Discussion Paper

Analyzing COVID-19 hospitalizations: fostering decision-making in future pandemics

Jonas Klingwort
Joep Burger
Jan van den Brakel

August 22, 2024

Effective and targeted decision-making during pandemics requires accurate forecasting of key health outcomes such as hospitalizations. This study investigates the relationship between the weekly number of COVID-19 hospitalizations and several indicators hypothesized to correlate with COVID-19 hospitalizations, including data from fixed and mobile sensors, wastewater treatment plants (WTPs), weather data, and policy measures in the Netherlands and Germany from 2020 to 2022. The fixed sensors provide data on pedestrian flows in metropolitan areas. The information from the mobile sensors is based on anonymized mobile phone data from apps like Google Maps. The WTP data informs about the coronavirus particles in wastewater. The weather data informs about temperature and humidity. The policy data provides daily policy measures that governments have implemented to address the pandemic. The relationship between the indicators and the weekly COVID-19 hospitalizations is estimated using structural time series (STS) modeling. The STS models decompose the observed hospitalizations into a trend and a regression component. Each time-dependent regression coefficient describes how the relation between an auxiliary series and the hospitalizations evolves during the COVID-19 pandemic. Considered trend components are the local level, smooth trend, and local linear trend models as well as models with a time-invariant intercept. Relevant auxiliary variables for the regression components are selected with a step-forward variable selection method. Models are fitted with the Kalman Filter after expressing them in state space form. The different model outcomes are extensively evaluated and discussed. Thus, besides informing on the potential relevance of indicators to predict COVID-19 hospitalizations, this paper aims to comprehensively compare the various models, their results, and their implications. Our analysis reveals important relations between various indicators and COVID-19 hospitalizations, highlighting the importance of some specific indicators that could be used for modeling and predicting hospitalization during a pandemic. By comprehensively assessing the predictive power of different indicators and model specifications, we foster informed decision-making during pandemics, facilitating more effective public health responses.

Keywords: non-pharmaceutical policy intervention, sensor, structural time series model, virus in wastewater

1 Introduction

Throughout the coronavirus disease 2019 (COVID-19) pandemic, a primary objective was to keep the number of hospitalizations low to prevent the healthcare infrastructure from collapsing. 'Corona Dashboards' were filled with indicators hypothesized to correlate with COVID-19 hospitalizations. However, which of these indicators help predict COVID-19 hospitalizations needs to be studied so that these indicators can be prioritized in future pandemics to foster data-driven decision-making.

To contribute to this question, this paper models the relationship between the weekly number of COVID-19 hospitalizations and indicators from fixed and mobile sensors, wastewater treatment plants (WTPs), weather data, and policy measures in the Netherlands and Germany between 2020 and 2022. The fixed sensors record pedestrian flows within metropolitan areas. The mobile sensors record population mobility with mobile phones. The virus particles contained in the wastewater are recorded. The weather data contains information about temperature and humidity. The policy data informs about several policy measures taken during the COVID-19 pandemic.

A large body of literature on predicting COVID-19 hospitalizations arose since the COVID outbreak in 2019. Many studies have, often for good reasons, limitations in different aspects. They focus on a certain region, or a specific time period of the COVID-19 pandemic, or a single data source for predicting hospitalization, or a single model. These aspects limit the external validity of these many empirical findings, i.e. the extent to which conclusions can be generalized to situation beyond the setting of a specific study. As time proceeds, more data become available. In an attempt to study COVID-19 hospitalization in a broader context we consider in this paper time series data that cover a time window of two-and-a-half years of the COVID-19 pandemic for two countries nation wide. To minimize confounding effects, such as population mobility and weather, a large set of different potential auxiliary variables are considered simultaneously using a general class of dynamic time series models. It is anticipated that this approach increases the external validity of empirical findings related to COVID-19 hospitalization.

It can be anticipated that relations between COVID-19 hospitalization and the auxiliary series that drive COVID-19 hospitalization change over time. Therefore a structural time series modeling approach is adopted since this class of models allow to define time-varying regression coefficients for the relationships between the auxiliary series and hospitalization during the COVID-19 pandemic. In this paper models with time-varying as well as time-invariant regression coefficients in combination with different trend models are applied with the purpose to shed light on the factors that drive COVID-19 hospitalization. The different model outcomes are extensively evaluated and discussed. Thus, besides informing on the potential relevance of indicators to predict COVID-19 hospitalizations, this paper aims to comprehensively compare the various models, their results, and their implications.

We provide evidence for which structural time series to choose and which indicators might predict hospitalizations during a future pandemic or other scenarios in which low hospitalization rate is essential. Therefore, we consider these results of practical importance for statisticians, data scientists, practitioners, and policymakers.

The paper is structured as follows. The research background is given in Section 2. Section 3 explains the methodology and related aspects. Section 4 with additional material in the Appendix describes the data used in this study. The results are presented in Section 5 with additional material in the Appendix. A discussion of the paper is given in Section 6. The paper concludes with Section 7.

2 Research Background

Non-pharmaceutical policy interventions (NPIs) were implemented during the COVID-19 pandemic in many countries to reduce infection rates and to avoid collapsing medical infrastructures (Perra 2021). Although many papers have studied the effects of NPIs, few have focused on hospitalizations. In a meta-analysis by Peters and Farhadloo (2023), only 4 of 44 papers looked at hospitalizations, and only 2 NPIs (shelter-in-place orders and mask-wearing) were studied, showing lower hospitalization rates 3-4 weeks after implementation. Another meta-analysis by Iezadi et al. (2021) also found only two studies on the effect of NPIs on COVID-19 intensive care unit admission, showing an overall reducing effect. Perra (2021) reviewed 348 papers on NPIs during the COVID-19 pandemic, 43 of which were on the efficacy of NPIs in the mitigation/suppression of COVID-19, but none on hospitalizations. A systematic review by Mendez-Brito et al. (2021) included 34 papers and 16 NPIs, but the hospitalization rate was not considered an outcome of interest.

In addition to pharmaceutical and non-pharmaceutical policy data, other features and their relations with COVID-19-related variables have been studied. For example, the effects of weather on COVID-19-related variables were studied by Ganslmeier et al. (2021) and Paraskevis et al. (2021). Turk et al. (2021), Lampos et al. (2021), and Wang et al. (2022) utilized Google Trends in their modeling approaches. The wastewater and the contained virus RNA level were used in studies by Twigg and Wenk (2022), Mantilla-Calderon et al. (2022), Li et al. (2023), and Hetebrij et al. (2024). Other objective measurements such as insurance claims, census information, and hospital resource usage were utilized in the study by Gao et al. (2023). In this article, we will use another objective feature based on measurements from sensors installed on facades in city centers that provide information about the mobility of the population during the COVID-19 pandemic (Klingwort et al. 2024).

Different modeling approaches have been used to predict short-term and long-term COVID-19 hospitalizations. For example, Zoest et al. (2024) used a negative binomial regression model. Other approaches are based on the SEIR models (Susceptible, Exposed, Infectious, Removed), for example Reno et al. (2020), Gerlee et al. (2021), and Gatto et al. (2021). Time series approaches are given by Perone (2022) and Klein et al. (2023). However, a systematic review on the effectiveness of NPIs found only 3 of 248 studies applying time series modeling (Banholzer et al. 2022).

As mentioned in the introduction, the literature on the effect of COVID-19 on the hospitalization rate is rather limited. This is mainly because there are not sufficiently long time series available from different regions or countries, sufficient explanatory variables, or model comparisons.

3 Methodology

3.1 Structural time series modeling

Let $y_t, t = 1, \dots, T$ denote a time series containing the average number of patients hospitalized due to a coronavirus infection observed on a weekly frequency. Structural time series models are developed to analyze the relation between the number of COVID-19 hospitalizations and auxiliary series that might explain or forecast the number of hospitalizations. A detailed description of the target variable can be found in Subsection 4.1. The auxiliary variables are described in detail in Subsections 4.2 until 4.7.

With a structural time series model, an observed time series is decomposed into a trend component, a seasonal component, other cyclic components, a regression component, and an irregular component. For each component, a stochastic model is assumed. This allows the trend, seasonal, and cyclic components but also the regression coefficients to be time-dependent. If necessary, ARMA components can be added to capture the autocorrelation in the series beyond these structural components. See Harvey (1989) and Durbin and Koopman (2012) for details about structural time series modeling.

The weekly hospitalization counts are described with a structural time series model that contains a trend and regression components:

$$y_t = \mu_t + \boldsymbol{x}_t^{'}\boldsymbol{\beta}_t + \epsilon_t \quad (1)$$
$$\epsilon_t \sim N(0, \sigma_\epsilon^2)$$

with μ_t the level of the trend component, $\boldsymbol{x}_t = \begin{bmatrix} x_{t,1} & \cdots & x_{t,q} \end{bmatrix}'$ a q vector with auxiliary variables, $\boldsymbol{\beta}_t = \begin{bmatrix} \beta_{t,1} & \cdots & \beta_{t,q} \end{bmatrix}'$ a q vector with time-varying regression coefficients and ϵ_t a measurement error that is independent and normally distributed. In this paper, dynamic trend models are considered that have the flexibility to capture cyclic patterns. Therefore, an additional seasonal component for this weekly series could be omitted in (1).

Different specifications for the trend component are considered. The most general specification is the so-called local linear trend (LLT) model:

$$\begin{aligned} \mu_t &= \mu_{t-1} + \nu_{t-1} + \eta_{t,\mu}, \\ \nu_t &= \nu_{t-1} + \eta_{t,\nu}, \\ \eta_{t,\mu} &\sim N(0, \sigma_\mu^2), \\ \eta_{t,\nu} &\sim N(0, \sigma_\nu^2) \end{aligned} \tag{2}$$

In (2) μ_t is referred to as the level of the trend while ν_t can be interpreted as the slope of the trend. A special case of the LLT model is the so-called smooth trend (ST) model, which is obtained by (2) with $\sigma_\mu^2 = 0$. This implies that there is no level disturbance term $\eta_{t,\mu}$ for μ_t. If both σ_μ^2 and σ_ν^2 are equal to zero, then the LLT reduces to linear trend, i.e. $y_t = \mu + \nu * t$, with μ a fixed intercept and ν a fixed slope.

