Time series estimation and forecasting of Covid in Norway

Gunnar Bårdsen Ragnar Nymoen

Part of NFR project no. 324472

25th Dynamic Econometrics conference
13 April 2023

Introduction

- Covid time series are typical examples of changing data-generating processes, because of
  - mutations and policy responses
  - over time and across economies
- Likely that constant-parameter models will fail in forecasting, even if they are correct in a constant parameter world.
- There might be strategies to derive forecasting models with a higher degree of robustness against breaks in relationships and parameters. Such approaches include trend models, stochastic trend models, autoregressive models, and robust forecasting methods, see Li and Linton (2021), Harvey and Kattuman (2021), Doornik et al. (2022), Mills (2022) amongst others.
- Here, a framework for forecasting new cases (incidence), hospital admissions and hospital beds is presented. This project, named CovidMod, produced 21 days ahead forecasts each working day from 17 March 2021 to 1 April 2022 (real time forecasts).

Introduction (continued)

- The model is next extended to allow for endogenous effects of policy responses, like lockdowns and vaccinations, to counter the events of exogenous effects, like new mutations. The threshold of the target variable for policy reactions is estimated. The forecasting performance is compared to the forecasts of the Norwegian Institute of Public Health (NIPH) as well as to the real time forecasts of CovidMod.

A model based on autoregressive processes

- Builds on Nymoen (2022).
- Let Y_t denote new cases on day t,

$$Y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \epsilon_t \tag{1}$$

where $\alpha_i, i = 0, 1., ...p$ are parameters and ϵ_t is an error term which we assume linearly independent of $Y_{t-1}, ..., Y_{t-1-p}$. Equation (1) can be re-written as:

$$Y_t = \alpha_0 + \beta(\sum_{i=1}^{p-1} Y_{t-i}) + \sum_{i=1}^{p-1} \alpha_i^{\ddagger} \Delta Y_{t-i} + \epsilon_t \tag{2}$$

where

$$\beta = \frac{\sum_{i=1}^{p} \alpha_i}{p-1} \tag{3}$$

and $\alpha_i^{\ddagger}$ $(i = 1, 2, ..., p-1)$ are combinations of the original autoregressive coefficients .

A model based on autoregressive processes (continued)

- The accumulated number of confirmed cases for day t is denoted S_{Yt}. It is given by the definition:

$$S_{Yt} = Y_t + S_{Yt-1}, \tag{4}$$

Equation (2) can therefore be expressed with the change $S_{Yt-1} - S_{Yt-p}$ on the right hand side:

$$Y_t = \alpha_0 + \beta(S_{Yt-1} - S_{Yt-p}) + \sum_{i=0}^{p-1} \alpha_i^{\ddagger} \Delta Y_{t-1-i} + \epsilon_t. \tag{5}$$

A working hypothesis could be that the amount of virus in the population, call it S_t^* (which is not observable) is increasing in the accumulated number of confirmed cases over a period of p days:

$$S_t^* = f(S_{Yt-1} - S_{Yt-p}),\ f' > 0, \tag{6}$$

Equations (4) and (5) highlight the positive feedback between between new cases and how much corona virus there is in the population.

The model contains four endogenous variables:

- Y_t, number of new infected with Covid-19, day t.
- S_{Yt}, accumulated number of new infected with Covid-19, day t.
- HA_t, number of new hospitalisations wih Covid-19, day t.[1]
- HB_t, number of hospitalisations with Covid-19, day t.[2]

[1] https://www.fhi.no/sv/smittsomme-sykdommer/corona/dags--og-ukerapporter

[2] https://www.helsedirektoratet.no/statistikk/antall-innlagte-med-pavist-covid-19-for-nedlasting

