t]$ and $\beta' Y_t - \mathcal{E}[\beta' Y_t]$ can be given initial distributions such that they become $I(0)$.

Both the Akaike information criterion and likelihood ratio tests of sequential model reduction support that the underlying VAR model should include six lags. A battery of diagnostic tests also indicates no serious departures from white noise errors with $k = 6$. Using the trace test statistic for rank determination and critical values derived from the response surface estimated in Johansen *et al.* (2000),[4] we

[3] The indicators for the structural break around the time of the global financial crisis, which is not the main focus of this paper, are, along with centred seasonal dummies, suppressed in (1) for ease of exposition. See Appendix B for details.

[4] See Table 4 and equations 3.12 and 3.13 in Johansen *et al.* (2000).

find support for the hypothesis of a single cointegrating vector among c_t, y_t and w_t at the 5 per cent significance level.

After conducting preliminary tests of restrictions under the assumption of $r = 1$ on $\Pi = \alpha\beta'$, where $\alpha' = (\alpha_c, \alpha_y, \alpha_w)$ and $\beta' = (\beta_c, \beta_y, \beta_w)$, we arrive at a long-run structure where the hypothesis of homogeneity between consumption, income and wealth ($\beta_y + \beta_w = 1$) is not rejected by a likelihood ratio test statistic of $\chi^2(1) = 0.79$ (p-value = 0.37). The restricted estimates of the parameters of interest, α and β, then become

$$(2) \qquad \hat{\alpha}' = \underset{}{\left(\underset{(0.085)}{-0.238}, \underset{(0.071)}{0.228}, \underset{(0.119)}{0.131}\right)} \quad \text{and} \quad \hat{\beta}' = \left(\underset{(-)}{1}, \underset{(0.027)}{-0.787}, \underset{(0.027)}{-0.213}\right),$$

where the figures in parentheses are standard errors. A comparison with the findings of Boug *et al.* (2021) shows that the estimated long-run coefficients for income and wealth are nearly reproduced when extending the sample period to $2019q4$. Having established a benchmark long-run structure for the Norwegian consumption function using the refreshed data set, we proceed to the estimation of CVAR models with and without outlier corrections for the extreme observations during the pandemic.

2.3 CVAR models with and without outlier corrections

As pointed out above, we estimate a CVAR model with additive outlier corrections (AO model), a CVAR model with innovational outlier corrections (IO model) and a CVAR model without any outlier corrections (NC model, short for no-correction model), all of which are subject to the cointegrating rank $r = 1$ and the restriction of homogeneity between c_t, y_t and w_t. To account for the aforementioned nine quarters of influential observations during the pandemic in the estimation of the former two models, we define the cluster of impulse dummies

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

where $1_{t=s}$ denotes an indicator function taking the value of one when $t = s$ and zero otherwise. The unusual consumption patterns observed during the pandemic, contrasted with the relatively stable trajectories of income and wealth, suggest that only consumption requires outlier corrections in the AO model. Thus, following the

notation in Nielsen (2008), the AO model is specified as

$$\Delta(Y_t - \theta D_t) = \alpha\beta'(Y_{t-1} - \theta D_{t-1}) + \sum_{i=1}^{k-1} \Gamma_i \Delta(Y_{t-i} - \theta D_{t-i}) + \mu + \varepsilon_t, \tag{3}$$

where the CVAR model in (1) is augmented with D_t as additive dummies and the 3×9 parameter matrix θ. This parameter matrix imposes zero elements in its second and third rows through $\theta = (1, 0, 0)' \theta_c$, where θ_c is a 9-dimensional vector corresponding to the extreme observations of c_t. Since (3) imposes non-linear restrictions on the parameters, estimation relies on the algorithm described in Nielsen (2004). Essentially, the additive dummies in (3) mitigate the impact of extreme observations on consumption during the pandemic by replacing the observed extreme observations with their model-based predictions. Conversely, the IO model in this context is specified as

$$\Delta Y_t = \alpha\beta' Y_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta Y_{t-i} + \phi D_t + \mu + \varepsilon_t, \tag{4}$$

where the CVAR model in (1) is augmented with D_t as unrestricted innovational dummies and the 3×9 parameter matrix ϕ. This parameter matrix imposes no restrictions as there is no direct correspondence between its elements and the variables in Y_t. Unlike the additive dummies, the innovational dummies in (4) do not remove the impact of extreme observations on consumption (and income and wealth) during the pandemic. Instead, they act as indicators of shocks to ε_t, with their effects dissipating over time through the dynamics of the model.

Before presenting estimates of the CVAR models in (1), (3) and (4), we repeat from Section 1 that using a "normal" observation after the pandemic for estimation is likely to enhance the forecasting performance of the AO model relative to the IO and NC models. To explore this issue in the subsequent forecasting analysis, we provide two sets of estimates of the three CVAR models. The first and second sets assume, respectively, that the last extreme observation of the pandemic, $2022q1$, and a "normal" observation after the pandemic, $2022q2$, represent the end of the estimation period. These assumptions are justified by the fact that consumption was back on track with the pre-pandemic trajectory by $2022q2$, as shown in Figure 1.

Tables 1 and 2 report the first and second sets of estimates of α and β of the three CVAR models along with the corresponding $\chi^2(1)$ test statistics for the

Table 1: Estimates of three CVAR models — without a "normal" observation prior to forecasting

AO model

$$\hat{\alpha}' = \underset{(0.081)}{-0.219}, \underset{(0.073)}{0.276}, \underset{(0.114)}{0.106}\Big), \hat{\beta}' = \Big(\underset{(-)}{1}, \underset{(0.025)}{-0.788}, \underset{(0.025)}{-0.212}\Big), \chi^2(1) = 0.11[0.74]$$

IO model

$$\hat{\alpha}' = \underset{(0.085)}{-0.238}, \underset{(0.071)}{0.228}, \underset{(0.119)}{0.131}\Big), \hat{\beta}' = \Big(\underset{(-)}{1}, \underset{(0.027)}{-0.787}, \underset{(0.027)}{-0.213}\Big), \chi^2(1) = 0.84[0.36]$$

NC model

$$\hat{\alpha}' = \underset{(0.073)}{-0.265}, \underset{(0.051)}{0.232}, \underset{(0.077)}{0.031}\Big), \hat{\beta}' = \Big(\underset{(-)}{1}, \underset{(0.033)}{-0.801}, \underset{(0.033)}{-0.199}\Big), \chi^2(1) = 2.01[0.16]$$

Notes: Estimation period: $t = 1983q3, \ldots, 2022q1$. AO model, IO model and NC model denote the CVAR models in (3), (4) and (1), respectively, figures in parentheses are standard errors and $\chi^2(1)$ denotes the likelihood ratio test statistic for the hypothesis $\beta_y + \beta_w = 1$ with the p-value in square brackets.

hypothesis $\beta_y + \beta_w = 1$. Overall, the two sets of estimates of the parameters of interest, irrespective of which CVAR model is estimated, are quite similar to those from the benchmark cointegration analysis. Not surprisingly, the first set of estimates of the IO model is identical to those in (2), since carrying out innovational outlier corrections are equivalent to excluding the extreme observations during the pandemic from the estimation period. However, because only c_t is subject to additive outlier corrections during the pandemic, the first set of estimates of the AO model differs slightly from those of the IO model, which would otherwise not be the case if c_t, y_t and w_t were all subject to additive outlier corrections. Nevertheless, the corresponding estimates of θ and ϕ for c_t as our time series of interest, $\hat{\theta}_c = (1, 0, 0)\,\hat{\theta}$ and $\hat{\phi}_c = (1, 0, 0)\,\hat{\phi}$, using the two estimation periods, are

$$\hat{\theta}_c = \begin{cases} \frac{1}{100}\,(-5.1, -18.5, -8.4, -11.3, -13.2, -12.1, -5.5, -1.0, -7.5)\,, & t = 1983q3, \ldots, 2022q1 \\ \frac{1}{100}\,(-5.6, -18.5, -8.4, -11.4, -14.0, -11.9, -5.2, -1.0, -7.3)\,, & t = 1983q3, \ldots, 2022q2 \end{cases}$$

and

$$\hat{\phi}_c = \begin{cases} \frac{1}{100}\,(-6.8, -15.8, -3.2, -1.8, -2.7,\ 4.6,\ 4.6,\ 1.0,\ \ 0.1)\,, & t = 1983q3, \ldots, 2022q1 \\ \frac{1}{100}\,(-6.7, -16.2, -3.3, -1.4, -2.6,\ 4.3,\ 4.1,\ 0.1, -0.2)\,, & t = 1983q3, \ldots, 2022q2. \end{cases}$$

Altogether, the estimates of θ_c and ϕ_c are fairly large and seem to align well with the quarters of extreme observations in c_t during the pandemic. For example, the largest estimated element of θ_c and ϕ_c both occur in $2020q2$ during which c_t, as shown in

Table 2: Estimates of three CVAR models — with a "normal" observation prior to forecasting

AO model

$$\hat{\alpha}' = \underset{(0.080)}{-0.216}, \underset{(0.073)}{0.275}, \underset{(0.113)}{0.111}\Bigg), \hat{\beta}' = \Bigg(\underset{(-)}{1}, \underset{(0.025)}{-0.788}, \underset{(0.025)}{-0.212}\Bigg), \chi^2(1) = 0.29[0.59]$$

IO model

$$\hat{\alpha}' = \underset{(0.082)}{-0.276}, \underset{(0.069)}{0.169}, \underset{(0.113)}{0.174}\Bigg), \hat{\beta}' = \Bigg(\underset{(-)}{1}, \underset{(0.026)}{-0.783}, \underset{(0.026)}{-0.217}\Bigg), \chi^2(1) = 0.15[0.70]$$

NC model

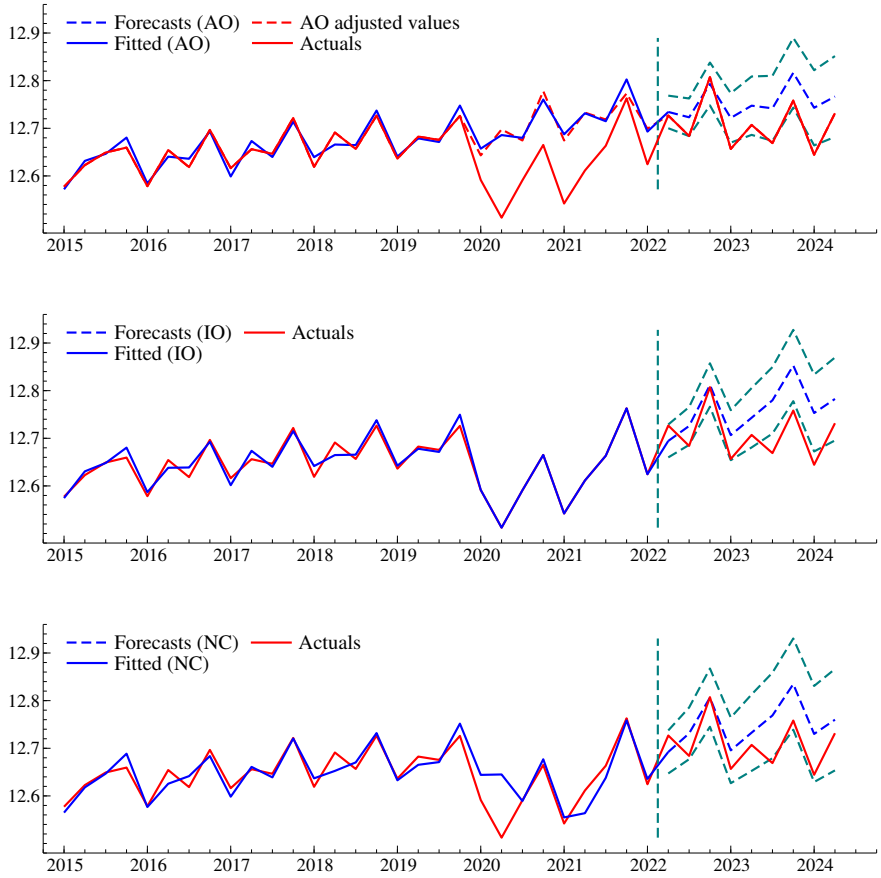
$$\hat{\alpha}' = \underset{(0.071)}{-0.285}, \underset{(0.050)}{0.201}, \underset{(0.075)}{0.047}\Bigg), \hat{\beta}' = \Bigg(\underset{(-)}{1}, \underset{(0.035)}{-0.800}, \underset{(0.035)}{-0.200}\Bigg), \chi^2(1) = 1.66[0.20]$$

Notes: Estimation period: $t = 1983q3, \ldots, 2022q2$. AO model, IO model and NC model denote the CVAR models in (3), (4) and (1), respectively, figures in parentheses are standard errors and $\chi^2(1)$ denotes the likelihood ratio test statistic for the hypothesis $\beta_y + \beta_w = 1$ with the p-value in square brackets.