Another trend specification considered in this paper is the local level (LL) model, which is based on a random walk for the level of the trend:

$$\begin{aligned} \mu_t &= \mu_{t-1} + \eta_{t,\mu}, \\ \eta_{t,\mu} &\sim N(0, \sigma_\mu^2). \end{aligned} \tag{3}$$

In the case that $\sigma_\mu^2 = 0$, there is no level disturbance term and the LL model is equal to a horizontal straight line, i.e. $y_t = \mu$, with μ a fixed intercept or constant level (CL).

The regression coefficients are modeled time-varying. This is achieved by assuming a random walk for each regression coefficient:

$$\begin{aligned} \beta_{t,k} &= \beta_{t-1,k} + \eta_{t,\beta_k}, \\ \eta_{t,\beta_k} &\sim N(0, \sigma_{\beta_k}^2). \end{aligned} \tag{4}$$

In the case that $\sigma_{\beta_k}^2 = 0$, the regression coefficient becomes time-invariant, since $\beta_{t,k} = \beta_{t-1,k} \equiv \beta_k$.

To fit structural time series model (1) with the Kalman filter, it is expressed as a state space model that consists of a measurement equation and a transition equation. The measurement equation states how the observed series depends on the trend and the regression components and is defined as:

$$\begin{aligned} y_t &= \mathbf{Z}_t \boldsymbol{\alpha}_t + \epsilon_t \\ \epsilon_t &\sim N(0, \sigma_\epsilon^2) \end{aligned} \tag{5}$$

with $\boldsymbol{\alpha}_t$ a p-dimensional vector with the state variables, i.e. μ_t, ν_t and $\beta_{t,k}$ and $\mathbf{Z}_t$ a $1 \times p$ design matrix that defines how the observed series depends on the state variables in $\boldsymbol{\alpha}_t$. The transition equation describes how the state variables evolve from one period to the next period and is defined as:

$$\begin{aligned} \boldsymbol{\alpha}_{t+1} &= \mathbf{T}\boldsymbol{\alpha}_t + \mathbf{R}\boldsymbol{\eta}_t \\ \boldsymbol{\eta}_t &\sim N(\mathbf{0}_p, \boldsymbol{\Sigma}_\eta) \end{aligned} \tag{6}$$

with $\mathbf{T}$ a $p \times p$ design matrix, $\boldsymbol{\eta}_t$ an l-dimensional vector with state disturbance terms, i.e. $\eta_{t,\mu}$, $\eta_{t,\nu}$ and $\eta_{t,\beta}$, $\mathbf{R}$ a $p \times l$-dimensional selection matrix that selects the appropriate disturbances from $\boldsymbol{\eta}$ and $\boldsymbol{\Sigma}_\eta$ an $l \times l$ covariance matrix for the state disturbance terms $\boldsymbol{\eta}_t$.

Furthermore let $\boldsymbol{\eta}_{t,\beta} = \begin{bmatrix} \eta_{t,\beta_1} & \cdots & \eta_{t,\beta_q} \end{bmatrix}'$ denote a vector with regression coefficient disturbance terms and $\boldsymbol{\Sigma}_\beta = \oplus_{k=1}^{q} \sigma^2_{\beta_k}$ a diagonal covariance matrix for the regression coefficient disturbance terms $\boldsymbol{\eta}_{t,\beta}$. Note that $\oplus$ denotes the direct sum, that defines a (block-)diagonal matrix. Finally, let $\mathbf{I}_k$ denote the identity matrix of size k, and $\mathbf{0}_k$ is a k-dimensional vector of zeros. Now the state space representation of the different trend models in combination with regression components is specified in Table 3.1.

Table 3.1 Specification of arrays in measurement and transition equation by model

Model	$\mathbf{Z}_t$	$\boldsymbol{\alpha}_t$	$\mathbf{T}$	$\mathbf{R}$	$\boldsymbol{\eta}_t$	$\boldsymbol{\Sigma}_\eta$
CL	$\boldsymbol{x}_t'$	$\boldsymbol{\beta}_t$	$\mathbf{I}_q$	$\mathbf{I}_q$	$\boldsymbol{\eta}_{t,\beta}$	$\boldsymbol{\Sigma}_\beta$
LL	$\begin{bmatrix} \boldsymbol{x}_t' & 1 \end{bmatrix}$	$\begin{bmatrix} \boldsymbol{\beta}_t' & \mu_t \end{bmatrix}'$	$\mathbf{I}_{q+1}$	$\mathbf{I}_{q+1}$	$\begin{bmatrix} \boldsymbol{\eta}_{t,\beta}' & \eta_{t,\mu} \end{bmatrix}'$	$\boldsymbol{\Sigma}_\beta \oplus \sigma^2_\mu$
ST	$\begin{bmatrix} \boldsymbol{x}_t' & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} \boldsymbol{\beta}_t' & \mu_t & \nu_t \end{bmatrix}'$	$\mathbf{I}_q \oplus \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$	$\mathbf{I}_q \oplus \begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} \boldsymbol{\eta}_{t,\beta}' & \eta_{t,\nu} \end{bmatrix}'$	$\boldsymbol{\Sigma}_\beta \oplus \sigma^2_\nu$
LLT	$\begin{bmatrix} \boldsymbol{x}_t' & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} \boldsymbol{\beta}_t' & \mu_t & \nu_t \end{bmatrix}'$	$\mathbf{I}_q \oplus \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$	$\mathbf{I}_{q+2}$	$\begin{bmatrix} \boldsymbol{\eta}_{t,\beta}' & \eta_{t,\mu} & \eta_{t,\nu} \end{bmatrix}'$	$\boldsymbol{\Sigma}_\beta \oplus \sigma^2_\mu \oplus \sigma^2_\nu$

In the case of time-varying (dynamic) regression coefficients, $\boldsymbol{\Sigma}_\beta = \oplus_{k=1}^{q} \sigma^2_{\beta_k}$. In case of time-invariant (static) regression coefficients, $\boldsymbol{\Sigma}_\beta = 0\mathbf{I}_q$, i.e. $\sigma^2_{\beta_k} = 0$ for all $k = 1, \ldots, q$.

3.2 Kalman filtering and smoothing

Once the model is in state space form, the Kalman filter can be applied to estimate the state variables. The state space representation distinguishes between state variables and hyperparameters. The state variables define the trend μ_t, ν_t and the regression coefficients $\beta_{t,k}$. The hyperparameters define the dynamics of the processes for the state variables, which are the variance components of the state disturbance terms, i.e. σ^2_μ, σ^2_ν and $\sigma^2_{\beta_k}$ and the variance of the measurement errors, i.e. σ^2_ϵ.

The Kalman filter is a recursive algorithm that starts at the beginning of the series and provides optimal estimates and standard errors for the state variables for each period t based on the time series observed until period t. These are referred to as the Kalman filter estimates. The Kalman filter estimates, can be updated with the information that became available after period t. This procedure is called smoothing and is based on a recursive algorithm that starts with the last observation of the observed series and updates the filtered estimates, including their standard errors, for the state variables of all preceding periods. These are referred to as the Kalman smoother estimates. Under the assumption that the disturbance terms and the initial state vector are normally distributed, the Kalman filter provides optimal estimates in the sense that they minimise the mean squared error. If the normality assumption doesn't hold, the Kalman filter is still an optimal estimator in the sense that it minimise the mean squared error within the class of all linear estimators (Harvey 1989, Section 3.2).

To start the Kalman filter, initial values for the state variables as well as values for the hyperparameters are required. The models defined in Subsection 3.1 contain non-stationary state variables, which are initialized with a diffuse initialization. This implies that the initial values of all state variables are equal to zero with a diagonal covariance matrix with diagonal elements diverging to ∞. The exact initial solution for the Kalman filter with diffuse initial conditions, proposed by Koopman (1997), is used. With this diffuse initialization of the Kalman filter the first d observations are required to construct a proper prior for the Kalman filter, where d equals the

number of non-stationary state variables of the state space model. For this reason, the Kalman filter estimates for the first d time periods are ignored in the analysis and also in the model evaluation of the one-step-ahead prediction errors.

The values of the hyperparameters are also unknown. They are estimated by means of maximum likelihood (Harvey 1989, Section 3.4). The likelihood function is optimized by repeatedly running the Kalman filter in a numerical optimization procedure using MaxBFGS. Since the hyperparameters are variances, which cannot take negative values, they are estimated on the log-scale. The unknown values of the hyperparameters of the state-space model are replaced by their maximum likelihood estimates in the Kalman filter. The standard errors of the Kalman filter estimates for the state variables do not reflect the additional uncertainty of replacing the true values of the hyperparameters by their maximum likelihood estimates, which is the common approach in state space time series analysis.

3.3 Forward variable selection

The auxiliary variables for the regression components are selected with a forward variable selection procedure that minimizes Akaike Information Criterion. See Durbin and Koopman (2012), Section 7.4 for a definition of the AIC for state space models. This model selection technique is used to select the most appropriate variables for the eight models. It is opted for a model that explains the data well while avoiding overfitting. A separate model selection procedure was applied to each model.

It was decided to begin with the null model, add features to the model iteratively, and evaluate the effect on the AIC. The model with the lowest AIC is selected as the final model. The selected final models are shown in Figures B.1 and B.2.

3.4 Model diagnostics

The stated normality assumption in Subsection 3.2 implies that the one-step-ahead prediction errors are independent and normally distributed. This is evaluated by testing the standardized one-step-ahead prediction errors for 1) normality using a Bowman-Shenton test, 2) heteroscedasticity using an F-test of equality of variances between the first and second half of the time series, and 3) autocorrelation using a Ljung-Box test for 21 lags (Durbin and Koopman 2012, Section 2.12). The following visualizations are made: 1) time series of the standardized innovations, 2) QQ-plot of the standardized innovations, 3) histogram of the standardized innovations, and 4) autocorrelation plot of the standardized innovations.