A flow chart of the model

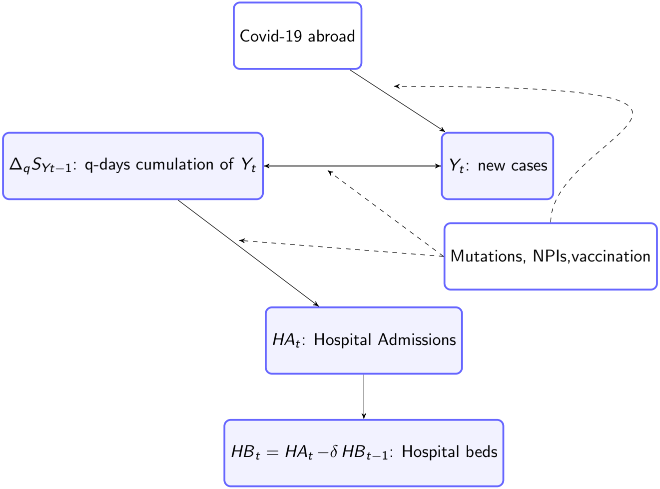


Figure: Variables and relationships of CovidMod. Relationships that may change (structural breaks) as a consequence of virus mutations, NPIs (inducing behavioural changes) and vaccination, are indicated by dotted lines.

Estimation of the CovidMod model

Since March 2021 and until early in May 2022, forecasts were produced five times weekly (ie., work-days). During that period, abrupt and more gradual changes in the dynamics of the pandemic had been represented in the model by step-indicator variables (step dummies) and an by variables that measures the percentage of the population that have received vaccination (one of more doses). The step-indicator variables can be divided in two categories:

- Abrupt exogenous events, like mutations (D).
- Non-pharmaceutical policy responses (G).

Estimation of the CovidMod model (continued)

The estimated equation for new cases Y_t, the empirical counterpart to equation (5), was:

$$
\begin{aligned}
\widehat{Y}_t = & \underset{(0.001)}{0.072}\left(S_{Yt-1} - S_{Yt-14}\right) \\
& + \underset{(0.003)}{0.019}\left(S_{Yt-1} - S_{Yt-14}\right) D_{\alpha t} \\
& + \underset{(0.006)}{0.014}\left(S_{Yt-1} - S_{Yt-14}\right) D_{\delta t} \\
& - \underset{(0.002)}{0.013}\left(S_{Yt-1} - S_{Yt-14}\right) G_{Mar21t} \\
& + \underset{(0.002)}{0.017}\left(S_{Yt-1} - S_{Yt-14}\right) G_{Oct21t} \\
& - \underset{(0.011)}{0.051}\left(S_{Yt-1} - S_{Yt-14}\right) VAC_t \\
& + \sum_{i=0}^{13} \hat{\alpha}_i^{\ddagger} \Delta Y_{t-1-i} + \text{other short term factors}
\end{aligned} \tag{7}
$$

Estimation period: 21.2.2020 – 22.11.2021, 641 obs.

Estimation of the model (continued)

- The estimation method was OLS. Heteroscedasticity and autocorrelation consistent standard errors (Newey-West) are reported in parentheses below the estimated coefficients.
- $D_{\alpha t}$ and $D_{\delta t}$ are two step-dummies that are zero before the Alpha and Delta variants became dominant, and 1 after.
- G_{Mar21t} represents the renewed-lockdown in March 2021.
- G_{Oct21t} is a dummy for a partial opening-up of the society in October 2021.
- VAC_t represents higher vaccination rate: increasing in a continuous manner towards 1.

Estimation of the model (continued)

The second estimated equation in the model is the conditional model of new admissions to hospital with Covid-19 (HA):

$$
\begin{aligned}
\widehat{HA}_t = & \underset{(0.081)}{0.276} HA_{t-1} + \underset{(0.0014)}{0.0035} Y_t \\
& + \underset{(0.00013)}{0.00039} (S_{Yt-3} - S_{Yt-19}) \\
& + \underset{(0.0001)}{0.0004} (S_{Yt-3} - S_{Yt-19}) D_{\alpha t} \\
& + \underset{(0.0001)}{0.0005} (S_{Yt-3} - S_{Yt-19}) G_{Oct21t} \\
& - \underset{(0.0002)}{0.0009} (S_{Yt-1} - S_{Yt-14}) VAC_{t-7} \\
& + \sum_{i=0}^{2} \hat{\gamma}_i \Delta HA_{t-1-i} + \text{other short term factors}
\end{aligned}
\tag{8}
$$

Estimation period: $14.7.2020 - 22.11.2021$, 497 obs.