Figure 1, was at its lowest observed level.[5] According to the estimates of θ_c, the negative effects on consumption through the five first quarters of the pandemic increase when $2022q2$ is also included in the estimation period. Through the remaining four quarters, however, the negative effects on consumption are slightly smaller. Overall, the updated estimates of θ_c when $2022q2$ is included in the estimation period imply a flatter slope for the AO adjusted consumption series, $c_t^{AO} = c_t - \hat{\theta}_c D_t$, which will affect the forecasts of consumption. We note from the $\chi^2(1)$ test statistics that the homogeneity restriction between c_t, y_t and w_t remains clearly significant with the implied outlier corrections in the AO and IO models as well as without any outlier corrections in the NC model across both estimation periods.

In the forecasting analysis below, we use the two sets of estimates from the three CVAR models to compare their out-of-sample properties. We emphasise that only the extreme observations during the pandemic, and not the level shifts in the growth rates of the time series after the pandemic, are explicitly modelled by different applications of the cluster of impulse dummies.

Figure 2: Actual values of c_t as well as fitted values and dynamic forecasts (with forecast error bands) from three CVAR models — without a "normal" observation prior to forecasting



Notes: Estimation period: $t = 1983q3, \ldots, 2022q1$ and forecasting period: $t = 2022q2, \ldots, 2024q2$. The AO, IO and NC fitted values and dynamic forecasts are derived from the CVAR models in (3), (4) and (1), respectively. The AO adjusted values of c_t during the pandemic in the upper panel are calculated as $c_t^{AO} = c_t - \hat{\theta}_c D_t$. The forecast error bands are 95 per cent confidence intervals.

2.4 Comparison of forecasting performance

Our comparison of forecasting performance is, in line with the two sets of estimates from the three CVAR models, based on the two forecasting periods $t = 2022q2, \ldots, 2024q2$ and $t = 2022q3, \ldots, 2024q2$. Accordingly, we compute dynamic forecasts of c_t from the three CVAR models using nine and eight available out-of-sample observations, respectively.

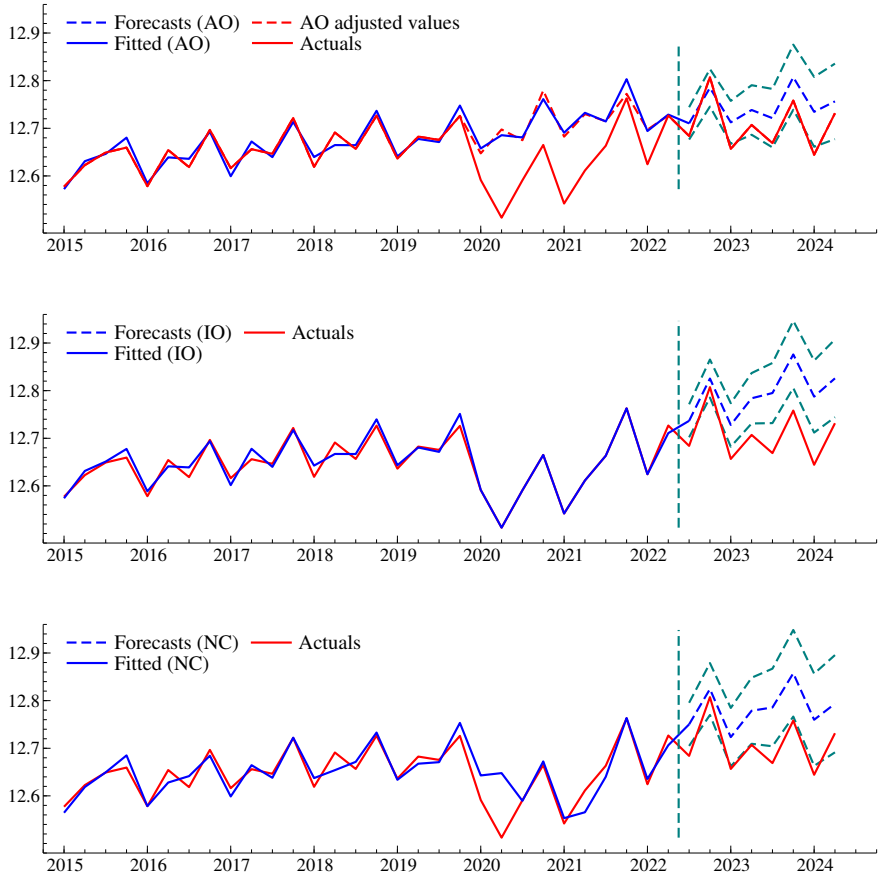
[5]The estimates of θ_c do not change much if y_t and w_t are also subject to additive outlier corrections across both estimation periods, which reflects that the developments of these two time series were not much affected by the pandemic, cf. Figure 1.

Figure 2 displays actual values of c_t, along with fitted values and dynamic forecasts (with forecast error bands) from the AO, IO and NC models, based on the estimation period $t = 1983q3, \ldots, 2022q1$ and the forecasting period $t = 2022q2, \ldots, 2024q2$. The upper panel of Figure 2 also records additive outlier adjusted values of c_t, $c_t^{AO} = c_t - \hat{\theta}_c D_t$, during the pandemic. Generally, the forecasting performance of the three models is quite good in the sense that the majority of the actual values of c_t fall inside or on the corresponding forecast error bands. A closer look at the different forecast paths, however, reveals that the directions of the forecasts from the AO model match those of the actual values perfectly throughout the forecasting period, while the directions of the forecasts from the IO and NC models are opposite to those of the actual values during the two quarters $2022q3$ and $2023q3$. All in all, the forecasts from the AO model seem to outperform the forecasts from the two model contenders.

These different forecasting properties are attributable to the varying influences of the extreme observations during the pandemic on the parameter estimates and the distinct conditioning values of consumption on the forecasts. As opposed to the estimates of the AO model, which are effectively based on observations of c_t prior to the pandemic and observations of y_t and w_t including the pandemic, the estimates of the IO and NC models are, respectively, based on observations of c_t, y_t and w_t prior to and including the pandemic. Whereas the forecasts from the AO model are conditioned on the predicted values of c_t during the pandemic, which are effectively counterfactual predicted values, the forecasts from the IO and NC models are conditioned on the actual values of c_t during the pandemic.

Figure 3 shows actual values of c_t, along with fitted values and dynamic forecasts (with forecast error bands) from the AO, IO and NC models, based on the estimation period $t = 1983q3, \ldots, 2022q2$ and the forecasting period $t = 2022q3, \ldots, 2024q2$. Again, the upper panel records additive outlier adjusted values of c_t during the pandemic. Having a "normal" observation instead of an extreme observation at the end of the estimation period, the forecasting performance of the AO model improves, while the forecasting performance of the IO and NC models worsens. Most actual values of c_t now fall inside the forecast error bands of the AO model, while most fall outside the forecast error bands of the IO and NC models. These improved forecasting properties of the AO model arise from the "normal" observation at the end of the estimation period, which provides a more accurate estimate of θ_c and thereby a more accurate forecasting path of consumption.

Figure 3: Actual values of c_t as well as fitted values and dynamic forecasts (with forecast error bands) from three CVAR models — with a "normal" observation prior to forecasting



Notes: Estimation period: $t = 1983q3, \ldots, 2022q2$ and forecasting period: $t = 2022q3, \ldots, 2024q2$. The AO, IO and NC fitted values and dynamic forecasts are derived from the CVAR models in (3), (4) and (1), respectively. The AO adjusted values of c_t during the pandemic in the upper panel are calculated as $c_t^{AO} = c_t - \hat{\theta}_c D_t$. The forecast error bands are 95 per cent confidence intervals.

We end the comparison of forecasting performance by means of the root mean squared forecast error (RMSFE) and mean absolute percentage error (MAPE) metrics for c_t. Table 3 presents values of the RMSFE and MAPE for c_t calculated by the forecasts from the three models across the forecasting periods $t = 2022q2, \ldots, 2024q2$ and $t = 2022q3, \ldots, 2024q2$.[6] Consistent with the evidence from Figures 2 and 3, the AO model outperforms the two model contenders in terms of having smaller values of RMSFE and MAPE for c_t across both forecasting periods. The forecasting

[6] $\text{RMSFE}_{c_t} = \left[\frac{1}{h}\sum_{t=1}^{h}(c_t - \hat{c}_t)^2\right]^{1/2}$ and $\text{MAPE}_{c_t} = \frac{100}{h}\sum_{t=1}^{h}\left|\frac{c_t - \hat{c}_t}{c_t}\right|$, where $h = 9, 8$ and $\hat{c}_t$ denotes the forecast of consumption.

Table 3: Forecasting performance for c_t from three CVAR models

(i) without a "normal" observation				(ii) with a "normal" observation			
	AO	IO	NC		AO	IO	NC
RMSFE	0.055	0.069	0.057	RMSFE	0.049	0.096	0.083
MAPE	0.378	0.464	0.384	MAPE	0.347	0.690	0.606

Notes: Forecasting period (i) $t = 2022q2, \ldots, 2024q2$ and (ii) $t = 2022q3, \ldots, 2024q2$. AO, IO and NC denote the CVAR models in (3), (4) and (1), respectively. Figures are root mean squared forecast error (RMSFE) and mean absolute percentage error (MAPE).

superiority of the AO model is, as discussed above, particularly pronounced in the case of $t = 2022q3, \ldots, 2024q2$, since the RMSFE and MAPE values from this model are only between 50 and 60 per cent of those from the IO and NC models.

Although empirical models, as emphasised by *inter alia* Clements and Hendry (1999, p. 25), never coincide perfectly with the underlying data generation process (DGP), they can match the data evidence in all measurable ways, and thereby be congruent prior to forecasting. However, when a structural change occurs after the in-sample period, the empirical models will not coincide with the DGP in the forecasting period. Generally, forecasting may or may not fail significantly when structural breaks occur. As demonstrated by Castle *et al.* (2024), a model with poor fit in-sample may produce far better forecasts than a model with good fit in-sample depending on the origins and characteristics of the structural breaks. Our empirical findings on the forecasting properties of the three CVAR models, both with and without outlier corrections during the pandemic, nevertheless motivate the theoretical formulation of forecasts in the next section, assuming that the AO model coincides with the DGP. The implications for forecast accuracy of simplified versions of the three CVAR models, when allowing for the unmodelled level shifts in the growth rates of c_t, y_t and w_t after the pandemic, are investigated through the Monte Carlo simulation experiments in Section 4.

3 Theoretical formulations

In this section, we focus exclusively on one-step-ahead forecasts. Although multi-step-ahead forecasts can also be derived, Clements and Hendry (1993, p. 631) point out that it is usually sufficient to consider one-step-ahead forecasts for ranking different models. First, we specify the general model for the DGP. Second, we examine forecasts produced using the AO, IO and NC models when the extreme observations occur at the end of the sample period. Third, we also assume the

presence of a "normal" observation prior to making the forecasts. Note that the parameters for the outliers in the AO and IO models are estimated using maximum likelihood.

3.1 A DPG with additive outliers

We assume the presence of n additive outliers, occurring towards the end of the sample period, which affect all variables in the system. This assumption differs from the assumption in the empirical example, where only one variable (consumption) exhibits such extreme observations. The fact that all the variables are affected by extreme values simplifies the formulation of the forecasts.

Initially, we assume that these outliers occur as the last n observations of the sample period. Later, we assume that there is one "normal" observation after the additive outliers. The observed values in the AO model can be decomposed by

$$Y_t = X_t + \theta D_t, \tag{5}$$

where Y_t is the p-dimensional vector of variables with the extreme observations from period $T-n+1$ to period T, X_t is the p-dimensional vector of variables without the extreme observations and $D_t = (1_{t=T-n+1}, \ldots, 1_{t=T})'$ is a vector of indicator variables. The corresponding additive outlier coefficients are given by the matrix $\theta = (\theta_{T-n+1}, \ldots, \theta_T)$, where θ_t is the p-dimensional vector of the coefficients for the additive outliers in period t (for $t = T-n+1, \ldots, T$). The observations net of the extreme observations are assumed to follow a CVAR model of order k satisfying Assumption 1 in Section 2.2:

$$\Delta X_t = \alpha\beta' X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu + \varepsilon_t, \tag{6}$$

where the parameters and the distribution of the error vector ε_t are defined in Section 2.2. The DGP is described by (5) and (6). We assume the absence of extreme observations or structural breaks during the forecasting period.