3.5 Software

All computations are performed in R. The state space models are implemented using the R package KFAS (Helske 2017).

4 Data

This section describes the data used in this study. All data are visualized in Figures 4.1 and 4.2. In the Appendix, lists with all used features and additional explanations are given in Tables A.1 and A.2.

4.1 COVID-19 hospitalizations

The target variable of this study is the average weekly number of COVID-19 hospitalizations. The German data was obtained from the GitHub repository of the Robert Koch-Institut (2023a) and the Dutch data from the Dutch governmental corona dashboard (Rijksoverheid 2023a). The period considered for this study is 138 weeks (W10 2020 – W43 2022) for Germany, and 140 weeks (W09 2020 – W44 2022) for the Netherlands. The daily number of hospitalizations was obtained and averaged per year and week.

4.2 Weather

Weather data were obtained from the German national weather service and the Royal Dutch Meteorological Institute (KNMI 2023). The data for the first was downloaded using the R-library developed by Boessenkool (2023). The daily temperature and relative humidity were averaged per year and week.

4.3 Policy stringency

The Oxford COVID-19 Government Response Tracker (OxCGRT) provides daily policy measures that governments have implemented to address the pandemic (Hale, Angrist, et al. 2021; Hale, Anania, et al. 2023). It provides information on 24 policy indicators: containment and closure policies (C), economic policies (E), health system policies (H), and vaccination policies (V). Four indices aggregate the 24 indicators: overall government response index (all indicators), containment and health index (all C and H indicators), stringency index (all C indicators, plus H1), and economic support index (all E indicators). More details can be found in Hale, Anania, et al. (2023). The daily numbers of all indicators and indices were averaged per year and week.

4.4 Mobile sensors

The Google COVID-19 Mobility Reports were provided during the pandemic (Google 2023). These reports are based upon anonymized mobile phone data provided by apps such as Google Maps, of users that have turned on the location history setting. The value in this data represents the daily comparison of visitors relative to a baseline day. The baseline day is based on the five weeks from January 3 to February 6, 2020. The data contains the change in total visitors provided for retail and recreation, grocery and pharmacy, parks, transit stations, and workplaces. The change in length of stay is provided for places of residence.

4.5 Local sensors

The pedestrian flows in metropolitan areas are recorded by location-based sensors and are provided by the company *hystreet.com GmbH*. The sensors are located in metropolitan areas, economically relevant areas, and highly frequented streets and places. The sensors are attached to facades and generate an invisible, eye-safe quadruple light curtain to measure pedestrian flows on a minute level. Pedestrians are not aware of being recorded, so consciously avoiding the recording is unlikely. The data do not contain individual information, only aggregate counts. Details on the methodology can be found in hystreet GmbH (2022). The data are used as a proxy for population mobility in this study, and 82 cities in Germany with 197 locations were selected. The sensor system is also installed in the Netherlands, but this data is not sufficiently available for the period considered in this study. The daily number of pedestrian counts was averaged per year and week.

4.6 Wastewater treatment plants

Wastewater is collected at treatment plants in the Netherlands and tested for SARS-CoV-2 RNA particles (Rijksoverheid 2023b). The daily average number of coronavirus particles per 100,000 inhabitants is obtained. One of the main sources of virus particles in wastewater is the stool of infected people. Monitoring of the wastewater is also done in Germany (Robert Koch-Institut 2023b), but this data is publicly available only since week 22 in 2022, thus it is not sufficiently available for the period considered in this study.[1)]

4.7 Surveys

The survey data contains responses on risk perception, protective behavior, supportive behavior, and trust during the COVID-19 pandemic. For Germany, a non-probability stratified sample from an online consumer/access panel was used. The sample size is $n \approx 1\,000$ and data was collected at weekly to four-weekly intervals (COSMO 2023). For the Netherlands, the data is also based on an online consumer/access panel. The sample size is $n \approx 5\,000$ and data was collected at three-weekly intervals (RIVM 2023). Both surveys contained several questions on the same topic. To reduce the number of features, the domain-specific questions with identical response scales were averaged.

Initially, the intention was to utilize the survey data. However, several problems were encountered with this data (in both countries), so it was decided not to use survey features in the models, but the data is still shown in Figures 4.1 and 4.2. This point will be addressed again in the discussion (Section 6).

[1)] see the project AMELAG: Wastewater Monitoring for Epidemiological Situation Assessment `https://www.rki.de/EN/Content/Institute/DepartmentsUnits/InfDiseaseEpidem/Div32/WastewaterSurveillance/WastewaterSurveillance.html`.

4.8 Visualized time series data

Figures 4.1 and 4.2 show the observed weekly averaged hospitalizations y_t in red and the 42 independent variables $x_{t,j}$ in theme-specific colors (weather in blue, policy in green, mobile sensors in purple, local sensors in orange (Germany), wastewater virus particles in orange (Netherlands), and survey in pink) for Germany and the Netherlands. The x-axis shows the time (in weeks), and the y-axis shows the measurements (min-max normalized for visual comparison).

The weekly hospitalizations show different developments for both countries. In Germany, the maximum was reached around the end of 2020 and the beginning of 2021, while the Netherlands had the maximum in mid-2020. It also shows that weekly hospitalizations are lower in summer than in winter. The weekly average temperature and humidity show a comparable pattern across countries, with clearly seasonal effects.

In the case of the policy measures, some show clear variation over time (e.g., school closing or internal movement restrictions), others somewhat less (e.g., contact tracing), and still, others were only in place for a very short time and were then no longer used (e.g., investment in healthcare) or others were 'activated once' and were then constantly in place (e.g., public information campaigns).

For mobile sensors, for example, the change in total visitors in parks is comparable with the temperature series, while retail and recreation or residential sensors show a time series that more closely follows the development of hospitalizations. For Germany, the local sensors also mirror to some extent the developments in hospitalizations: low pedestrian counts in periods with high hospitalizations and vice versa. For the Netherlands, peaks and troughs in the virus particles in the wastewater roughly coincide with those in hospitalizations, although the relative number of particles increases over time. For both countries, the survey features show a lower frequency of observations and a shorter time series. In the Appendix, a list of all the variables and descriptions is shown.

Figure 4.1 Observed time series for Germany.

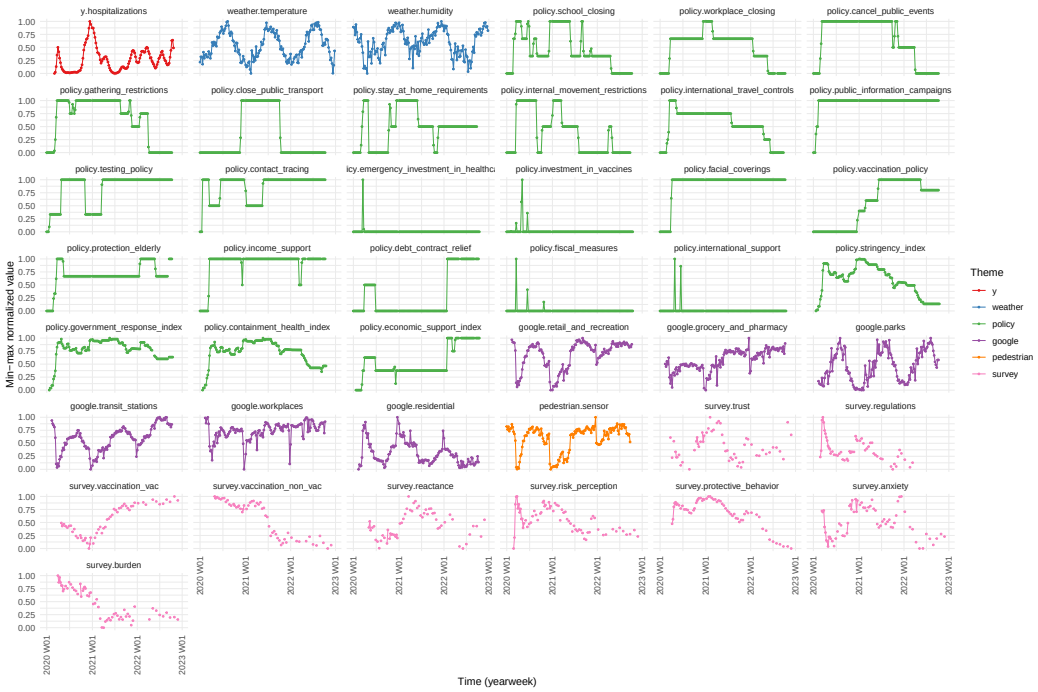
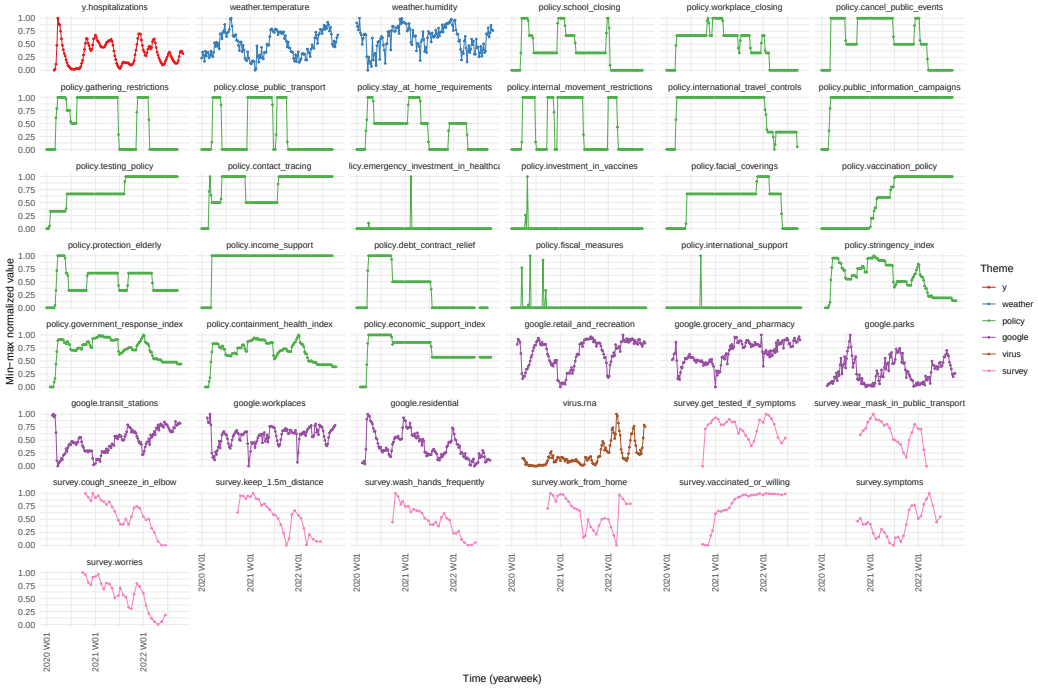


Figure 4.2 Observed time series for the Netherlands.