Estimation of the model (continued)

The third empirical equation in the model estimated on 22 November 2021 was the simple law of motion for the number of hospital beds with Covid-19:

$$\widehat{HB}_t = HA_t + \underset{(0.0029)}{0.898}\ HB_{t-1} \tag{9}$$

$$\text{Estimation period:} 14.7.2020 - 22.11.2021,\ 497 \text{ obs.}$$

- A realistic picture of the DGP of new cases Y and hospital admissions (HA), is that it is complex and changing.
- In comparison the model equations are simple.
- Hence, there are several factors that end up the the error-terms of the estimated equations.
- As mentioned, import of virus trough international travel has played a major role during the pandemic. This factor is not represented in CovidMod and is a source of simulation and forecast errors.

Forecast performance: New cases

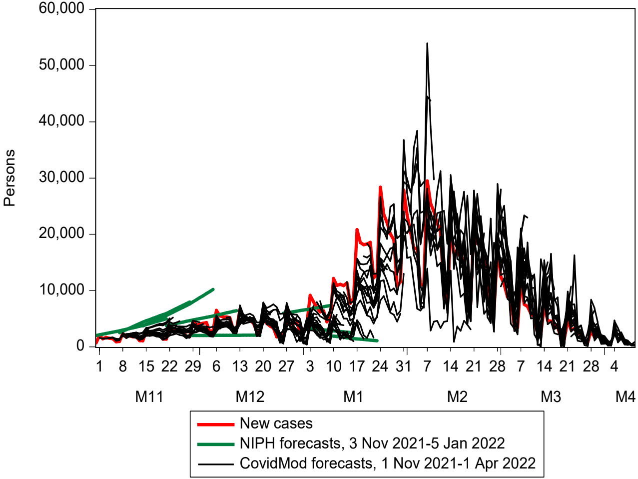


Figure: Forecasted number of new cases (daily data) and forecasts from CovidMod and NIPH (National regional model).

Forecast performance: Hospital beds

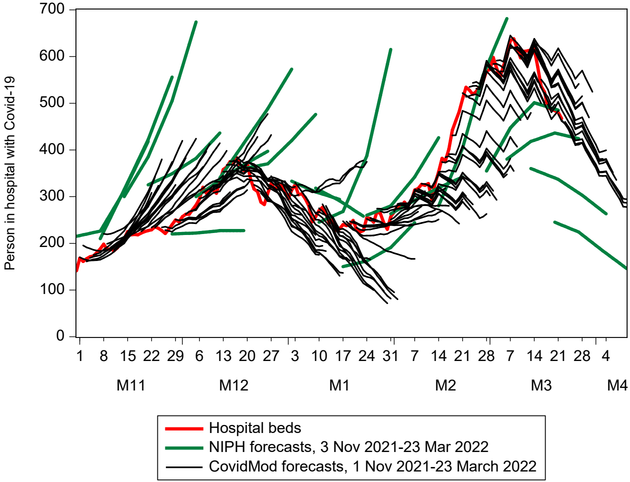


Figure: Forecasted number of hospital beds and forecasts from CovidMod and NIPH (National regional model).

Extending the model with endogenous effects of policy interventions: CovidMod-STR

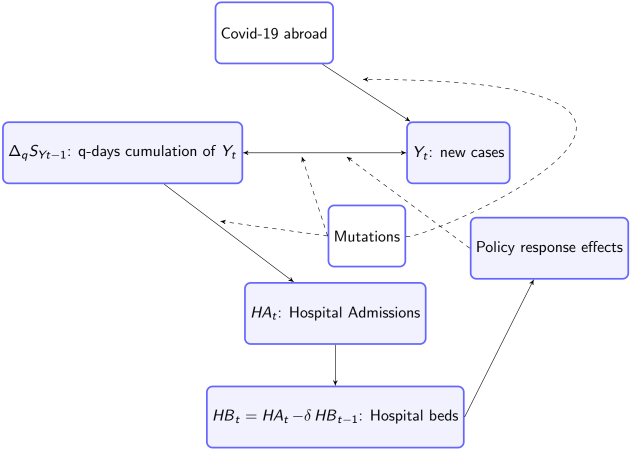


Figure: Variables and relationships of CovidMod-STR.