3.2 Absence of "normal" observations

In the absence of "normal" observations following the additive outliers, period T is the last observation prior to forecasting. In this case, we simplify by setting $k = 1$, so that the CVAR model includes only one lag. With $k = 1$, we can reformulate (6)

as

$$X_t = \mu + A_1 X_{t-1} + \varepsilon_t,$$

where $A_1 = I_p + \alpha\beta'$ as implicitly defined in Assumption 1 in Section 2.2.

Substituting this into (5), with the definitions of D_t and θ, yields

$$Y_t - \sum_{i=0}^{n-1} \theta_{T-i} 1_{t=T-i} = \mu + A_1 \left(Y_{t-1} - \sum_{i=0}^{n-1} \theta_{T-i} 1_{t-1=T-i} \right) + \varepsilon_t,$$

which for periods T and $T+1$ are given by

$$Y_T - \theta_T = \mu + A_1 \left(Y_{T-1} - \theta_{T-1} \right) + \varepsilon_T, \tag{7a}$$

$$Y_{T+1} = \mu + A_1 \left(Y_T - \theta_T \right) + \varepsilon_{T+1}. \tag{7b}$$

For later use, we also express these equations conditioned on Y_{T-n}, *i.e.* the last observation before the additive outliers, as

$$Y_T - \theta_T = \sum_{i=0}^{n-1} A_1^i \mu + A_1^n Y_{T-n} + \sum_{i=0}^{n-1} A_1^i \varepsilon_{T-i}, \tag{8a}$$

$$Y_{T+1} = \sum_{i=0}^{n} A_1^i \mu + A_1^{n+1} Y_{T-n-1} + \sum_{i=0}^{n} A_1^i \varepsilon_{T-i}. \tag{8b}$$

3.2.1 Forecasts with the AO model

Assume that the AO model, which is correctly specified, is estimated using observations from period $t = 1$ to period $t = T$. With n additive outliers, the observations from period $t = T-n+1$ are used to estimate the coefficient vectors for the additive outliers, denoted θ_t. Consequently, the estimates of the remaining parameters (μ and A_1) are effectively based on observations from period $t = 1$ to period $t = T-n$. We denote these estimates with a hat ($\hat{}$), as

$$Y_t = \hat{\mu} + \hat{A}_1 \left(Y_{t-1} - \sum_{i=0}^{n-1} \hat{\theta}_{T-i} 1_{t-1=T-i} \right) + \hat{\varepsilon}_t.$$

The one-step-ahead forecasts of Y_{T+1} using information up to and including period T, is given by

$$\hat{Y}^{AO}_{T+1|T} = \hat{\mu} + \hat{A}_1 \left(Y_T - \hat{\theta}_T \right). \tag{9}$$

The corresponding one-step-ahead forecast error, using (7b) and (9), can be expressed as

$$\begin{aligned}\hat{\varepsilon}^{AO}_{T+1|T} &= Y_{T+1} - \hat{Y}^{AO}_{T+1|T} \\ &= [\mu - \hat{\mu}] + \left[A_1 - \hat{A}_1\right] Y_T - \left[A_1\theta_T - \hat{A}_1\hat{\theta}_T\right] + \varepsilon_{T+1}.\end{aligned}$$

The first equality represents the definition of the one-step ahead forecast error, while the second equality decomposes the forecast error into biases in the estimates of μ, A_1 and $A_1\theta_T$ along with the unpredictable error term ε_{T+1}.

We can rewrite the one-step-ahead forecast error as

$$\hat{\varepsilon}^{AO}_{T+1|T} = [\mu - \hat{\mu}] + \left[A_1 - \hat{A}_1\right](Y_T - \theta_T) - \hat{A}_1\left(\theta_T - \hat{\theta}_T\right) + \varepsilon_{T+1}.$$

The estimate of the extreme observation is $\hat{\theta}_T = Y_T - \hat{Y}^{AO}_{T|T-n}$, *i.e.* the prediction error of Y in period T conditioned on observations up to period $T-n$. This is equivalent to replacing $Y_T - \hat{\theta}_T$ with its predicted value $\hat{Y}^{AO}_{T|T-n}$.

The "true" and estimated θ_T are, respectively, given by

$$\theta_T = Y_T - \left(\sum_{i=0}^{n-1} A_1^i\right)\mu - A_1^n Y_{T-n} - \sum_{i=0}^{n-1} A_1^i \varepsilon_{T-i} \tag{10}$$

and

$$\hat{\theta}_T = Y_T - \left(\sum_{i=0}^{n-1} \hat{A}_1{}^i\right)\hat{\mu} - \hat{A}_1{}^n Y_{T-n}, \tag{11}$$

where (10) is obtained by solving (8a) for θ_T and (11) represents the corresponding observational counterpart.

By combining (10) and (11), we obtain

$$\begin{aligned}\theta_T - \hat{\theta}_T = &- \left[\left(\sum_{i=0}^{n-1} A_1^i\right)\mu - \left(\sum_{i=0}^{n-1} \hat{A}_1{}^i\right)\hat{\mu}\right] \\ &- \left[A_1^n - \hat{A}_1{}^n\right] Y_{T-n} - \sum_{i=0}^{n-1} A_1^i \varepsilon_{T-i}.\end{aligned} \tag{12}$$

Thus, the corresponding one-step-ahead forecast error is given by

$$\begin{aligned}\hat{\varepsilon}^{AO}_{T+1|T} = &\left[I_p + \hat{A}_1\left(\sum_{i=0}^{n-1}\hat{A}_1^i\right)\right][\mu - \hat{\mu}] + \hat{A}_1\sum_{i=0}^{n-1}\left(A_1^i - \hat{A}_1^i\right)\mu \\ &+ \left[A_1 - \hat{A}_1\right](Y_T - \theta_T) \\ &+ \hat{A}_1\left[A_1^n - \hat{A}_1{}^n\right]Y_{T-n} \\ &+ \varepsilon_{T+1} + \hat{A}_1\sum_{i=0}^{n-1}A_1^i\varepsilon_{T-i}.\end{aligned}$$

3.2.2 Forecasts with the IO model

While the AO model corrects extreme observations of Y_T by using predicted values based on the additive outliers, the IO model corrects extreme values through innovational outliers. Using observations from period $t = 1$ to period $t = T$, the estimated IO model is

$$Y_t = \hat{\mu} + \hat{A}_1 Y_{t-1} + \sum_{i=0}^{n-1}\hat{\phi}_{T-i}1_{t=T-i} + \hat{\varepsilon}_t,$$

where $\hat{\phi}_{T-n+1}, \ldots, \hat{\phi}_T$ are the estimated coefficient vectors for the innovational outliers. When the observations of Y from period $T-n+1$ to period T are used to estimate ϕ, we have $\hat{\phi}_s = \hat{\theta}_s - \hat{A}_1\hat{\theta}_{s-1}$ (for $s = T-n+1, \ldots, T$ with $\hat{\theta}_{T-n} = 0$). Furthermore, since the remaining parameters are effectively estimated by observations from period $t = 1$ to period $t = T-n$, they will be equal to those estimated when using the AO model. Therefore, we use the same notation for the estimates of these parameters; $\hat{\mu}$ and $\hat{A}_1$.

The one-step-ahead forecast given by the IO model is

$$\hat{Y}^{IO}_{T+1|T} = \hat{\mu} + \hat{A}_1 Y_T.$$

The corresponding one-step-ahead forecast error is

$$\hat{\varepsilon}^{IO}_{T+1|T} = Y_{T+1} - \hat{Y}^{IO}_{T+1|T} = [\mu - \hat{\mu}] + \left[A_1 - \hat{A}_1\right]Y_T - A_1\theta_T + \varepsilon_{T+1}.$$

3.2.3 Forecasts with the NC model

Contrary to the AO and IO models, the NC model does not apply any corrections for the extreme observations in the periods from $T-n+1$ to T. The one-step-ahead

forecast and forecast error of the NC model will therefore have the same form as those of the IO model, *i.e.*

$$\hat{Y}^{NC}_{T+1|T} = \hat{\mu} + \hat{A}_1 Y_T,$$
$$\hat{\varepsilon}^{NC}_{T+1|T} = [\mu - \hat{\mu}] + \left[A_1 - \hat{A}_1\right] Y_T - A_1\theta_T + \varepsilon_{T+1}.$$

If we assume that the parameters μ and A_1 are estimated from period $t = 1$ to period $t = T - n$ only, the estimates of these parameters will be identical to those obtained by the AO and IO models. Then the forecasts ($\hat{Y}^{NC}_{T+1|T} = \hat{Y}^{IO}_{T+1|T}$) and the forecast errors ($\hat{\varepsilon}^{NC}_{T+1|T} = \hat{\varepsilon}^{IO}_{T+1|T}$) will also be identical to those produced by the IO model.

3.2.4 Comparison of forecasts

When comparing the forecasts from the three models, we disregard the estimation uncertainty of μ, A_1 and Ω in order to isolate the effects of the extreme observations on the forecast errors. The expectations of the forecast errors for the three models, assuming no uncertainty in the parameter estimates except for those associated with the extreme observations, are given by

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

Thus, the correctly specified AO model yields unbiased forecasts, whereas the forecasts of the IO and NC models are biased. The bias in these models depends on the magnitude of the extreme observations, θ_T, and the persistence in the dynamics, A_1.

The forecast uncertainty, measured by the squared forecast errors, are given by

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

It is noteworthy that the forecast uncertainty of the AO model increases with the number of extreme observations, n. If $n = 1$, the forecast uncertainty of the AO model is the smallest, provided that the extreme observation is substantial, *i.e.* $A_1\theta_T\theta_T'A_1' - A_1\Omega A_1' = A_1\left(\theta_T\theta_T' - \Omega\right)A_1' \succ 0$, where $\succ$ is the succeeds operator such that $A \succ 0$ implies that the matrix A is positive definite (i.e. $x'Ax > 0, \forall x \neq 0$). If A_1 is non-singular, we can write this assumption as $\theta_T\theta_T' - \Omega \succ 0$, see Lemma 1 in Appendix C. If n is large, the forecast uncertainty of the AO model can be the largest, which occurs when $\sum_{i=1}^{n} A_1^i\Omega A_1^{i\prime} - A_1\theta_T\theta_T'A_1' \succ 0$, even though this is the correctly specified model.

3.3 Presence of a "normal" observation

Thus far, we have considered forecasts in cases where the extreme observations occur at the end of the sample period. We will now examine the case in which one "normal" observation follows the additive outliers at the end of the sample period, just prior to making forecasts. This "normal" observation can be used to improve the estimate of the additive outliers in the AO model and thereby reduce the forecast uncertainty.

When one "normal" observation is available after the additive outliers, the forecasts of Y in period $T+2$ use information up to and including period $T+1$, while the extreme observations remain within the n periods spanning from period $T-n+1$ to period T. In this case, the one-step-ahead forecasts, assuming that the estimates of μ and A_1 are the same for the three models, will be identical across all three models when $k = 1$. Therefore, we also consider the case when $k = 2$, where the forecasts are conditioned on an extreme observation. In this case, the forecasts of the AO model will differ from those of the IO and NC models.

3.3.1 One lag in the CVAR models

When $k = 1$ and one "normal" observation is available after the additive outliers, the forecasts of all three models are given by

$$\hat{Y}^i_{T+2|T+1} = \hat{\mu} + \hat{A}_1 Y_{T+1}, \quad i = AO, IO, NC,$$

where the estimates of μ and A_1 may differ across the models. Under certain conditions, however, these estimates can be identical across the models. This uniformity in estimates can be achieved by effectively estimating μ and A_1 based on observa-

tions prior to the extreme observations, implying that the AO and IO models are both estimated based on observations from period 1 to period $T-1$. The last n observations in this estimation sample are used to estimate θ (for the AO model) or ϕ (for the IO model). The remaining parameters are then estimated using observations from period 1 to period $T-n-1$. Therefore, we obtain the same estimates of these parameters in the NC model if they are estimated based on observations from period 1 to period $T-n-1$.[7] Thus, both the one-step-ahead forecasts and the forecast errors for period $T+2$, conditional on information available up to period $T+1$, will be identical across the three models.

These findings show that when a "normal" observation is available upon which the forecasts are conditioned, the forecasts of the three models will be the same, provided the following two conditions are met: First, the estimates of μ and A_1 are identical across the models. Second, the three estimated models use only one lag, so that they condition solely on a "normal" observation when making the forecasts. In the following, we will explore the importance of the last condition by considering the case where two lags are used in the estimated models.