4.9 Pre-processing

The data described previously was pre-processed to use not only the original features but also a lagged and leading version. The idea behind this is to predict the present with a value from the past (lag) and to predict the present with a value from the future (lead). For both the lagged and the leading versions, four weeks of lags/leads were used. Hence, each feature has nine versions (one original, four lagged, four lead). The optimal shift $u \in \{-4, \dots, +4\}$ was determined per feature by fitting a constant level model with the kth feature ($q = 1$, $Z_t = x_{k,t+u}$ and $\alpha_{t+u} = \beta_{k,t+u}$) with a static ($\sigma^2_\beta = 0$) or dynamic ($\sigma^2_\beta > 0$) regression coefficient. For comparability, the time series started four weeks later ($u = -4$) and ended four weeks earlier ($u = +4$) than the original series. The shift u that gave the lowest negative log-likelihood per feature was chosen for subsequent analyses. The features with only a single event (emergency_investment_in_healthcare, investment_in_vaccines, fiscal_measures and international_support) and features that were switched on once (public_information_campaigns, facial_coverings and income_support) were omitted (see Figures 4.1 and 4.2).

5 Results

Table 5.1 summarizes model diagnostics by country (DE = Germany, NL = the Netherlands), the type of β (static or dynamic), and the trend model (CL, LL, ST, or LLT). For each combination the table contains the AIC value (the AIC value is divided by the sample size for comparability across countries), number of regression coefficients of the auxiliary series selected with a step-forward variable selection procedure, check marks whether a model diagnostics test failed to reject the null hypothesis, how many model diagnostics tests failed to reject the null hypothesis, and finally whether the chosen features of a model can be interpreted meaningfully

Table 5.1 Summary of model diagnostics by country, β and model. Checkmark (✓) indicates null hypothesis of normality, no heteroscedasticity, or no auto correlation could not be rejected.

Country	β	Model	AIC	No. of βs	normality	heteroscedasticity	autocorrelation	No. of ✓'s	interpretable
DE	static	CL	1.0	11	✓	✓		2	
DE	static	LL	0.0	4	✓			1	yes
DE	static	ST	−0.4	2			✓	1	
DE	static	LLT	−0.4	2			✓	1	
DE	dynamic	CL	−0.2	2	✓	✓		2	
DE	dynamic	LL	−0.3	3	✓	✓		2	
DE	dynamic	ST	−0.6	3		✓	✓	2	
DE	dynamic	LLT	−0.6	3		✓	✓	2	
NL	static	CL	1.6	5		✓		1	yes
NL	static	LL	−0.2	4		✓		1	yes
NL	static	ST	−0.7	1		✓		1	yes
NL	static	LLT	−0.7	1		✓		1	yes
NL	dynamic	CL	−0.4	6			✓	1	
NL	dynamic	LL	−0.4	4			✓	1	
NL	dynamic	ST	−0.8	1		✓		1	yes
NL	dynamic	LLT	−0.8	1		✓		1	yes

For Germany, the AIC points towards choosing the ST and LLT models although the differences are small. The results of the feature selection using AIC can be found in Appendix B (Fig. B.1). Regarding the model diagnostics tests of the standardized innovations, the models fail either the normality test (ST and LLT) or the autocorrelation test (CL and LL). The models with static βs fail the heteroscedasticity test, except the CL trend model. There is no model successfully passing all diagnostics tests (maximum is two out of three), suggesting there is a trade-off between the different diagnostics. The diagnostics test results, summarized in this table, are visualized in the Appendix C. The Bowman-Shenton test for normality is shown in Figure C.1, the F-test for heteroscedasticity is shown in Figure C.2, and the Ljung-Box test for autocorrelation in Figure C.3. The Appendix D shows for Germany the standardized innovations (Fig. D.1), their Q-Q plots (Fig. D.2), their histograms (Fig. D.3) and their autocorrelations (Fig. D.4).

The smoothed estimates for the states are presented in Appendix F. A visual inspection of the states (Fig. F.1) reveals that the vast majority of features chosen are from the policy theme. From these, the testing policy, gathering restrictions, protection of the elderly, and the stringency index are chosen most often. The figure points to an LL model with static βs, three of which are selected: a positive coefficient for the stringency index four weeks ahead, a positive coefficient for the protection of the elderly, and a negative coefficient for internal movement restrictions four weeks ago. Note that selecting a feature a few weeks ago (lag) implies a causal effect, while selecting a feature one or more weeks ahead (lead) could imply a policy reaction to an increased infection rate. The random walk of the level of the LL model follows the movements of the observed target series. A CL model with static βs contains no less than eleven features from all themes, possibly because it lacks flexibility, particularly for the trend (supported by its high AIC). An LL model with dynamic βs seems to result in a model that overfits the observed series with too volatile coefficients that do not have a meaningful interpretation. The ST and LLT models contain a random slope variable that seems redundant. The ST and LLT models with dynamic βs initially showed a positive relation between the policy stringency index a week ago and the hospitalizations, but it disappeared quickly during the pandemic. A high hospitalization rate followed by a high stringency index four weeks later, suggested by the LL model with static βs, seems more plausible.

For the Netherlands, the AIC also points towards choosing the ST and LLT models although the differences are small. The feature selection results using AIC can be found in Appendix B (Fig. B.2). Regarding the model diagnostics tests of the standardized innovations, all models fail the normality test. Most models pass the heteroscedasticity test, except CL and LL with dynamic βs. However, these two are the only models that pass the autocorrelation test. No model is successfully passing all diagnostics tests, and only one test, at maximum, is passed per model.

The diagnostics test results, which are summarized in Table 5.1, are visualized in the Appendix C. The Bowman-Shenton test for normality is shown in Figure C.1, the F-test for heteroscedasticity is shown in Figure C.2, and the Ljung-Box test for autocorrelation in Figure C.3. The Appendix E shows for the Netherlands the standardized innovations (Fig. E.1), their Q-Q plots (Fig. E.2), their histograms (Fig. E.3) and their autocorrelations (Fig. E.4).

A visual inspection of the smoothed estimates of the states (Fig. F.2 in Appendix F) reveals that all models contain the number of virus particles in wastewater, a feature not available in Germany. The amount of virus is always a lagged version of one or two weeks ago, suggesting this feature could be a potential predictor of hospitalizations. The figure points again to an LL model with static βs, four of which are selected: a positive coefficient for the amount of virus in wastewater two weeks ago, a negative coefficient for the economic support index four weeks ago, a positive coefficient for Google mobility at places of residence one week ago, and a negative coefficient for Google mobility at transit stations in the current week. This model is the only one where a policy feature's lagged (predictive) version was selected. The random walk of the level of the LL model is sufficiently flexible to model the variation in the target series. A CL model with static βs contains five features, of which only the amount of virus in wastewater matches the LL model. This model has the highest AIC and can be seen as a benchmark to the other, more complex models. An LL model with dynamic βs again seems to result in too volatile coefficients, which do not have a meaningful interpretation. The ST and LLT models contain a random slope variable that again seems redundant.

In summary, none of the models passed all diagnostic tests, but the local level (LL) model with static βs seems to yield the most plausible and interpretable model for both the German and Dutch data. For the German data, only three features were selected, all policy measures, only one with a lagged version: internal movement restrictions. For the Dutch data, four features were selected from three themes, three of which had a lagged version: virus in wastewater, economic support policy, and Google mobility in residential areas. Weather and pedestrian flow in German metropolitan areas have little added value.

6 Discussion

This discussion chapter will answer the research question, outline the research's importance and relevance, address the study's limitations, and give recommendations for future research.

The main goal of this study was to analyze the relationship between the weekly number of COVID-19 hospitalizations and several indicators hypothesized to correlate with COVID-19 hospitalizations. Structural time series modeling was used to estimate the relationship between the indicators and the weekly COVID-19 hospitalizations. The constant level, local level, smooth trend, and local linear trend models were fitted using static or dynamic regression coefficients for the auxiliary variables. The indicators include data on weather, policy measures, mobility from mobile and local sensors, and virus in wastewater. The analyses were conducted with data from Germany and the Netherlands during two-and-a-half years of the COVID-19 pandemic (2020–2022). According to the models based on the German data, out of eight models, only the LL with constant regression coefficients was found appropriate considering AIC, diagnostic tests, visual inspection, and interpretability. For the Netherlands, most models did not pass the diagnostic tests, but most models could be interpreted meaningfully. The virus particles made a major contribution to interpretability here. This feature was selected in all models and was the only relevant feature in the ST and LLT models.