Extending the model with endogenous effects of policy interventions: CovidMod-STR (continued)

Figure 4 illustrates some of the differences between CovidMod, as illustrated by Figure 1 and the model with endogenous policy response effects, representing all policies that can be modelled as responses to the evolution of hospital beds with Covid-19 patients.

To allow for effects of endogenous policy responses, we divide the breaks or the switching part of the model into

- exogenous regime changes D_t, like mutations,
- and effects of endogenous policy regime changes G_t, like lockdowns.

Extending the model with endogenous effects of policy interventions: CovidMod-STR (continued)

In order to illustrate the approach, consider the following stylized version of equation (5), where $X_t = (S_{Yt-1} - S_{Yt-p})$:

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_t D_t + \beta_3 X_t G_t + u_t \tag{10}$$

To model the endogenous policy effects G_t, we use the standard Smooth Transition Regression model (STR), see e.g. van Dijk et al. (2002) and Teräsvirta et al. (2010, Ch. 3.4). The transition function is the logistic specification

$$G\left(HB_{t-k}; \theta, HB^*\right) = \frac{1}{1 + \exp\left[-\theta\left(HB_{t-k} - HB^*\right)\right]}, \tag{11}$$

that changes smoothly from 0 to 1 with an increasing policy target variable as argument. The policy variable is taken to be hospital beds HB_{t-k}, with a threshold value of HB^*.

Extending the model with endogenous effects of policy interventions: CovidMod-STR (continued)

- Imagine a contagious mutation, represented by $\beta_2 D_t > 0$ in (10), making the binary variable D_t change the effect of X_t on Y_t from β_1 to $\beta_1 + \beta_2$.
- The policy response might be a lockdown, bringing the infections down, with effect represented by $\beta_3 G_t < 0$.
- For illustration purposes, assume that $D_t = G_t = 1$. If then $\beta_3 = -\beta_2$, the effects of policy interventions mitigate the effects of the virus mutation.

Modelling new infections $\boldsymbol{Y_t}$

The equation for Y_t is estimated as:

$$
\begin{aligned}
\hat{Y}_t = \underset{(0.004)}{0.060}\left(S_{Y,t-1} - S_{Y,t-14}\right) \\
+ \underset{(0.002)}{0.007}\left(S_{Y,t-1} - S_{Y,t-14}\right) D_t \\
- \underset{(0.004)}{0.007}\left(S_{Y,t-1} - S_{Y,t-14}\right) G_t \\
+ \sum_{i=0}^{8} \hat{\alpha}_i^{\ddagger} \Delta Y_{t-1-i} + \text{ residual} \\
T = 15.2.2020 - 5.1.2022,\ 691 \text{ obs.}
\end{aligned}
\tag{12}
$$

where

$$
D_t = f(D_{It} + D_{\alpha t} + D_{\delta t} + D_{ot})
$$

$$
G_t = \frac{1}{1 + \exp\left[\underset{(0.032)}{-0.029}\left(HB_{t-5} - \underset{(48.575)}{295}\right)\right]}.
$$

Modelling new infections $\boldsymbol{Y_t}$ (continued)

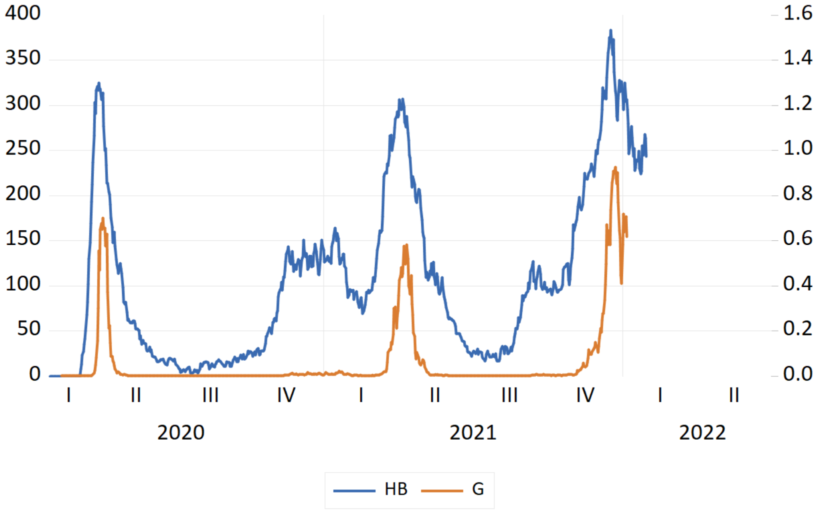
HAC standard errors are reported in parentheses below the estimates.