3.3.2 Two lags in the CVAR models

When $k=2$, the system in (6) combined with (5) alongside the definitions of D_t and θ can be formulated as

$$\begin{aligned} Y_t - \sum_{i=0}^{n-1} \theta_{T-i} 1_{t=T-i} =& \mu + A_1 \left(Y_{t-1} - \sum_{i=0}^{n-1} \theta_{T-i} 1_{t-1=T-i} \right) + \\ & A_2 \left(Y_{t-2} - \sum_{i=0}^{n-1} \theta_{T-i} 1_{t-2=T-i} \right) + \varepsilon_t, \end{aligned}$$

where $A_1 = (I_p + \Gamma_1) + \alpha\beta'$ and $A_2 = -\Gamma_1$. When there is only one extreme observation in period T, *i.e.* $n=1$, the DGP simplifies to

$$Y_t - \theta_T 1_{t=T} = \mu + A_1 \left(Y_{t-1} - \theta_T 1_{t-1=T} \right) + A_2 \left(Y_{t-2} - \theta_T 1_{t-2=T} \right) + \varepsilon_t, \tag{13}$$

[7] Under this assumption, there is no need to use different notations for the estimates across the three models.

which for the periods T, $T+1$ and $T+2$ becomes

$$Y_T - \theta_T = \mu + A_1 Y_{T-1} + A_2 Y_{T-2} + \varepsilon_T, \tag{14a}$$

$$Y_{T+1} = \mu + A_1 (Y_T - \theta_T) + A_2 Y_{T-1} + \varepsilon_{T+1}, \tag{14b}$$

$$Y_{T+2} = \mu + A_1 Y_{T+1} + A_2 (Y_T - \theta_T) + \varepsilon_{T+2}. \tag{14c}$$

For the AO model, the one-step-ahead forecast is given by

$$\hat{Y}^{AO}_{T+2|T+1} = \hat{\mu} + \hat{A}_1 Y_{T+1} + \hat{A}_2 \left(Y_T - \hat{\theta}_T \right)$$

and the corresponding forecast error is

$$\begin{aligned} \hat{\varepsilon}^{AO}_{T+2|T+1} = [\mu - \hat{\mu}] + \left[A_1 - \hat{A}_1 \right] Y_{T+1} + \left[A_2 - \hat{A}_2 \right] (Y_T - \theta_T) \\ - \hat{A}_2 \left[\theta_T - \hat{\theta}_T \right] + \varepsilon_{T+2}. \end{aligned} \tag{15}$$

The estimate of θ_T must be carefully considered. We consider two possible estimators of θ_T: the first, denoted $\hat{\theta}_{T|T}$, uses information up to period T; the second, denoted $\hat{\theta}_{T|T+1}$, also incorporates the additional information from period $T+1$. For the first estimator, we apply (12) with $n = 1$. For the second, we exploit the additional information from the observational counterpart to (14b).

The general estimator for θ_T, given observations up to period $T+s$ (where $s \geq 0$) and estimates of the other parameters (μ, A_1, A_2, and Ω), is implicitly given by

$$\hat{\theta}_{T|T+s} = \arg\min \sum_{t=1}^{T+s} \hat{\varepsilon}_t \left(\theta_{T|T+s} \right)' \hat{\Omega}^{-1} \hat{\varepsilon}_t \left(\theta_{T|T+s} \right), \text{ where} \tag{16}$$

$$\begin{aligned} \hat{\varepsilon}_t \left(\theta_{T|T+s} \right) = & Y_t - \theta_{T|T+s} 1_{t=T} - \hat{\mu} \\ & - \hat{A}_1 \left(Y_{t-1} - \theta_{T|T+s} 1_{t-1=T} \right) - \hat{A}_2 \left(Y_{t-2} - \theta_{T|T+s} 1_{t-2=T} \right). \end{aligned}$$

In Proposition 1, we consider the forecast bias and accuracy of the AO model with this estimator of $\theta_{T|T+s}$ for $s = 0$ and $s = 1$ when there is one extreme observation in period T only.

Proposition 1 *Given that the parameter vector θ_T in* (13) *is estimated through* (16) *conditioned on the estimates of μ, A_1, A_2 and Ω, the forecasts of period $T+2$ by the AO model with one extreme observation in period T and one "normal" observation in period $T+1$ exhibit the following properties:*

(i) If the estimator of θ_T where only information up to and including period T, denoted $\hat{\theta}_{T|T}$, is applied, the forecast is unbiased since

$$\mathcal{E}\left[\hat{\varepsilon}^{AO}_{T+2|T+1}\left(\hat{\theta}_{T|T}\right);\hat{\mu}=\mu,\hat{A}_1=A_1,\hat{A}_2=A_2\right]=0. \tag{17}$$

The corresponding covariance matrix is given by

$$\begin{aligned}\mathcal{E}&\left[\hat{\varepsilon}^{AO}_{T+2|T+1}\left(\hat{\theta}_{T|T}\right)\hat{\varepsilon}^{AO}_{T+2|T+1}\left(\hat{\theta}_{T|T}\right)';\hat{\mu}=\mu,\hat{A}_1=A_1,\hat{A}_2=A_2,\hat{\Omega}=\Omega\right]\\ =&\Omega+A_2\Omega A_2'.\end{aligned} \tag{18}$$

(ii) If the estimator of θ_T where information up to and including period $T+1$, denoted $\hat{\theta}_{T|T+1}$, is applied, the forecast is unbiased since

$$\mathcal{E}\left[\hat{\varepsilon}^{AO}_{T+2|T+1}\left(\hat{\theta}_{T|T+1}\right);\hat{\mu}=\mu,\hat{A}_1=A_1,\hat{A}_2=A_2\right]=0. \tag{19}$$

The corresponding covariance matrix is given by

$$\begin{aligned}\mathcal{E}&\left[\hat{\varepsilon}^{AO}_{T+2|T+1}\left(\hat{\theta}_{T|T+1}\right)\hat{\varepsilon}^{AO}_{T+2|T+1}\left(\hat{\theta}_{T|T+1}\right)';\hat{\mu}=\mu,\hat{A}_1=A_1,\hat{A}_2=A_2,\hat{\Omega}=\Omega\right]\\ =&\Omega+A_2\left[\Omega^{-1}+A_1'\Omega^{-1}A_1\right]^{-1}A_2'.\end{aligned} \tag{20}$$

Proof. *(i) When $k=2$ and $n=1$, the "true" and estimated values of θ_T, given information up to period T, are, respectively, given by*

$$\theta_T=Y_T-\mu-A_1Y_{T-1}-A_2Y_{T-2}-\varepsilon_T \tag{21}$$

and

$$\hat{\theta}_{T|T}=Y_T-\hat{\mu}-\hat{A}_1Y_{T-1}-\hat{A}_2Y_{T-2}, \tag{22}$$

where (21) *is derived by solving* (14a) *for θ_T and* (22) *represents the estimated conterpart. The bias is given by the difference*

$$\theta_T-\hat{\theta}_{T|T}=-\left[\mu-\hat{\mu}\right]-\left[A_1-\hat{A}_1\right]Y_{T-1}-\left[A_2-\hat{A}_2\right]Y_{T-2}-\varepsilon_T. \tag{23}$$

The results in (17) *and* (18) *follow simply by inserting* (23) *in* (15) *and taking the conditional expectation and variance.*

(ii) As shown in (16), *$\hat{\theta}_{T|T+1}$ is included only when $t=T$ and $t=T+1$.*

This estimator for θ_T minimises the sum of the corresponding two squared errors $\hat{\varepsilon}_T'\hat{\Omega}^{-1}\hat{\varepsilon}_T + \hat{\varepsilon}_{T+1}'\hat{\Omega}^{-1}\hat{\varepsilon}_{T+1}$. The estimator is a weighted average of $\hat{\theta}_{T|T}$ in (22) *and the observational counterpart to* (14b) *with $\hat{\varepsilon}_{T+1}$ set to zero:*

$$\begin{aligned}(24) \qquad \hat{\theta}_{T|T+1} = &\left[\hat{\Omega}^{-1} + \hat{A}_1'\hat{\Omega}^{-1}\hat{A}_1\right]^{-1}\hat{\Omega}^{-1}\left[Y_T - \hat{\mu} - \hat{A}_1 Y_{T-1} - \hat{A}_2 Y_{T-2}\right] \\ &- \left[\hat{\Omega}^{-1} + \hat{A}_1'\hat{\Omega}^{-1}\hat{A}_1\right]^{-1}\hat{A}_1'\hat{\Omega}^{-1}\left[Y_{T+1} - \hat{\mu} - \hat{A}_1 Y_T - \hat{A}_2 Y_{T-1}\right],\end{aligned}$$

which corresponds to equation 11 in Nielsen (2004). Substituting (24) *into* (15) *and taking the conditional expectation and variance yield the results in* (19) *and* (20). ■

Both AO estimators for θ give unbiased estimates. However, the estimator using information up to period $T+1$, denoted $\hat{\theta}_{T|T+1}$, results in the lowest forecast uncertainty. To demonstrate this, we compare the covariance matrices in (18) and (20). Subtracting the latter from the former yields

$$(25) \qquad A_2\Omega A_2' - A_2\left[\Omega^{-1} + A_1'\Omega^{-1}A_1\right]^{-1}A_2'$$

as the terms involving Ω alone cancel out. In Proposition 2, we show that the resulting matrix is positive definite if both A_1 and A_2 are non-singular.

Proposition 2 *Suppose both A_1 and A_2 are non-singular. Then*

$$A_2\Omega A_2' - A_2\left[\Omega^{-1} + A_1'\Omega^{-1}A_1\right]^{-1}A_2' \succ 0.$$

Proof. *We can rewrite the difference as follows:*

$$A_2\Omega A_2' - A_2\left[\Omega^{-1} + A_1'\Omega^{-1}A_1\right]^{-1}A_2' = A_2\left[\Omega - \left(\Omega^{-1} + A_1'\Omega^{-1}A_1\right)^{-1}\right]A_2',$$

which is positive definite if and only if

$$\Omega - \left(\Omega^{-1} + A_1'\Omega^{-1}A_1\right)^{-1}$$

is positive definite when A_2 is non-singular, see Lemma 1 in Appendix C.

Since A_1 is non-singular and Ω is positive definite, $A_1'\Omega^{-1}A_1$ is according to Lemma 1 in Appendix C positive definite and so is $\Omega^{-1} + A_1'\Omega^{-1}A_1$. Define

$M = \Omega^{-1} + A_1'\Omega^{-1}A_1$ *and* $N = \Omega^{-1}$ *to find*

$$M - N = \Omega^{-1} + A_1'\Omega^{-1}A_1 - \Omega^{-1} = A_1'\Omega^{-1}A_1 \succ 0,$$

so that

$$N^{-1} - M^{-1} = \Omega - \left(\Omega^{-1} + A_1'\Omega^{-1}A_1\right)^{-1} \succ 0,$$

according to Magnus and Neudecker (2019, Theorem 24 in Chapter 1). ■

For the IO and NC models, the forecasts are given by

$$\hat{Y}^i_{T+2|T+1} = \hat{\mu} + \hat{A}_1 Y_{T+1} + \hat{A}_2 Y_T, \quad i = IO, NC$$

and the corresponding forecast errors are

$$\begin{aligned}\hat{\varepsilon}^i_{T+2|T+1} = &[\mu - \hat{\mu}] + \left[A_1 - \hat{A}_1\right] Y_{T+1} \\ &+ \left[A_2 - \hat{A}_2\right] Y_T - A_2\theta_T + \varepsilon_{T+2}, \quad i = IO, NC.\end{aligned}$$

If the estimates of μ, A_1 and A_2 are the same, the forecasts of the IO and NC models are identical; $\hat{Y}^{IO}_{T+2|T+1} = \hat{Y}^{NC}_{T+2|T+1}$. The conditional forecast bias and the corresponding forecast uncertainty are given by

$$\mathcal{E}\left[\hat{\varepsilon}^i_{T+2|T+1}; \hat{\mu} = \mu, \hat{A}_1 = A_1, \hat{A}_2 = A_2\right] = -A_2\theta_T$$

and

$$\begin{aligned}\mathcal{E}\left[\hat{\varepsilon}^i_{T+2|T+1}\hat{\varepsilon}^i_{T+2|T+1}{}'; \hat{\mu} = \mu, \hat{A}_1 = A_1, \hat{A}_2 = A_2, \hat{\Omega} = \Omega\right] = &\Omega + A_2\theta_T\theta_T'A_2', \\ &i = IO, NC.\end{aligned}$$

Thus, both the IO and NC models still provide biased forecasts.