Several aspects drive the importance and relevance of the research. First, it is conjectured that the results in this study have a larger external validity compared to earlier studies since time series data are used that cover a larger time window of the COVID-19 pandemic, and a larger geographical area. Using a larger set of auxiliary data simultaneously and using a broad set of dynamic time series models also increases the internal validity of the study. The insights about the model selection and comparison especially reveal that it matters which model is chosen. Most models contained a different set of selected features, leading to different interpretations of content, and accordingly, different conclusions would be drawn. This shows how important it is to compare and test the models' assumptions.

Second, the models in this study suggest that there is a trade-off in model diagnostics between normality and autocorrelation in the German data, and between heteroscedasticity and autocorrelation in the Dutch data. In practice, a decision must be made and prioritized in the specific case as to which diagnostics must be fulfilled.

Third, most models for both countries showed a tendency for features from the policy theme to be good predictors for the weekly COVID-19 hospitalizations. However, a lead version of a policy feature was often chosen as the best predictor for COVID-19 hospitalizations in the current week, suggesting a reactive policy. A lagged version of a policy feature would suggest a preventive/predictive policy. Another problem is that lead versions are not available at present, and analysis with these features can only take place retrospectively. However, this also shows how important this work is for future political decisions.

Fourth, the choice of mainly policy theme features also suggests that most of the other indicators used in this work do not have predictive power to predict weekly COVID-19 hospitalizations. Thus, whether these indicators should be prioritized in future pandemics is questionable. It seems particularly useful to rely on indicators such as virus particles from the wastewater for predicting weekly COVID-19 hospitalizations in future pandemics. Accordingly, other environmental samples from soil and sediment might also be considered. However, it may also mean we have not included all relevant features in our study. For example, other features such as COVID-19 cases, testing rate, positive test rate, vaccination effectiveness and coverage, population demographics, or virus variant prevalence could be included in these models.

Next, it will be elaborated on the limitations and recommendations for future research. One limitation of this study is that the comparability of the data content from the two countries is only given for some features. The fact that no wastewater treatment plant data for Germany was available for the period under study is regrettable, given the findings for the Dutch data regarding this feature. A recommendation for future research is to include the survey data in the models. The survey data contains low-frequent and subjective measurements. It would be valuable to compare whether this data helps predict the weekly number of COVID-19 hospitalizations. However, the data collection started later and ended earlier than the period considered in this study. All other time series would have to be shortened to use this data, resulting in information loss. For example, the survey time series in both countries starts after the first peak of hospitalizations. Moreover, the data is not available on a weekly frequency. Therefore, more complex methods such as mixed-frequency time series models must be considered to combine weekly data with survey data measured on a frequency of three weeks. Alternatively, more pragmatic alternatives like imputation or interpolation could be considered to impute the weeks with missing survey data. Finally, KFAS (Helske 2017) also offers Poisson and negative binomial state space models, which might also be suitable to model the number of hospitalizations.

7 Conclusion

This research has uncovered important correlations between various indicators and COVID-19 hospitalizations, highlighting the potential role of specific indicators in building predictive models for future pandemics. By thoroughly exploring different indicators and model configurations, this study provides information in the context of a decision-making framework that could aid statisticians, data scientist, practitioners, and policymakers.

Acknowledgments

We like to thank Harm Jan Boonstra for useful comments on an earlier version of this manuscript. The views expressed in this paper are those of the author(s) and do not necessarily reflect the policies of Statistics Netherlands.

References

Banholzer, N. et al. (2022). "The methodologies to assess the effectiveness of non-pharmaceutical interventions during COVID-19: A systematic review." In: *European Journal of Epidemiology* 37, pp. 1003–1024. DOI: 10.1007/s10654-022-00908-y.

Boessenkool, B. (2023). *An R package to handle data from the German Weather Service*. https://bookdown.org/brry/rdwd/.

COSMO (2023). *COVID-19 Snapshot Monitoring*. https://projekte.uni-erfurt.de/cosmo2020/web/. (Visited on 11/12/2023).

Durbin, J. and S. J. Koopman (2012). *Time Series Analysis by State Space Methods*. 2nd. Oxford: Oxford University Press.

Ganslmeier, M., D. Furceri, and J. D. Ostry (2021). "The impact of weather on COVID-19 pandemic". In: *Scientific Reports* 11.22027, pp. 1–7. DOI: 10.1038/s41598-021-01189-3.

Gao, J. et al. (2023). "Evidence-driven spatiotemporal COVID-19 hospitalization prediction with Ising dynamics". In: *Nature Communications* 14.1, p. 3093. DOI: 10.1038/s41467-023-38756-3.

Gatto, A. et al. (2021). "Limits of Compartmental Models and New Opportunities for Machine Learning: A Case Study to Forecast the Second Wave of COVID-19 Hospitalizations in Lombardy, Italy". In: *Informatics* 8.3, p. 57. DOI: 10.3390/informatics8030057.

Gerlee, P. et al. (2021). "Predicting regional COVID-19 hospital admissions in Sweden using mobility data". In: *Scientific Reports* 11.1, p. 24171. DOI: 10.1038/s41598-021-03499-y.

Google (2023). *Mobility reports on the Corona crisis: Differences in movement behavior due to COVID-19*. https://www.google.com/covid19/mobility/. (Visited on 11/12/2023).

Hale, T., J. Anania, et al. (2023). *Variation in government responses to COVID-19*. BSG Working Paper 2020/032 Version 15.0. Oxford: Blavatnik School of Government.

Hale, T., N. Angrist, et al. (2021). "A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker)". In: *Nature Human Behaviour* 5.4, pp. 529–538. DOI: 10.1038/s41562-021-01079-8.

Harvey, A. C. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge: Cambridge University Press.

Helske, J. (2017). “KFAS: Exponential family state space models in R”. In: *Journal of Statistical Software* 78.10, pp. 1–39. DOI: 10.18637/jss.v078.i10.

Hetebrij, W. A. et al. (2024). “Inferring hospital admissions from SARS-CoV-2 virus loads in wastewater in The Netherlands, August 2020 – February 2022”. In: *Science of The Total Environment* 912, p. 168703. DOI: 10.1016/j.scitotenv.2023.168703.

hystreet GmbH (2022). *Pedestrian frequencies: Real time. Precise. Transparent.* https://hystreet.com. (Visited on 11/12/2023).

Iezadi, S. et al. (Nov. 2021). “Effectiveness of non-pharmaceutical public health interventions against COVID-19: A systematic review and meta-analysis”. In: *PLOS ONE* 16.11, pp. 1–19. DOI: 10.1371/journal.pone.0260371.

Klein, B. et al. (2023). “Forecasting hospital-level COVID-19 admissions using real-time mobility data”. In: *Communications Medicine* 3. DOI: 10.1038/s43856-023-00253-5.

Klingwort, J., J. Burger, and J. van den Brakel (2024). “Effectiveness of non-pharmaceutical policy interventions in reducing population mobility during the COVID-19 pandemic”. In: *Journal of the Royal Statistical Society: Series A*, qnae050. DOI: 10.1093/jrsssa/qnae050.

KNMI (2023). *Daily weather data*. https://www.knmi.nl/nederland-nu/klimatologie/daggegevens. (Visited on 11/12/2023).

Koopman, S. J. (1997). “Exact initial Kalman filtering and smoothing for non-stationary time series models.” In: *Journal of the American Statistical Association* 92, pp. 1630–1638. DOI: 10.1080/01621459.1997.10473685.

Lampos, V. et al. (2021). “Tracking COVID-19 using online search”. In: *npj Digital Medicine* 4.17, pp. 1–11. DOI: 10.1038/s41746-021-00384-w.

Li, X. et al. (2023). “Correlation between SARS-CoV-2 RNA concentration in wastewater and COVID-19 cases in community: A systematic review and meta-analysis”. In: *Journal of Hazardous Materials* 441, p. 129848. DOI: 10.1016/j.jhazmat.2022.129848.

Mantilla-Calderon, D. et al. (2022). “Emerging investigator series: meta-analyses on SARS-CoV-2 viral RNA levels in wastewater and their correlations to epidemiological indicators”. In: *Environmental Science: Water Research & Technology* 8.7, pp. 1391–1407. DOI: 10.1039/D2EW00084A.

Mendez-Brito, A., C. El Bcheraoui, and F. Pozo-Martin (2021). “Systematic review of empirical studies comparing the effectiveness of non-pharmaceutical interventions against COVID-19”. In: *Journal of Infection* 83.3, pp. 281–293. DOI: 10.1016/j.jinf.2021.06.018.

Paraskevis, D. et al. (2021). “A review of the impact of weather and climate variables to COVID-19: In the absence of public health measures high temperatures cannot probably mitigate outbreaks”. In: *Science of The Total Environment* 768, p. 144578. DOI: 10.1016/j.scitotenv.2020.144578.

Perone, G. (2022). “Comparison of ARIMA, ETS, NNAR, TBATS and hybrid models to forecast the second wave of COVID-19 hospitalizations in Italy”. In: *European Journal of Health Economics* 23.6, pp. 917–940. DOI: 10.1007/s10198-021-01347-4.

Perra, N. (2021). “Non-pharmaceutical interventions during the COVID-19 pandemic: A review”. In: *Physics Reports* 913, pp. 1–52. DOI: 10.1016/j.physrep.2021.02.001.

Peters, J. A. and M. Farhadloo (2023). “The effects of nonpharmaceutical interventions on COVID-19 cases, hospitalizations, and mortality: A systematic literature review and meta-analysis”. In: *AJPM Focus* 2.4, article 100125. DOI: 10.1016/j.focus.2023.100125.

Reno, C. et al. (2020). “Forecasting COVID-19-Associated Hospitalizations under Different Levels of Social Distancing in Lombardy and Emilia-Romagna, Northern Italy: Results from an Extended SEIR Compartmental Model”. In: *Journal of Clinical Medicine* 9.5, p. 1492. DOI: 10.3390/jcm9051492.

Rijksoverheid (2023a). *Coronavirus Dashboard: View on the hospitals*. https://coronadashboard.government.nl/landelijk/ziekenhuizen-en-zorg. (Visited on 11/12/2023).