The step dummies in $D_t = f(D_{It} + D_{\alpha t} + D_{\delta t} + D_{ot})$ represent exogenous breaks due to the occurrence of imported infections and the Alpha, Delta, and Omicron mutations, respectively:

$$D_{It} = 1 \text{ from } 09.03.2020$$
$$D_{\alpha t} = 1 \text{ from } 02.03.2021$$
$$D_{\delta t} = 1 \text{ from } 16.07.2021$$
$$D_{ot} = 1 \text{ from } 28.12.2021$$

- For the policy effects function G_t, the threshold value $HB^{\star}$ is estimated to be 295 hospital beds.
- Note that the effects of exogenous pandemic shocks D_t are mitigated by the endogenous policy effects G_t.
- The chosen specification indicates five days lag in the effects of policy.

Hospitalisations *HB* and smooth transition function *G*



The *HA* and *HB* equations

To complete the model, the estimated versions of *HA* (hospital admissions) and *HB* (hospital beds) equations follows.

$$
\begin{aligned}
\widehat{HA}_t = & -\underset{(0.0001)}{0.0006}\left(S_{Y,t-3} - S_{Y,t-9}\right) D_{o(t-6)} \\
& + \underset{(0.001)}{0.012}\, Y_t \\
& - \underset{(0.001)}{0.007}\, Y_t \times D_{1,t} \\
& - \underset{(0.001)}{0.002}\, Y_t \times D_{2,t} \\
& + \sum_{j=1}^{3} \alpha_j^{\ddagger\ddagger} HA_{t-j} + \text{residual}
\end{aligned}
\tag{13}
$$

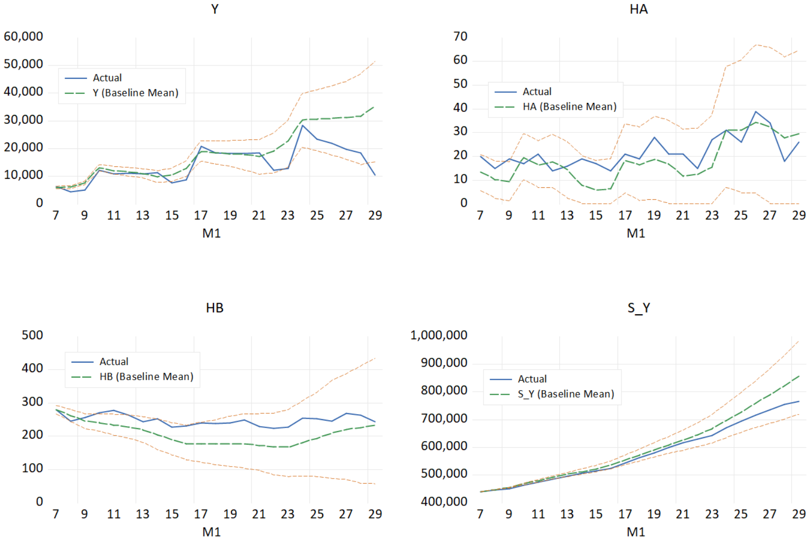
$$
\widehat{HB}_t = HA_t + \underset{(0.004)}{0.901}\, HB_{t-1} \tag{14}
$$

$$
T = 14.7.2020 - 5.1.2022, \; 541 \text{ obs.}
$$

The *HA* and *HB* equations (continued)

$D_{1,t}$ = Step-dummy, 1 from 28 April 2021.
$D_{2,t}$ = Step-dummy, 1 from 15 November 2021.
Although qualitatively similar in terms of variables, the specification in (13) has a more complex structure of the effects of incidence, Y_t, than (8).
A possible interpretation of the dummies might be that more young people became infected, but with a lower frequency of hospital admittance.
Equation (14) is almost the same as (9).