In the above analysis, we have considered the case where there is only one extreme observation at the end of the sample period. As in the case when $k = 1$, the forecast uncertainty of the AO model is typically smaller than that of the IO and NC models. However, similar to the case when $k = 1$, the forecast uncertainty of the AO model will increase if there are more extreme observations close to the end of the sample period. Then, the AO model may not longer provide the most accurate forecasts. Nonetheless, if one (or more) "normal" observations are available at the end of the sample period, the estimate of the additive outlier will be more precise,

leading to a reduction in the forecast uncertainty of the AO model compared to the forecast uncertainty of the AO model without this information. Consequently, the forecasts of the AO model with one "normal" observation at the end of the sample period may be more accurate than those of the IO and NC models, even if the forecasts of the AO model without one "normal" observation are less precise than those of the IO and NC models.

3.4 Theoretical versus empirical findings

The theoretical framework for comparing forecasting performance, assuming that a CVAR model with additive outliers in all variables of Y_t coincides with the DGP in-sample, has considered four cases: (i) $k = 1$, n extreme observations and no "normal" observations at the end of the sample period; (ii) $k = 1$, n extreme observations and one "normal" observation at the end of the sample period; (iii) $k = 2$, one extreme observation and no "normal" observations at the end of the sample period and (iv) $k = 2$, one extreme observation and one "normal" observation at the end of the sample period. For simplicity, the theoretical framework has ignored uncertainty of parameter estimates, except for θ and ϕ, across the AO, IO and NC models in all four cases.

In contrast, the framework for the empirical example has considered two cases: (i) $k = 6$, $n = 9$ and no "normal" observations at the end of the sample period and (ii) $k = 6$, $n = 9$ and one "normal" observation at the end of the sample period. In addition, the framework for the empirical example has assumed that only one of the variables of Y_t contains extreme observations. For these reasons, the estimated parameters of interest reported in Tables 1 and 2 differ somewhat, though not much, across the three CVAR models. As discussed in Section 2.4, the varying impacts of the extreme observations during the pandemic on the parameter estimates and the distinct conditioning values of consumption on the forecasts give rise to the different forecasting properties reported in Table 3 across the three CVAR models.

Although the theoretical and empirical frameworks differ in terms of the lag length, the number of extreme observations and the number of variables affected by extreme observations, the theoretical results support the empirical findings. Specifically, the forecast accuracy of the AO model improves significantly relative to the forecast accuracy of the IO and NC models when one "normal" observation is available prior to forecasting.

In the theoretical formulation of forecasts, we have relied on simplifying assumptions of a short lag length ($k = 1$ or $k = 2$), a small number of extreme observations ($n = 1$) for some of the results and no out-of-sample structural breaks. The implications of relaxing these assumptions for the relative forecast accuracy of the three CVAR models are explored through the Monte Carlo experiments presented in the next section.

4 Monte Carlo experiments

This section undertakes Monte Carlo simulation experiments with a view to giving quantitative support to the theoretical results developed above. The simulation settings are formulated by partially referring to the preceding empirical study of the Norwegian macroeconomic data for household consumption, income and wealth, a study which has proved to be a noteworthy example for further research. First, we discuss a set of simulation settings related to a DGP and a class of three CVAR models. Second, we conduct a benchmark comparative study of forecasting performance of the different models. Third, we compare forecasting performance in the presence of unmodelled level shifts.[8]

4.1 DGP and three models

The DGP is formulated here as a trivariate CVAR process $Y_t = (Y_{1,t}, Y_{2,t}, Y_{3,t})'$ for $t = 1, \ldots, T + j + h$, which comprises a combination of (i) a dynamic process, (ii) a cluster of nine additive outliers and (iii) a pair of deterministic shifts in the level (also referred to as an unrestricted window function). The dynamic part is formulated as $X_t = (X_{1,t}, X_{2,t}, X_{3,t})'$ subject to $r = 1$, the cluster of additive outliers is given as $D_t = (1_{t=T-8}, \ldots, 1_{t=T})'$ and the unrestricted window function is given as $d_t = 1_{t_0 \leq t \leq t_1}$, in which $1_{(\cdot)}$ denotes an indicator function. The values for T, h, t_0 and t_1 vary in the experiments. Note that (i) $T + j$ represents the total number of observations for estimation with the default $T = 100$ and $j = 0$, (ii) h corresponds to a forecast horizon and (iii) $d_t = 0$ for all t is assumed in the benchmark study, implying there are no level shifts in the process. The DGP is conceived to be a simplified version of the Norwegian consumption-function model, so that the triplet $(Y_{1,t}, Y_{2,t}, Y_{3,t})'$ is viewed as corresponding to the triplet $(c_t, y_t, w_t)'$

[8]The Ox (see Doornik, 2013) code written to conduct the Monte Carlo simulations in this section is available from the authors upon request.

studied in the empirical example. The first series $Y_{1,t}$ of the triplet is thus seen as the variable of most interest. The specification of the DGP is

$$Y_t = X_t + \theta D_t$$

and

$$\Delta X_t = \alpha\left(\beta' X_{t-1} + \rho\right) + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \eta + \omega d_t + \varepsilon_t,$$

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

$$\theta_1 = \begin{pmatrix} -0.10 & -0.12 & -0.12 & -0.12 & -0.10 & -0.10 & -\phi & -\phi & -\phi \end{pmatrix}$$

for $\phi = 0.06$ or $\phi = 0.08$,

$$\alpha = \begin{pmatrix} -0.2 \\ 0.2 \\ 0.0 \end{pmatrix},\ \beta = \begin{pmatrix} 1.0 \\ -0.8 \\ -0.2 \end{pmatrix},\ \rho = 0.68,\ \eta = \begin{pmatrix} 0.0 \\ 0.014 \\ 0.006 \end{pmatrix},\ \omega = \begin{pmatrix} -0.01 \\ 0.0 \\ 0.0 \end{pmatrix}$$

and

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

Note that the unrestricted constant is given as $\mu = \alpha\rho + \eta$, ensuring that the dynamic part of the DGP is formulated in the same way as in (1) under Assumption 1 in Section 2.2. The lag length k is assigned 2 or 4, so that only the corresponding parameter matrices of Γ_1, Γ_2 and Γ_3 are involved in the DGP. The error process ε_t is a trivariate *i.i.d.* pseudo normal process, $N(0,\Omega)$, in which Ω is a positive definite symmetric matrix, where each diagonal and off-diagonal element is assigned a unit value and a quarter, respectively. All the elements are multiplied by a scalar damping factor 0.015^2. The starting values of $X_{1,t}$, $X_{2,t}$ and $X_{3,t}$ for $t = -3,\ldots,0$ are assigned values close to the initial values for c_t, y_t and w_t in the empirical example, respectively. In line with common practice (see *e.g.* Castle *et al.*, 2024), the number of Monte Carlo replications in each experiment is set to $10,000$.

Using the DGP above, we conduct a range of Monte Carlo experiments to compare forecasting performance of various models. The performance is assessed using the root mean squared forecast error (RMSFE) metric. We study in principle

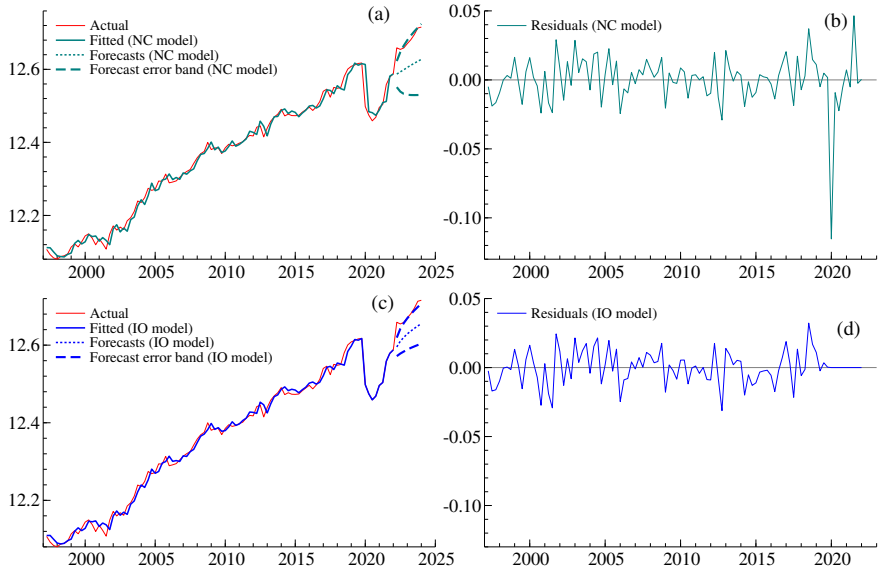
the performance of three CVAR models, which are all well-specified in terms of lag length k and cointegrating rank r but vary in terms of the manners of addressing outliers in the DGP. The three models are, in line with Sections 2 and 3, as follows: a CVAR model not corrected with any dummy variables (NC model), a CVAR model corrected with nine innovational outlier (dummy) variables (IO model) and a CVAR model corrected with nine additive outlier (dummy) variables (AO model). Note that the estimated AO model is correctly specified in that it encompasses the underlying DGP when $d_t = 0$, *i.e.* when there are no level shifts in the DGP. This means that we only estimate additive outliers for the first variable of Y_t in accordance with the empirical example. Also consistent with the empirical study, the estimation of the AO model employs the restricted maximum likelihood method, aligning with the structure of the parameter matrix θ given above.

4.2 Benchmark comparative study

We begin the comparative study by discussing the results of a class of benchmark experiments, in which the DGPs have 100 observations ($T = 100$), no "normal" observations prior to forecasting ($j = 0$), a relatively short forecasting horizon ($h = 8$), no presence of level shifts ($d_t = 0$), relatively small additive outliers at the end of the sample period ($\phi = 0.06$) and both short and long lag lengths ($k = 2$ and $k = 4$). Note that no "normal" observations prior to forecasting ($j = 0$) imply that the end of the additive outlier series coincides with the end of the sample period effective for estimation.

We first check a set of sample paths of $Y_{1,t}$, the variable of most interest in this study, based on a single realisation (replication) derived from the three models, before launching a detailed RMSFE-based comparative study. Figure 4 displays a set of sample paths of $Y_{1,t}$ (actual values) as well as dynamic forecasts, fitted values and residuals derived from the NC and IO models with a short lag length ($k = 2$). Note that the time range $1997q1 - 2025q1$ is assigned to the horizontal axis to make the artificial data conceivable with the quarterly data presented in the empirical study. A series of additive outliers in the DGP causes a huge drop in the residuals of the NC model, as shown in panel (b). In contrast, those outliers are fully addressed in the IO model, resulting in a series of zero residuals in panel (d). The forecast paths from the two models (see panels (a) and (c)) appear to be comparable to some extent. However, the forecast error bands are rather different, reflecting the influences of the outliers on the parameter estimates.

Figure 4: Actual values of $Y_{1,t}$ as well as dynamic forecasts, fitted values and residuals from the NC and IO models — without a "normal" observation prior to forecasting

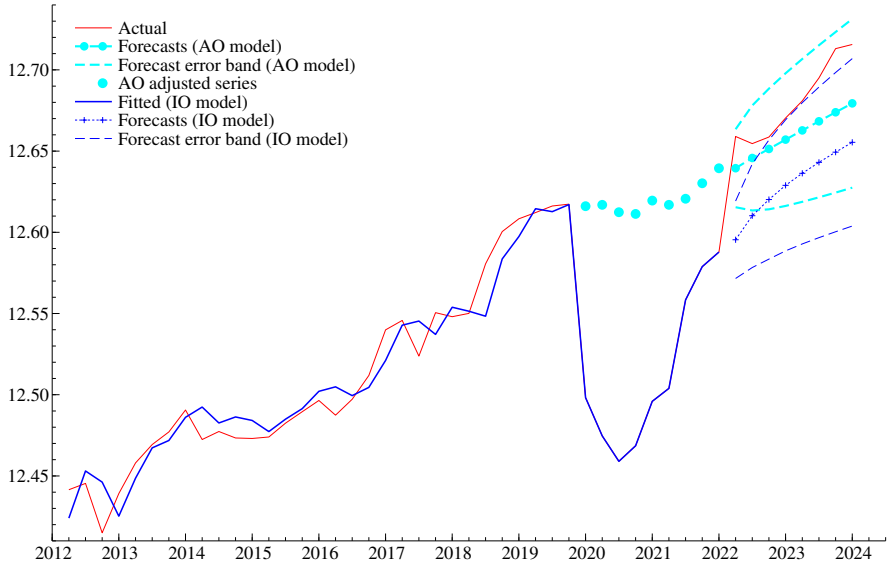


Notes: (a) Dynamic forecasts (NC model), (b) Residuals (NC model), (c) Dynamic forecasts (IO model) and (d) Residuals (IO model). $T = 100$, $j = 0$, $h = 8$, $d_t = 0$, $\phi = 0.06$ and $k = 2$. Based on a single replication. The forecast error bands are 95 per cent confidence intervals.