Rijksoverheid (2023b). *Coronavirus Dashboard: Virus particles in wastewater*. https://coronadashboard.government.nl/landelijk/rioolwater. (Visited on 11/12/2023).

RIVM (2023). *Background information trend research*. https://www.rivm.nl/gedragsonderzoek/trendonderzoek/achtergrondinformatie. (Visited on 11/12/2023).

Robert Koch-Institut (2023a). *COVID-19 hospitalizations in Germany*. https://github.com/robert-koch-institut/COVID-19-Hospitalisierungen_in_Deutschland. (Visited on 11/12/2023).

– (2023b). *Wastewater surveillance for SARS-CoV-2 [in German: Abwassersurveillance auf SARS-CoV-2]*. https://www.rki.de/DE/Content/Institut/OrgEinheiten/Abt3/FG32/Abwassersurveillance/Abwassersurveillance.html. (Visited on 11/12/2023).

Turk, P. J. et al. (2021). "A predictive internet-based model for COVID-19 hospitalization census". In: *Scientific Reports* 11.5106, pp. 1–12. DOI: 10.1038/s41598-021-84091-2.

Twigg, C. and J. Wenk (2022). "Review and Meta-Analysis: SARS-CoV-2 and Enveloped Virus Detection in Feces and Wastewater". In: *ChemBioEng Reviews* 9.2, pp. 129–145. DOI: 10.1002/cben.202100039.

Wang, T. et al. (2022). "COVID-19 hospitalizations forecasts using internet search data". In: *Scientific Reports* 12.9661, pp. 1–10. DOI: 10.1038/s41598-022-13162-9.

Zoest, V. van et al. (2024). "Predicting COVID-19 hospitalizations: The importance of healthcare hotlines, test positivity rates and vaccination coverage". In: *Spatial and Spatio-temporal Epidemiology* 48, p. 100636. DOI: 10.1016/j.sste.2024.100636.

A Variable descriptions

Table A.1 Target variable, weather, and policy features.

Theme	Name	Description
Target	hospitalizations	COVID-19 hospitalizations in Germany or the Netherlands
Weather	temperature	Temperature in Germany or the Netherlands
Weather	humidity	Humidity in Germany or the Netherlands
Policy	school_closing	Record closings of schools and universities
Policy	workplace_closing	Record closings of workplaces
Policy	cancel_public_events	Record cancelling public events
Policy	gathering_restrictions	Record the cut-off size for limits on gatherings
Policy	close_public_transport	Record closing of public transport
Policy	stay_at_home_requirements	Record orders to 'shelter-in-place' and otherwise confine to the home
Policy	internal_movement_restrictions	Record restrictions on internal movement between cities/regions
Policy	international_travel_controls	Record restrictions on international travel
Policy	public_information_campaigns	Record presence of public info campaigns
Policy	testing_policy	Record government policy on who has access to testing
Policy	contact_tracing	Record government policy on contact tracing after a positive diagnosis
Policy	emergency_investment_in_healthcare	Announced short term spending on healthcare system
Policy	investment_in_vaccines	Announced public spending on COVID-19 vaccine development
Policy	facial_coverings	Record policies on the use of facial coverings outside the home
Policy	vaccination_policy	Record policies for vaccine delivery for different groups
Policy	protection_elderly	Record policies for protecting elderly people (as defined locally) in Long Term Care Facilities and/or the community and home setting
Policy	income_support	Record if the government is providing direct cash payments to people who lose their jobs or cannot work
Policy	debt_contract_relief	Record if the government is freezing financial obligations for households
Policy	fiscal_measures	Announced economic stimulus spending
Policy	international_support	Announced offers of COVID-19 related aid spending to other countries
Policy	stringency_index	overall impression of government activity, based on 9 policy measures
Policy	government_response_index	overall impression of government activity, based on 16 policy measures
Policy	containment_health_index	overall impression of government activity, based on 14 policy measures
Policy	economic_support_index	overall impression of government activity, based on 2 policy measures

Table A.2 Mobile sensors, local sensors, and survey features.

Theme	Name	Description
Mobile sensor	retail_and_recreation	Mobility trends for places like restaurants, cafes, shopping centers, theme parks, museums, libraries, and movie theaters
Mobile sensor	grocery_and_pharmacy	Mobility trends for places like grocery markets, food warehouses, farmers markets, specialty food shops, drug stores, and pharmacies
Mobile sensor	parks	Mobility trends for places like local parks, national parks, public beaches, marinas, dog parks, plazas, and public gardens
Mobile sensor	transit_stations	Mobility trends for places like public transport hubs such as subway, bus, and train stations
Mobile sensor	workplaces	Mobility trends for places of work
Mobile sensor	residential	Mobility trends for places of residence
Local sensor	cities	Pedestrian counts in cities
Survey (Germany)	trust	Trust in institutions
Survey (Germany)	regulations	Attitude towards regulations
Survey (Germany)	vaccination_vac	Vaccination willingness and attitude (vaccinated persons)
Survey (Germany)	vaccination_non_vac	Vaccination willingness and attitude (non-vaccinated persons)
Survey (Germany)	reactance	Reactance and conspiracy thinking
Survey (Germany)	risk_perception	Risk perception
Survey (Germany)	protective_behavior	Applied protective behavior
Survey (Germany)	anxiety	Worries and fears
Survey (Germany)	burden	Perceived personal burden
Survey (The Netherlands)	get_tested_if_symptoms	Compliance with and support for getting tested if COVID-19-related symptoms
Survey (The Netherlands)	wear_mask_in_public_transport	Compliance with and support for wearing face mask in public transport
Survey (The Netherlands)	cough_sneeze_in_elbow	Compliance with and support for coughing and sneezing in elbow
Survey (The Netherlands)	keep_1.5m_distance	Compliance with and support for keeping 1.5 m distance
Survey (The Netherlands)	wash_hands_frequently	Compliance with and support for washing hands frequently
Survey (The Netherlands)	work_from_home	Compliance with and support for working from home if possible
Survey (The Netherlands)	vaccinated_or_willing	Vaccinated or willing to get vaccinated
Survey (The Netherlands)	symptoms	At least one airway symptom
Survey (The Netherlands)	worries	Worries about the coronavirus

B Feature selection

Figure B.1 Feature selection for the German data, by trend model (row panels) and time-dependency of the regression components (column panels).

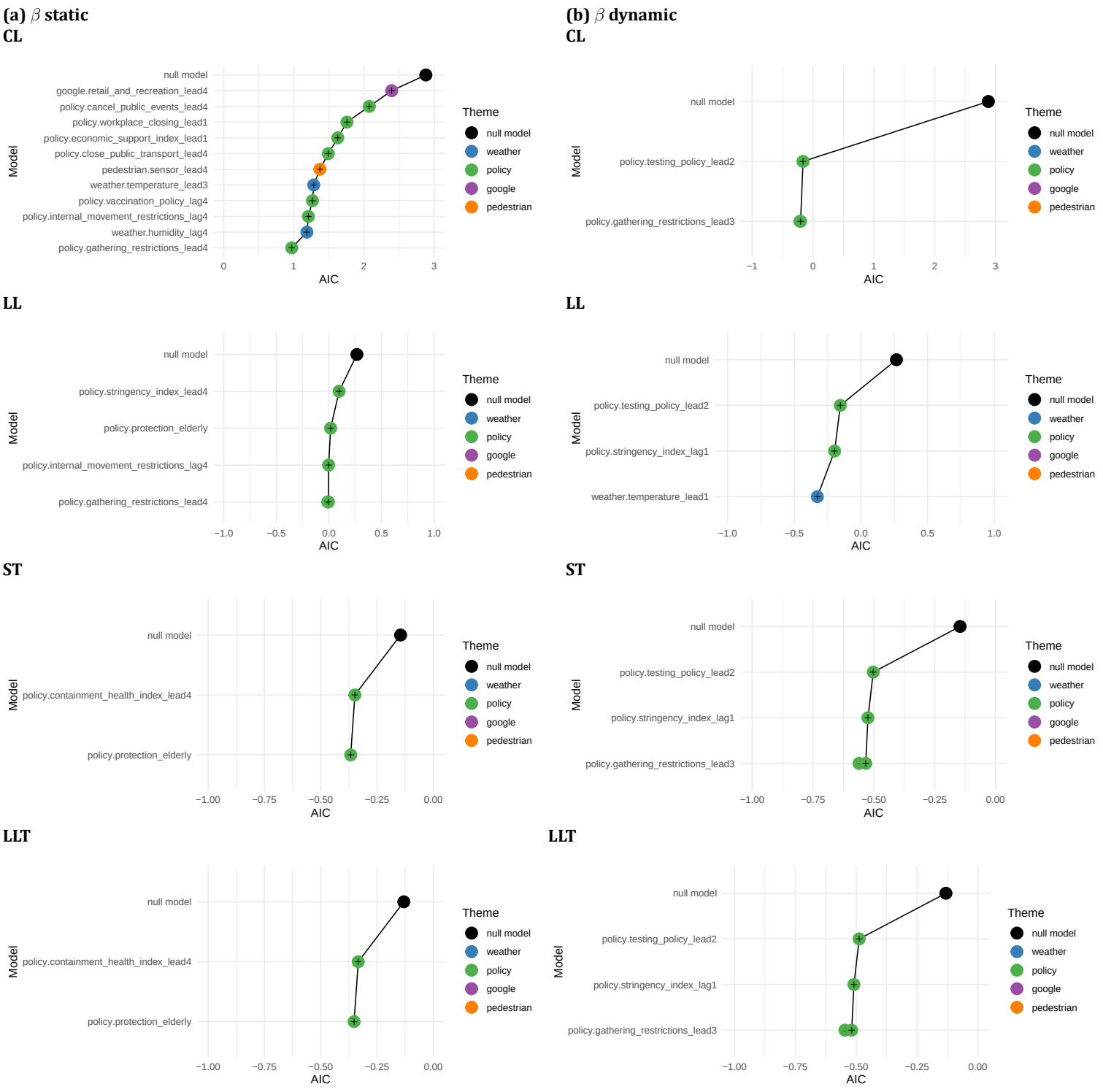
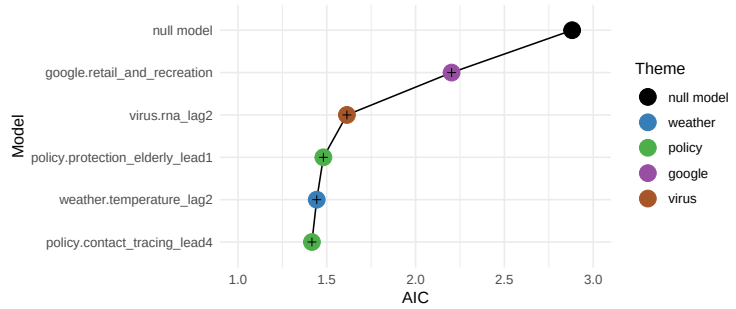