The forecasts and realizations with parameter uncertainty



Comparisons of forecasts

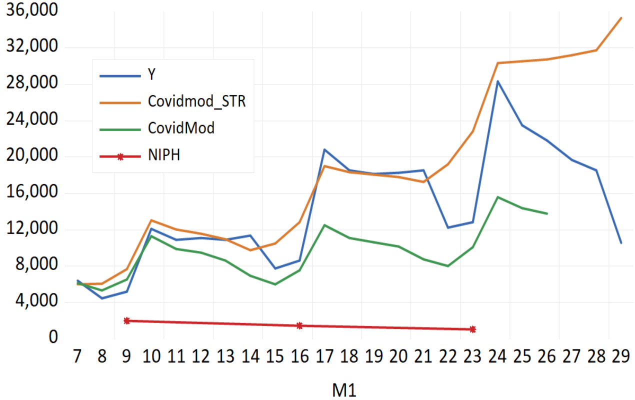


Figure: Incidents. Source: "Situational awareness and forecasting for Norway". FHI COVID-19 modelling team Week 1, 5 January 2022. Table 2. Linear interpolation between 7, 14 and 21 days ahead forecasts from forecast origin date.

Comparisons of forecasts (continued)

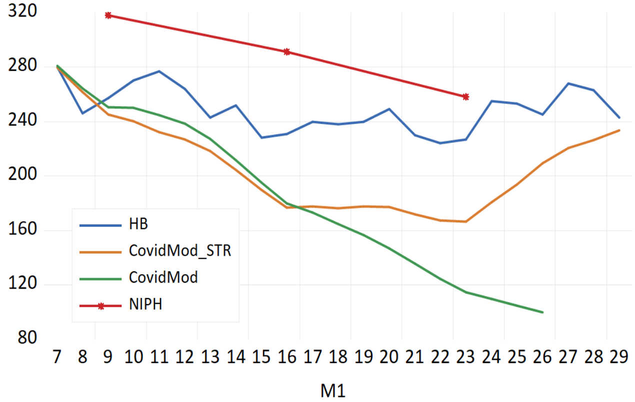


Figure: Hospitalisations. Source: "Situational awareness and forecasting for Norway".FHI COVID-19 modelling team Week 1, 5 January 2022. Table 2 and Figure 5. Linear interpolation between 7, 14 and 21 days ahead forecasts from forecast origin date.

Summary

- When modelling Covid time series, it is likely that constant-parameter models will fail in forecasting.
- A framework for forecasting new cases (incidence), hospital admissions and hospital beds is presented.
- The model is extended to allow for endogenous effects of policy responses.
- The approach, based on forecasting methodology of macroeconomic time series, is to be considered as a complement to other forecasting approaches used.

References

Doornik, J. A., J. L. Castle and D. F. Hendry (2022). Short-term forecasting of the coronavirus pandemic. *International Journal of Forecasting*, *38*(2), 453–466.

Harvey, A. and P. Kattuman (2021). A farewell to *R*: time-series models for tracking and forecasting epidemics. *Journal of The Royal Society Interface*, *18*(182), 20210179.

Li, S. and O. Linton (2021). When will the Covid-19 pandemic peak? *Journal of Econmetrics*, *220*(1), 130–157.

Mills, T. C. (2022). Modelling the link Between Covid-19 ccase, hospital admissions and deaths in England. *National Accounting Review*, *4*(1), 38–58. DOI: 10.3934/NAR.2022003.

Nymoen, R. (2022). Dynamisk modellering og framskrivning av nye smittede og innlagte med Covid-19 i Norge. *Samfunnsøkonomen*, (1), 5–13.

References (continued)

Teräsvirta, T., D. Tjøstheim and C. W. J. Granger (2010). *Modelling Nonlinear Economic Time Series*. Oxford University Press.

van Dijk, D., T. Teräsvirta and P. H. Franses (2002). Smooth Transition Autoregressive Models — A Survey Of Recent Developments. *Econometric Reviews*, *21*(1), 1–47.