Figure 5 records a set of sample paths of $Y_{1,t}$ as well as dynamic forecasts and fitted values derived from the AO model alongside those derived from the IO model, taken from Figure 4 (c). Note that the series are displayed only over the latter half of the sample period, so as to focus on the comparison of the two forecast paths. The class of dots indicates a series of AO adjusted values, $X_{1,t} = Y_{1,t} - \hat{\theta}_1 D_t$, which are free from the effects of the underlying outliers due to the AO adjustment. The path of dynamic forecasts from the AO model tracks the actual value path much better than that from the IO model, an encouraging feature attributable to the AO adjustment.

We proceed to a detailed RMSFE-based comparative study with the number of Monte Carlo replications being $10,000$. Table 4 presents a battery of RMSFEs in the benchmark cases with short ($k = 2$) and long ($k = 4$) lag length, respectively. Consistent with the evidence in Figures 4 and 5, the AO model outperforms the other two models in terms of RMSFEs for $Y_{1,t}$. This advantage is also observed in almost all of the other RMSFEs for $Y_{2,t}$ and $Y_{3,t}$. Table 5 reports the results for a long forecasting horizon ($h = 20$) with all the other settings identical to those in

Figure 5: Actual values of $Y_{1,t}$ as well as dynamic forecasts and fitted values from the AO and IO models — without a "normal" observation prior to forecasting



Notes: The AO adjusted series, indicated by a class of dots, is calculated as $X_{1,t} = Y_{1,t} - \hat{\theta}_1 D_t$. $T = 100$, $j = 0$, $h = 8$, $d_t = 0$, $\phi = 0.06$ and $k = 2$. Based on a single replication. The forecast error bands are 95 per cent confidence intervals.

Table 4: 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 ($\phi = 0.06$)

(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). $T = 100$, $j = 0$, $h = 8$, $d_t = 0$ and $\phi = 0.06$. Based on $10,000$ replications.

Table 4. We arrive at the same conclusion as above in terms of the assessment of the RMSFEs.

Next, we compare a class of RMSFEs when the magnitude of the last three additive outliers is increased from $\phi = 0.06$ to $\phi = 0.08$. All the other settings in the DGP are unchanged from those in the previous experiments. Tables 6 and 7 show that the superiority of the AO model is much enhanced, in comparison with Tables 4 and 5, reflecting a larger magnitude of additive outliers at the end of the

Table 5: Forecasting performance of three CVAR models — without a "normal" observation ($j = 0$), long forecasting horizon ($h = 20$) and small additive outliers at the end of the sample period ($\phi = 0.06$)

(i) short lag length ($k = 2$)				(ii) long lag length ($k = 4$)			
	NC	IO	AO		NC	IO	AO
$Y_{1,t}$	0.0731	0.0464	0.0420	$Y_{1,t}$	0.0789	0.0583	0.0522
$Y_{2,t}$	0.0332	0.0384	0.0333	$Y_{2,t}$	0.0420	0.0496	0.0422
$Y_{3,t}$	0.0627	0.0674	0.0622	$Y_{3,t}$	0.0814	0.0931	0.0797

Notes: Figures are root mean squared forecast errors (RMSFEs). $T = 100,\ j = 0,\ h = 20,\ d_t = 0$ and $\phi = 0.06$. Based on $10,000$ replications.

Table 6: 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 ($\phi = 0.08$)

(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). $T = 100,\ j = 0,\ h = 8,\ d_t = 0$ and $\phi = 0.08$. Based on $10,000$ replications.

Table 7: Forecasting performance of three CVAR models — without a "normal" observation ($j = 0$), long forecasting horizon ($h = 20$) and large additive outlier at the end of the sample period ($\phi = 0.08$)

(i) short lag length ($k = 2$)				(ii) long lag length ($k = 4$)			
	NC	IO	AO		NC	IO	AO
$Y_{1,t}$	0.0876	0.0514	0.0421	$Y_{1,t}$	0.0927	0.0639	0.0522
$Y_{2,t}$	0.0335	0.0416	0.0333	$Y_{2,t}$	0.0424	0.0533	0.0422
$Y_{3,t}$	0.0630	0.0696	0.0622	$Y_{3,t}$	0.0819	0.0974	0.0797

Notes: Figures are root mean squared forecast errors (RMSFEs). $T = 100,\ j = 0,\ h = 20,\ d_t = 0$ and $\phi = 0.08$. Based on $10,000$ replications.

sample period.

As the final analysis of the benchmark comparative study, we evaluate quantitatively the strategy of extending the sample period such that an additional "normal" observation is taken into account in the estimation. This strategy was employed with success in the empirical study and its validity was assessed from a theoretical point of view in the previous section. The sample period for estimation is now extended to $T + j = 101$ for $j = 1$ (instead of $j = 0$ so far), and the forecasts are then

Table 8: 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 ($\phi = 0.08$)

(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). $T = 100,\ j = 1,\ h = 8,\ d_t = 0$ and $\phi = 0.08$. Based on $10,000$ replications.

Table 9: Forecasting performance of three CVAR models — with a "normal" observation ($j = 1$), long forecasting horizon ($h = 20$) and large additive outlier at the end of the sample period ($\phi = 0.08$)

(i) short lag length ($k = 2$)				(ii) long lag length ($k = 4$)			
	NC	IO	AO		NC	IO	AO
$Y_{1,t}$	0.0462	0.0440	0.0377	$Y_{1,t}$	0.0568	0.0625	0.0472
$Y_{2,t}$	0.0365	0.0349	0.0333	$Y_{2,t}$	0.0473	0.0514	0.0420
$Y_{3,t}$	0.0648	0.0636	0.0620	$Y_{3,t}$	0.0855	0.0884	0.0804

Notes: Figures are root mean squared forecast errors (RMSFEs). $T = 100,\ j = 1,\ h = 20,\ d_t = 0$ and $\phi = 0.08$. Based on $10,000$ replications.

calculated over $T + j + h = 101 + h$. We refer to this as having one "normal" observation ($j = 1$) prior to forecasting. Note that the start of dynamic forecasting is pushed forward one quarter as a result of extending the sample period for estimation. We continue to choose a relatively large value of the additive outliers ($\phi = 0.08$) at the end of the sample period. The RMSFEs in the cases of a short ($h = 8$) and long ($h = 20$) forecasting horizon are recorded in Tables 8 and 9, respectively.

Overall, the RMSFEs when having one "normal" observation ($j = 1$) prior to forecasting become smaller than those in Tables 6 and 7. It is particularly noteworthy that the superiority of the AO model now appears more distinctive, quantitatively showing that adding one "normal" observation at the end of the estimation sample is a useful method for improving the forecasting performance of this model.

4.3 Presence of unmodelled level shifts

We have so far studied a class of benchmark cases where only a set of additive outliers is present in the DGP, *i.e.* no level shifts take place, resulting in an ideal situation where all the AO models employed encompass the underlying DGP. Since

the additive-type breaks are well captured by the AO adjustment in the benchmark study, it may seem natural that the AO models outperform the two other model contenders. The potentials of AO models in forecasting capability are, however, not necessarily limited to such a setting as in the benchmark study.

We therefore proceed to the study of cases where d_t involves level shifts in the process in such a manner that

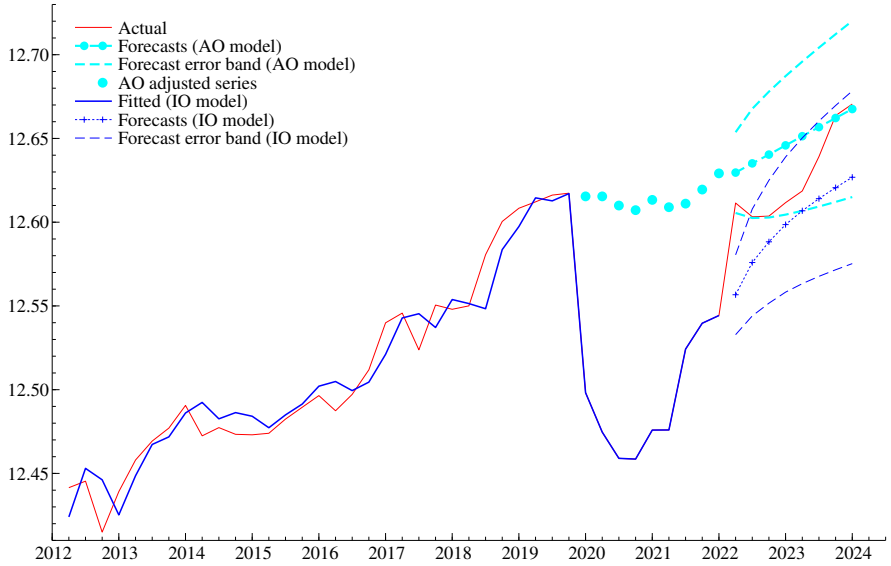
$$d_t = 1_{t_0 \le t \le t_1}, \quad \text{for} \quad t = 1, \ldots, T + j + h, \tag{26}$$

where $t_0 = T - 5$ and $t_1 = T + 3$ are assigned here. Recall that the coefficient of d_t is $\omega = (-0.01, 0, 0)'$. In other words, a window function $d_t = 1_{t_0 \le t \le t_1}$ is introduced in the DGP unrestrictedly and in such a manner that the range of its non-zero values partially overlaps the series of additive outliers $D_t = (1_{t=T-8}, \ldots, 1_{t=T})'$ and slightly continues after it. As a result of this re-specification of the DGP, level shifts (going down at t_0 and moving back to the original level at $t_1 + 1$) are allowed to occur in the underlying process. It should be noted that the end point t_1 is greater than T, so that the level-shifting effects on the process continue even after the additive outliers are removed. The seasonally adjusted consumption path shown in Figure 1 provides the impetus for introducing the level shifts given by (26) to account for a plausible shift in the slope of the underlying trend following the COVID-19 pandemic period. Since an unrestricted intercept in the CVAR process generates a linear trend via the Granger-Johansen moving average representation, see *e.g.* Johansen (1996, Theorem 4.2) and Hansen (2005), it seems natural to approximate the plausible slope shift with (26) in the DGP for the simulation study.

Despite this change in the DGP, all the estimated models are the same as those in the benchmark study and thus fail to capture the level shifts. This corresponds to a realistic situation where the mixture of the various underlying breaks in the process renders it difficult for econometricians to identify all of the breaks correctly when building their models. In other words, all the models are misspecified to the underlying DGP, which is subject to undetected level shifts.

Figure 6 presents a set of sample paths of $Y_{1,t}$ as well as dynamic forecasts and fitted values derived from the AO and IO models in the case of unmodelled level shifts ($d_t = 1_{T-5 \le t \le T+3}$) and no "normal" observations prior to forecasting ($j = 0$). We set $\phi = 0.06$ here to focus on the study of the undetected level shifts. In contrast to Figure 5, the dynamic forecast path from the IO model aligns more closely with the actual value path compared to the AO model, likely due to the level shifts in the

Figure 6: Actual values of $Y_{1,t}$ as well as dynamic forecasts and fitted values from the AO and IO models in the presence of unmodelled level shifts — without a "normal" observation prior to forecasting

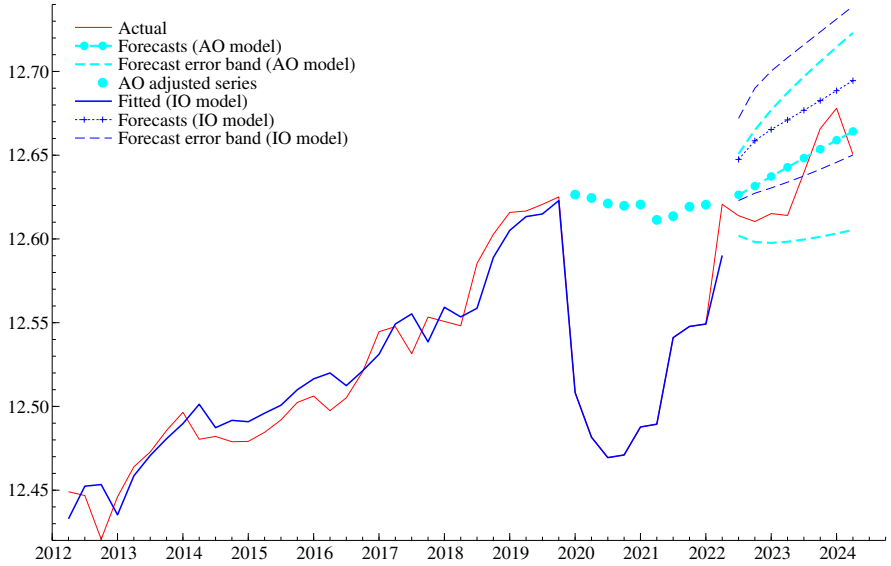


Notes: The AO adjusted series, indicated by a class of dots, is calculated as $X_{1,t} = Y_{1,t} - \hat{\theta}_1 D_t$. $T = 100$, $j = 0$, $h = 8$, $d_t = 1_{T-5 \leq t \leq T+3}$, $\phi = 0.06$ and $k = 2$. Based on a single replication. The forecast error bands are 95 per cent confidence intervals.