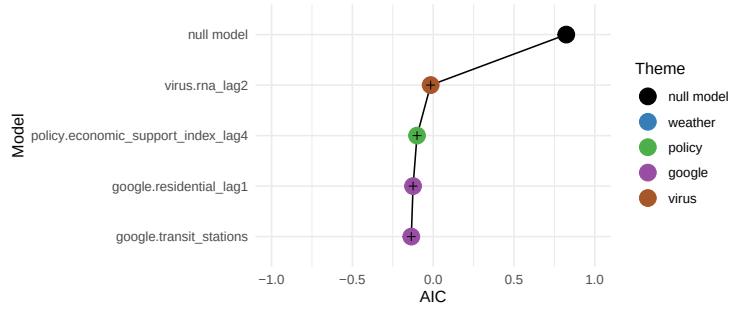
Figure B.2 Feature selection for the Dutch data, by trend model (row panels) and time-dependency of the regression components (column panels).

(a) β static

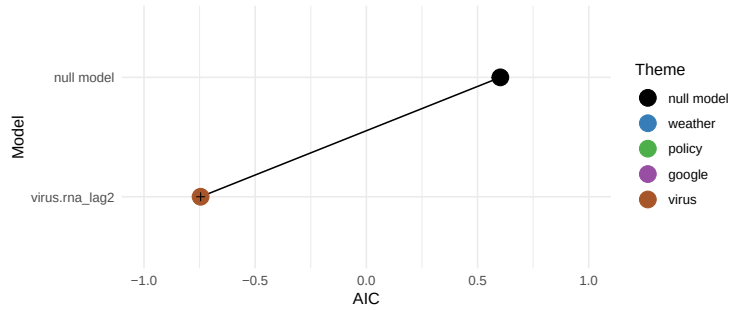
CL



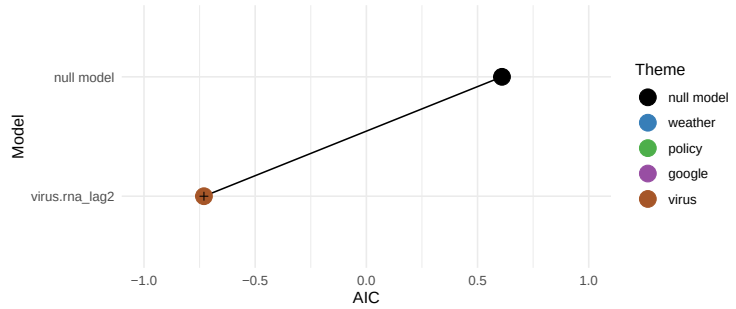
LL



ST

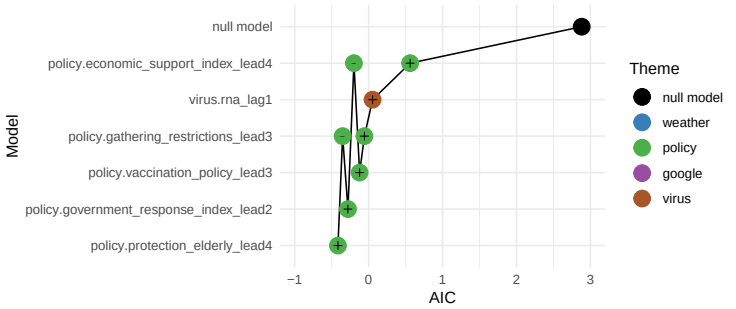


LLT

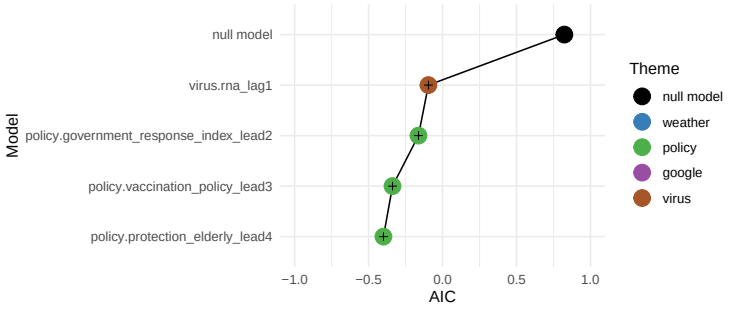


(b) β dynamic

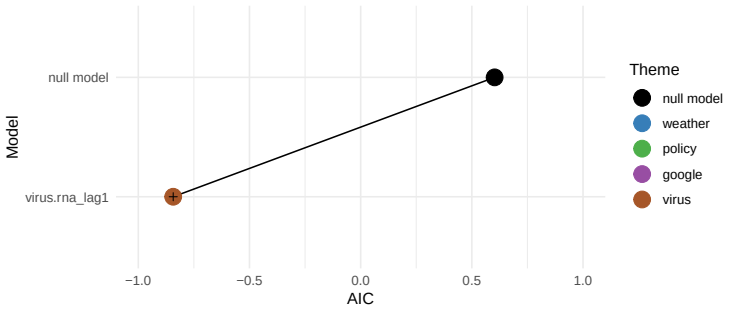
CL



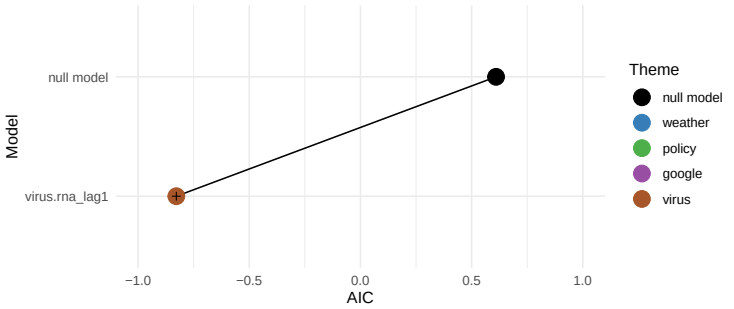
LL



ST



LLT



C Model diagnostics

Figure C.1 Bowman-Shenton test for normality of innovations, by trend model (x-axis), time-dependence of the regression components (legend) and country (column panels).

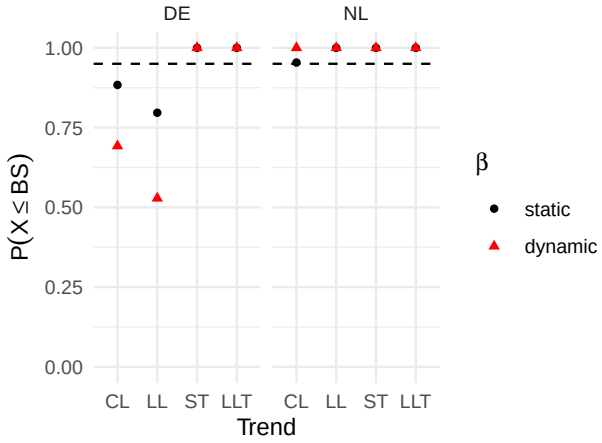


Figure C.2 F-test for heteroscedasticity of innovations, by trend model (x-axis), time-dependence of the regression components (legend) and country (column panels).

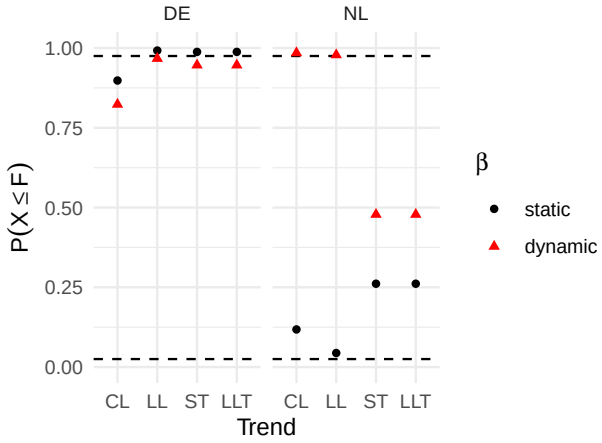
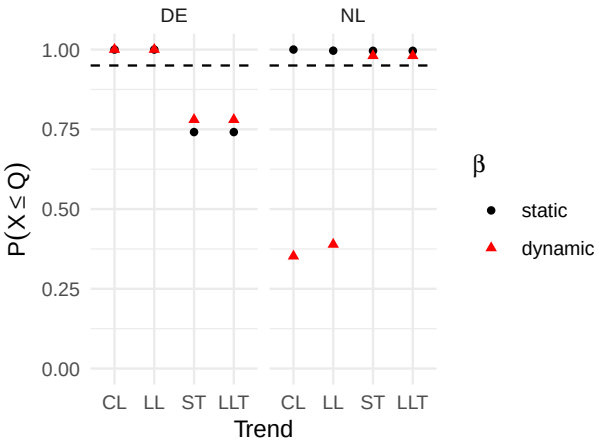


Figure C.3 Ljung-Box test for autocorrelation of innovations, by trend model (x-axis), time-dependence of the regression components (legend) and country (column panels).