DGP. The AO model has only partially adjusted for the overall breaks, which may have negatively impacted its forecasting performance compared to the IO model.

Table 10 records a class of RMSFEs from the three models. As anticipated from Figure 6, we observe that the RMSFEs of $Y_{1,t}$ based on the AO model are larger than those based on the IO model. We are therefore justified in concluding that the advantage of the IO model over the AO model is a general property to be noted under the present DGP. Based on the overall arguments presented in this paper, we predict that having one "normal" observation ($j = 1$) prior to forecasting will be useful to effectively address this property. We will investigate the efficacy of this adjustment in resolving the issues observed in Figure 6 and Table 10. Specifically, we aim to evaluate whether this adjustment can serve as an effective method for improving forecasts during breaks, as discussed by Castle *et al.* (2019). We continue to use the term 'a "normal" observation' here, despite the presence of unmodelled level shifts, on the grounds that the influential additive outliers have become zero from $T + 1$ onwards.

Figure 7: Actual values of $Y_{1,t}$ as well as dynamic forecasts and fitted values from the AO and IO models in the presence of unmodelled level shifts — with a "normal" observation prior to forecasting



Notes: The AO adjusted series, indicated by a class of dots, is calculated as $X_{1,t} = Y_{1,t} - \hat{\theta}_1 D_t$. $T = 100$, $j = 1$, $h = 8$, $d_t = 1_{T-5 \leq t \leq T+3}$, $\phi = 0.06$ and $k = 2$. Based on a single replication. The forecast error bands are 95 per cent confidence intervals.

Table 10: Forecasting performance of three CVAR models in the presence of unmodelled level shifts — without a "normal" observation ($j = 0$), short forecasting horizon ($h = 8$) and small additive outliers at the end of the sample period ($\phi = 0.06$)

(i) short lag length ($k = 2$)				(ii) long lag length ($k = 4$)			
	NC	IO	AO		NC	IO	AO
$Y_{1,t}$	0.0559	0.0381	0.0487	$Y_{1,t}$	0.0553	0.0528	0.0570
$Y_{2,t}$	0.0275	0.0279	0.0295	$Y_{2,t}$	0.0339	0.0392	0.0363
$Y_{3,t}$	0.0381	0.0465	0.0372	$Y_{3,t}$	0.0478	0.0689	0.0460

Notes: Figures are root mean squared forecast errors (RMSFEs). $T = 100$, $j = 0$, $h = 8$, $d_t = 1_{T-5 \leq t \leq T+3}$ and $\phi = 0.06$. Based on 10,000 replications.

Figure 7 is an updated version of Figure 6 when having one "normal" observation ($j = 1$) prior to forecasting. As a consequence of having one "normal" observation ($j = 1$), the forecast path of $Y_{1,t}$ derived from the AO model is now back on the right track, and the model appears to surpass the IO model in forecast tracking. This is an encouraging finding, leading us to a detailed study of RMSFEs.

Table 11 reports a set of RMSFEs when there is a "normal" observation prior

Table 11: Forecasting performance of three CVAR models in the presence of unmodelled level shifts — with a "normal" observation ($j = 1$), short forecasting horizon ($h = 8$) and small additive outliers at the end of the sample period ($\phi = 0.06$)

(i) short lag length ($k = 2$)				(ii) long lag length ($k = 4$)			
	NC	IO	AO		NC	IO	AO
$Y_{1,t}$	0.0311	0.0486	0.0302	$Y_{1,t}$	0.0383	0.0680	0.0362
$Y_{2,t}$	0.0339	0.0293	0.0254	$Y_{2,t}$	0.0432	0.0445	0.0309
$Y_{3,t}$	0.0390	0.0405	0.0381	$Y_{3,t}$	0.0483	0.0548	0.0471

Notes: Figures are root mean squared forecast errors (RMSFEs). $T = 100,\ j = 1,\ h = 8,\ d_t = 1_{T-5\leq t\leq T+3}$ and $\phi = 0.06$. Based on $10,000$ replications.

Table 12: Forecasting performance of three CVAR models in the presence of unmodelled level shifts — without a "normal" observation ($j = 0$), long forecasting horizon ($h = 20$) and small additive outliers at the end of the sample period ($\phi = 0.06$)

(i) short lag length ($k = 2$)				(ii) long lag length ($k = 4$)			
	NC	IO	AO		NC	IO	AO
$Y_{1,t}$	0.0715	0.0456	0.0511	$Y_{1,t}$	0.0761	0.0622	0.0614
$Y_{2,t}$	0.0393	0.0378	0.0415	$Y_{2,t}$	0.0510	0.0547	0.0533
$Y_{3,t}$	0.0633	0.0711	0.0622	$Y_{3,t}$	0.0835	0.1022	0.0809

Notes: Figures are root mean squared forecast errors (RMSFEs). $T = 100,\ j = 0,\ h = 20,\ d_t = 1_{T-5\leq t\leq T+3}$ and $\phi = 0.06$. Based on $10,000$ replications.

to forecasting. In line with Figure 7, the RMSFEs of $Y_{1,t}$ from the AO model are smaller than those from the IO model. This is also true for the RMSFEs of $Y_{2,t}$ and $Y_{3,t}$. Accordingly, having one "normal" observation ($j = 1$) prior to forecasting works well in this study. The same specific adaptation also leads to a range of successful outcomes when the forecasting horizon is extended to $h = 20$. The RMSFEs in this case are reported in Tables 12 and 13 for $j = 0$ and $j = 1$, respectively. A comparison of these two tables indicates the usefulness of having one "normal" observation ($j = 1$) in the context of long-range forecasting.

The statistical theory developed in Section 3, based on a class of simplifying assumptions, has demonstrated the advantage of the AO model for out-of-sample forecasting. The Monte Carlo experiments in this section have shown that the theoretical arguments hold true in the context of a quantitative comparative analysis using generalised settings, including unmodelled level shifts. These simulation results lend weight to the practicality of the AO model as a reliable forecasting device.

Table 13: Forecasting performance of three CVAR models in the presence of unmodelled level shifts — with a "normal" observation ($j = 1$), long forecasting horizon ($h = 20$) and small additive outliers at the end of the sample period ($\phi = 0.06$)

(i) short lag length ($k = 2$)			
	NC	IO	AO
$Y_{1,t}$	0.0446	0.0511	0.0396
$Y_{2,t}$	0.0450	0.0411	0.0365
$Y_{3,t}$	0.0635	0.0640	0.0626

(ii) long lag length ($k = 4$)			
	NC	IO	AO
$Y_{1,t}$	0.0542	0.0686	0.0501
$Y_{2,t}$	0.0592	0.0618	0.0470
$Y_{3,t}$	0.0841	0.0888	0.0825

Notes: Figures are root mean squared forecast errors (RMSFEs). $T = 100,\ j = 1,\ h = 20,\ d_t = 1_{T-5 \leq t \leq T+3}$ and $\phi = 0.06$. Based on $10,000$ replications.

5 Conclusions

In this paper, we have presented key findings that offer notable improvements to forecasting following extreme events, such as pandemics. Specifically, we have demonstrated that additive outlier (AO) corrections outperform both innovational outlier (IO) corrections and no outlier corrections within the CVAR framework when macroeconomic variables rapidly return to their initial trajectories following transient extreme observations.

Using data from the COVID-19 pandemic, we empirically have shown that CVAR models incorporating AO corrections yield more accurate forecasts of post-pandemic Norwegian household consumption than those incorporating IO corrections or no outlier adjustments. The forecasting superiority of the AO corrections is especially pronounced when the estimation sample includes a "normal" observation following the extreme observations in-sample. Furthermore, we have developed a theoretical framework that explains the superior forecasting performance of AO corrections when the extreme observations are short-lived. Through simplifying assumptions of a short lag length, a limited number of extreme observations and no out-of-sample structural breaks, we have demonstrated that while CVAR models with IO and no outlier corrections produce biased forecasts, the model with AO corrections provides unbiased forecasts with potentially lower forecast uncertainty. Notably, forecast uncertainty may diminish further when the model conditions on a "normal" observation following the extreme observations. Our Monte Carlo simulations, conducted under more general settings involving longer lag lengths, a higher number of extreme observations and structural breaks during the forecasting period, support the robustness of the theoretical findings.

The accurate forecasting of macroeconomic variables is critical for the effec-

tive implementation of fiscal and monetary policies. Therefore, incorporating AO corrections into CVAR models is crucial when rapid normalisation of macroeconomic variables is expected following extreme observations. Failure to account for such corrections could lead to policy recommendations that inadvertently exacerbate, rather than mitigate, business cycle fluctuations.

Bibliography

Ahumada HA, Garegnani ML (2012). "Forecasting a monetary aggregate under instability: Argentina after 2001." *International Journal of Forecasting*, **28**(2), 412–427. doi:10.1016/j.ijforecast.2011.01.008.

Akpan EA, Lasisi KE, Adamu A, Rann HB (2019). "Application of Iterative Approaches in Modeling the Efficiency of ARIMA-GARCH Processes in the Presence of Outliers." *Applied Mathematics*, **10**(03), 138–158. doi:10.4236/am.2019.103012.

Battaglia F, Cucina D, Rizzo M (2020). "Detection and estimation of additive outliers in seasonal time series." *Computational Statistics*, **35**(3), 1393–1409. doi:10.1007/s00180-019-00928-5.

Bauwens L, Koop G, Korobilis D, Rombouts JV (2015). "The Contribution of Structural Break Models to Forecasting Macroeconomic Series." *Journal of Applied Econometrics*, **30**(4), 596–620. doi:10.1002/jae.2387.

Boug P, Cappelen Å, Jansen ES, Swensen AR (2021). "The Consumption Euler Equation or the Keynesian Consumption Function?" *Oxford Bulletin of Economics and Statistics*, **83**(1), 252–272. doi:10.1111/obes.12394.

Castle JL, Clements MP, Hendry DF (2015). "Robust approaches to forecasting." *International Journal of Forecasting*, **31**(1), 99–112. doi:10.1016/j.ijforecast.2014.11.002.

Castle JL, Clements MP, Hendry DF (2016). "An Overview of Forecasting Facing Breaks." *Journal of Business Cycle Research*, **12**(1), 3–23. doi:10.1007/s41549-016-0005-2.

Castle JL, Clements MP, Hendry DF (2019). *Forecasting: An Essential Introduction*. Yale University Press. doi:10.2307/j.ctvfjd0c2.

Castle JL, Doornik JA, Hendry DF (2024). "Improving models and forecasts after equilibrium-mean shifts." *International Journal of Forecasting*, **40**(3), 1085–1100. doi:10.1016/j.ijforecast.2023.09.006.

Catalán B, Trívez FJ (2007). "Forecasting volatility in GARCH models with additive outliers." *Quantitative Finance*, **7**(6), 591–596. doi:10.1080/14697680601116872.

Clements MP, Hendry DF (1993). "On the limitations of comparing mean square forecast errors." *Journal of Forecasting*, **12**(8), 617–637. doi:10.1002/for.3980120802.

Clements MP, Hendry DF (1996). "Intercept corrections and structural change." *Journal of Applied Econometrics*, **11**(5), 475–494. doi:10.1002/(SICI)1099-1255(199609)11:5⟨475::AID-JAE409⟩3.0.CO;2-9.

Clements MP, Hendry DF (1998). *Forecasting Economic Time Series*. Cambridge University Press. doi:10.1017/CBO9780511599286.

Clements MP, Hendry DF (1999). *Forecasting Non-Stationary Economic Time Series*. MIT Press.

De Gooijer JG, Hyndman RJ (2006). "25 years of time series forecasting." *International Journal of Forecasting*, **22**(3), 443–473. doi:10.1016/j.ijforecast.2006.01.001.

Doornik J, Hendry DF (2018). *Modelling Dynamic Systems - PcGive 15: Volume II*. Timberlake.

Doornik JA (2013). *Ox 7 : an object-oriented matrix programming language*. Timberlake.

Eitrheim Ø, Jansen E, Nymoen R (2002). "Progress from forecast failure-the Norwegian consumption function." *The Econometrics Journal*, **5**(1), 40–64. doi:10.1111/1368-423X.t01-1-00072.

Findley DF, Monsell BC, Bell WR, Otto MC, Chen BC (1998). "New capabilities and methods of the x-12-arima seasonal-adjustment program." *Journal of Business and Economic Statistics*, **16**(2), 127–152. doi:10.1080/07350015.1998.10524743.

Fox AJ (1972). "Outliers in Time Series." *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **34**(3), 350–363. doi:10.1111/j.2517-6161.1972.tb00912.x.

Franses PH, Ghijsels H (1999). "Additive outliers, GARCH and forecasting volatility." *International Journal of Forecasting*, **15**(1), 1–9. doi:10.1016/S0169-2070(98)00053-3.

Franses PH, Haldrup N (1994). "The effects of additive outliers on tests for unit roots and cointegration." *Journal of Business and Economic Statistics*, **12**(4), 471–478. doi:10.1080/07350015.1994.10524569.

Franses PH, Lucas A (1998). "Outlier detection in cointegration analysis." *Journal of Business and Economic Statistics*, **16**(4), 459–468. doi:10.1080/07350015.1998.10524785.

Hansen PR (2005). "Granger's representation theorem: A closed-form expression for I(1) processes." *The Econometrics Journal*, **8**(1), 23–38. doi:10.1111/j.1368-423X.2005.00149.x.

Hotta LK, Trucíos C (2018). "Inference in (M)GARCH Models in the Presence of Additive Outliers: Specification, Estimation, and Prediction." *Advances in Mathematics and Applications*, (M), 179–202. doi:10.1007/978-3-319-94015-1{_}8.

Hungnes H (2010). "Identifying Structural Breaks in Cointegrated Vector Autoregressive Models." *Oxford Bulletin of Economics and Statistics*, **72**(4), 551–565. doi:10.1111/j.1468-0084.2010.00586.x.

Jansen ES (2013). "Wealth effects on consumption in financial crises: The case of Norway." *Empirical Economics*, **45**(2), 873–904. doi:10.1007/s00181-012-0640-y.

Johansen S (1996). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press.

Johansen S, Mosconi R, Nielsen B (2000). "Cointegration analysis in the presence of structural breaks in the deterministic trend." *The Econometrics Journal*, **3**(2), 216–249. doi:10.1111/1368-423X.00047.

Johansen S, Nielsen MØ (2018). "The cointegrated vector autoregressive model with general deterministic terms." *Journal of Econometrics*, **202**(2), 214–229. doi:10.1016/j.jeconom.2017.10.003.

Kurita T, Nielsen B (2019). "Partial cointegrated vector autoregressive models with structural breaks in deterministic terms." *Econometrics*, **7**(4). doi:10.3390/econometrics7040042.

Laome L, Adhi Wibawa GN, Raya R, Makkulau, Asbahuna AR (2021). "Forecasting time series data containing outliers with the ARIMA additive outlier method." *Journal of Physics: Conference Series*, **1899**(1), 8–13. doi:10.1088/1742-6596/1899/1/012106.

Ledolter J (1989). "The effect of additive outliers on the forecasts from ARIMA models." *International Journal of Forecasting*, **5**(2), 231–240. doi:10.1016/0169-2070(89)90090-3.

Ljung GM (1993). "On Outlier Detection in Time Series." *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **55**(2), 559–567. doi:10.1111/j.2517-6161.1993.tb01924.x.

Lütkepohl H, Saikkonen P, Trenkler C (2004). "Testing for the Cointegrating Rank of a VAR Process with Level Shift at Unknown Time." *Econometrica*, **72**(2), 647–662. doi:10.1111/j.1468-0262.2004.00505.x.

Magnus JR, Neudecker H (2019). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley Series in Probability and Statistics. Wiley. doi:10.1002/9781119541219.

Nielsen HB (2004). "Cointegration analysis in the presence of outliers." *The Econometrics Journal*, **7**(1), 249–271. doi:10.1111/j.1368-423X.2004.00130.x.

Nielsen HB (2008). "Influential observations in cointegrated VAR models: Danish money demand 1973–2003." *The Econometrics Journal*, **11**(1), 39–57. doi:10.1111/j.1368-423X.2007.00226.x.

Patrocinio PF, Reisen VA, Bondon P, Monte EZ, Danilevicz IM (2024). "M-Quantile Estimation for GARCH Models." *Computational Economics*, **63**(6), 2175–2192. doi:10.1007/s10614-023-10398-z.

Petropoulos F, Apiletti D, Assimakopoulos V, Babai MZ, Barrow DK, Ben Taieb S, Bergmeir C, Bessa RJ, Bijak J, Boylan JE, Browell J, Carnevale C, Castle JL, Cirillo P, Clements MP, Cordeiro C, Cyrino Oliveira FL, De Baets S, Dokumentov A, Ellison J, Fiszeder P, Franses PH, Frazier DT, Gilliland M, Gönül MS,

Goodwin P, Grossi L, Grushka-Cockayne Y, Guidolin M, Guidolin M, Gunter U, Guo X, Guseo R, Harvey N, Hendry DF, Hollyman R, Januschowski T, Jeon J, Jose VRR, Kang Y, Koehler AB, Kolassa S, Kourentzes N, Leva S, Li F, Litsiou K, Makridakis S, Martin GM, Martinez AB, Meeran S, Modis T, Nikolopoulos K, Önkal D, Paccagnini A, Panagiotelis A, Panapakidis I, Pavía JM, Pedio M, Pedregal DJ, Pinson P, Ramos P, Rapach DE, Reade JJ, Rostami-Tabar B, Rubaszek M, Sermpinis G, Shang HL, Spiliotis E, Syntetos AA, Talagala PD, Talagala TS, Tashman L, Thomakos D, Thorarinsdottir T, Todini E, Trapero Arenas JR, Wang X, Winkler RL, Yusupova A, Ziel F (2022). "Forecasting: theory and practice." *International Journal of Forecasting*, **38**(3), 705–871. doi: 10.1016/j.ijforecast.2021.11.001.

Appendices

A Data definitions and sources

The seasonally unadjusted data used for estimation and forecasting purposes (c_t, y_t and w_t) are taken from the databank of the KVARTS model of Statistics Norway and the seasonally unadjusted and adjusted data used for descriptive purposes (c_t^{sa}, y_t^{sa}, S_t and S_t^{sa}) are taken from the quarterly national and non-financial sector accounts in the Statbank of Statistics Norway. The data were downloaded in September 2024 and are available from the authors upon request. Details about data definitions and sources are listed below.

Seasonally unadjusted data

c_t The log of household real consumption of goods and services excluding housing (fixed 2021 prices). Housing consumption is, as in Boug *et al.* (2021), excluded because the imputed value of housing consumption is closely related to the imputed value of housing income by construction in the national accounts. Source: The databank of the KVARTS model of Statistics Norway.

y_t The log of household real disposable income excluding equity dividends (fixed 2021 prices), defined as household nominal disposable income excluding equity dividends deflated by the price deflator for household consumption of goods and services excluding housing (2021 = 1). Equity dividends are, as in Boug *et al.* (2021), excluded due to some years during the sample period where tax-increases announced in advance implied substantial tax related fluctuations in this income component, which are likely to be less motivating for consumption than other income sources. Source: The databank of the KVARTS model of Statistics Norway.

w_t The log of household real net wealth (fixed 2021 prices), defined as household nominal net wealth, the sum of net financial wealth and housing wealth, deflated by the price deflator for household consumption of goods and services excluding housing (2021 = 1). See Boug *et al.* (2021) for further details. Source: The databank of the KVARTS model of Statistics Norway.

S_t The household savings ratio excluding equity dividends (per cent). Source: The quarterly non-financial sector accounts in the Statbank of Statistics Norway, Table 11020.

We follow Eitrheim *et al.* (2002), Jansen (2013) and Boug *et al.* (2021) and work with non-per capita consumption, income and wealth in the empirical example because $C_t/N_t = (Y_t/N_t)^{(1-\beta_y)} \cdot (W_t/N_t)^{\beta_y}$, where N_t denotes the population, is equivalent to $c_t = (1-\beta_y)y_t + \beta_y w_t$ in the case of homogeneity between consumption, income and wealth. As shown in Section 2.2, the homogeneity restriction is indeed supported by the data.

Seasonally adjusted data

c_t^{sa} The log of household real consumption of goods and services excluding housing (fixed 2021 prices). Source: The quarterly national accounts in the Statbank of Statistics Norway, Table 09173.

y_t^{sa} The log of household real disposable income excluding equity dividends (fixed 2021 prices). Source: The quarterly non-financial sector accounts in the Statbank of Statistics Norway, Table 11020.

S_t^{sa} The household savings ratio excluding equity dividends (per cent). Source: The quarterly non-financial sector accounts in the Statbank of Statistics Norway, Table 11020.

B The CVAR model covering the global financial crisis

The CVAR model in (1), covering the global financial crisis, is specified as

$$\begin{aligned}\Delta Y_t &= \alpha \begin{pmatrix}\beta \\ \gamma\end{pmatrix}' \begin{pmatrix}Y_{t-1} \\ t \times SD_t\end{pmatrix} + \sum_{i=1}^{k-1} \Gamma_i \Delta Y_{t-i} + \mu^* SD_t + \sum_{i=1}^{k} \delta_i ID_{t-i} \\ &\quad + \psi CS_t + \varepsilon_t,\end{aligned}$$

where $SD_t = (SD_{1,t}, SD_{2,t})'$, $SD_{1,t} = 1_{t \leq 2008q3}$ and $SD_{2,t} = 1_{t \geq 2010q2}$ denote step dummies which equal one for, respectively, the sub-periods before and after the global financial crises, zero otherwise, and $ID_t = (1_{t=2008q4}, 1_{t=2009q1}, \ldots, 1_{t=2010q1})'$ denote impulse dummies which equal one for single quarters in the period of the global financial crises, zero otherwise. Accordingly, (1) allows for separate deterministic trends (a broken deterministic trend) for the sub-periods before and after the global financial crisis. The associated parameters are $\gamma = (\gamma_1', \gamma_2')'$, $\mu^* = (\mu_1, \mu_2)$ and δ_i for $i = 1, \ldots, k$. Also, (1) includes ψ which contains the three parameters of centred seasonal dummies CS_t.

The benchmark cointegration analysis considers a deterministic specification in which SD_t, ID_t and CS_t enter as unrestricted terms in (1) and $t \times SD_t$ enters as a restricted term in the cointegration space. We find, assuming $r = 1$, that the restriction $\gamma_2 = 0$ is not rejected since $\chi^2(1) = 2.14$ (p-value $= 0.14$), implying that there is no deterministic trend in the cointegrating vector after the global financial crisis. The joint hypothesis of $\beta_y + \beta_w = 1$ and $\gamma_2 = 0$ is also not rejected since $\chi^2(2) = 2.26$ (p-value $= 0.32$).

Since $\gamma_2 = 0$, we can ignore $tSD_{2,t}$ in (1). In addition, we can ignore both $SD_{1,t}$ and ID_t for observations after the global financial crises as $SD_{1,t} = 0$ and $ID_t = 0_{6\times1}$ for $t > 2010q1$. We therefore use $\mu = \mu_2$ in (1).

C Results in Section 3

In Section 3, the following lemma is applied to compare forecast accuracy between the three models:

Lemma 1 *If A is non-singular, then it follows that S is positive definite if and only if ASA' is positive definite.*

Proof. *Proof that S is positive definite implies that ASA' is positive definite when A is non-singular: S is positive definite implies that $x'Sx > 0$ for all non-zero x. Then $y'(ASA')y > 0$ for $y' = x'A^{-1}$ and x spans the entire space.*

Proof that ASA' is positive definite implies that S is positive definite when A is non-singular: ASA' is positive definite implies that $x'ASA'x > 0$ for all non-zero x. Then $y'Sy > 0$ for $y = A'x$ and y spans the entire space. ■