D Model diagnostics---Germany

Figure D.1 Standardized innovations, by trend model (row panels) and time-dependence of the regression components (column panels). German data.

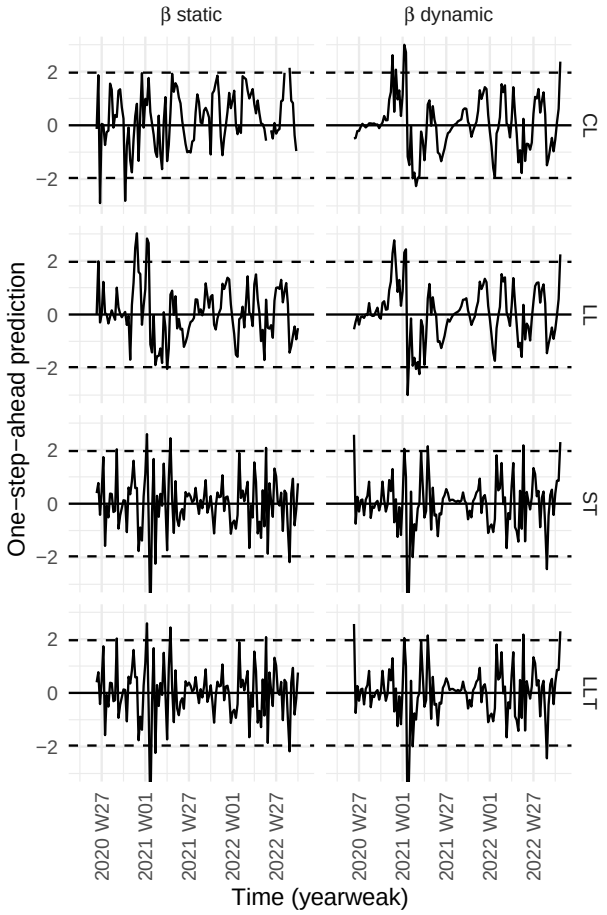


Figure D.2 Q-Q plot of standardized innovations, by trend model (row panels) and time-dependence of the regression components (column panels). German data.

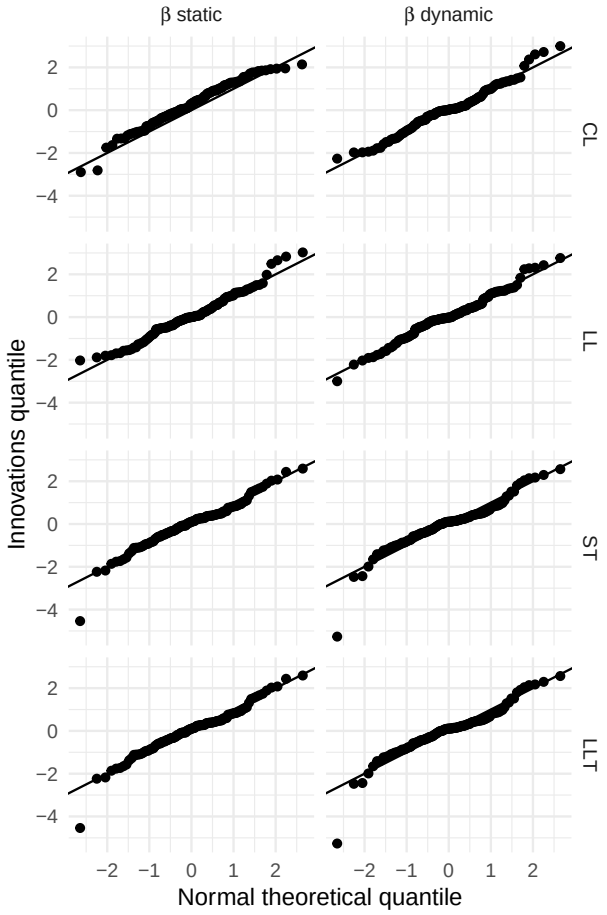


Figure D.3 Histogram of standardized innovations, by trend model (row panels) and time-dependence of the regression components (column panels). German data.

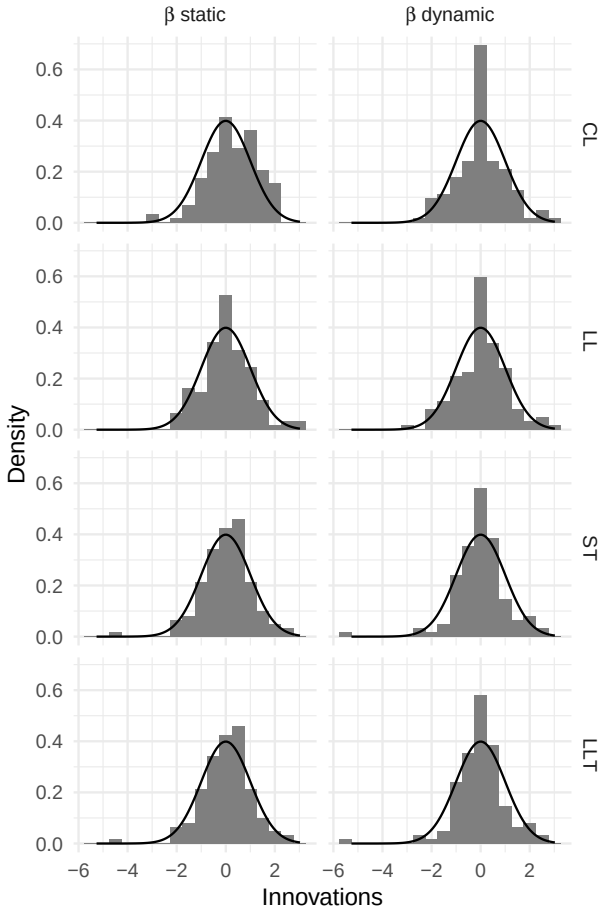
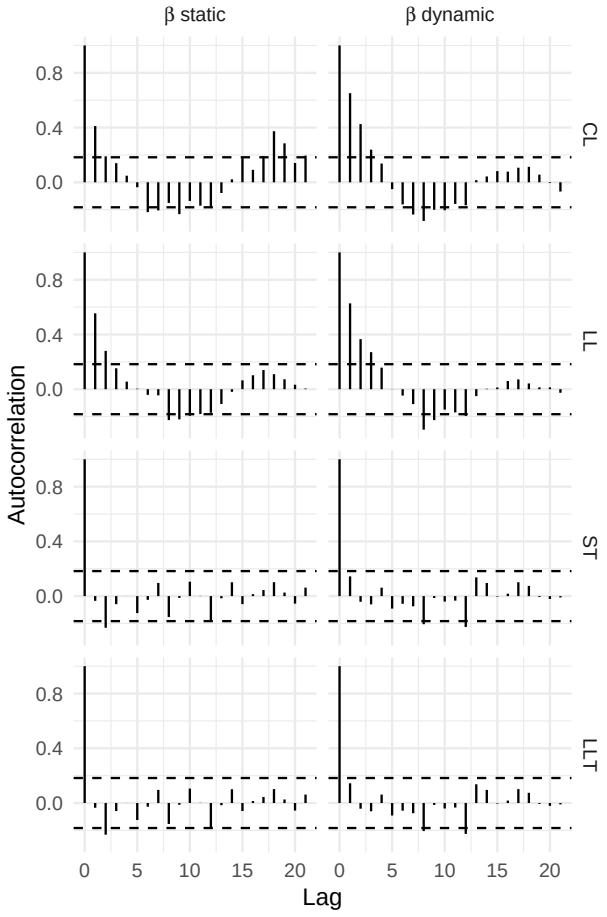


Figure D.4 Autocorrelation of standardized innovations, by trend model (row panels) and time-dependence of the regression components (column panels). German data.



E Model diagnostics---the Netherlands

Figure E.1 Standardized innovations, by trend model (row panels) and time-dependence of the regression components (column panels). Dutch data.

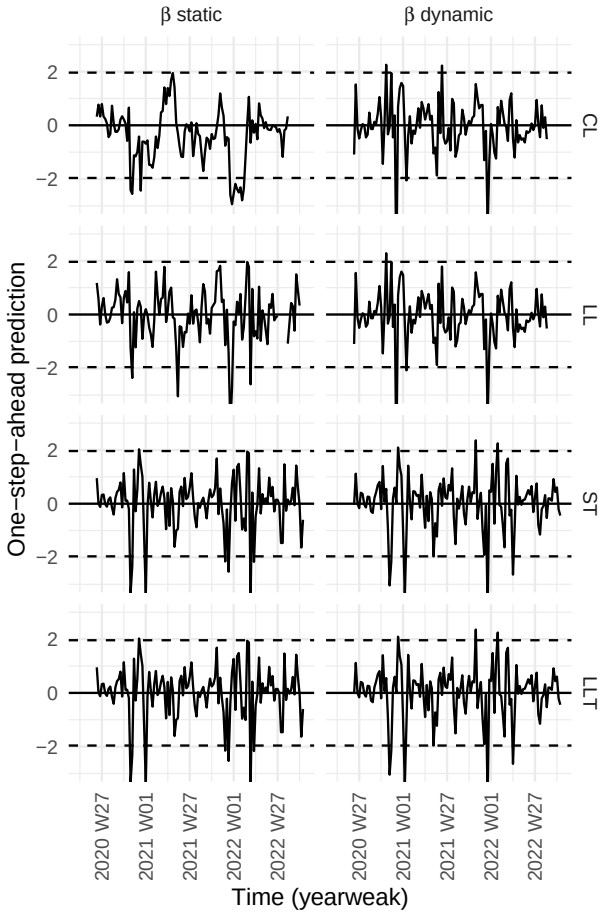


Figure E.2 Q-Q plot of standardized innovations, by trend model (row panels) and time-dependence of the regression components (column panels). Dutch data.

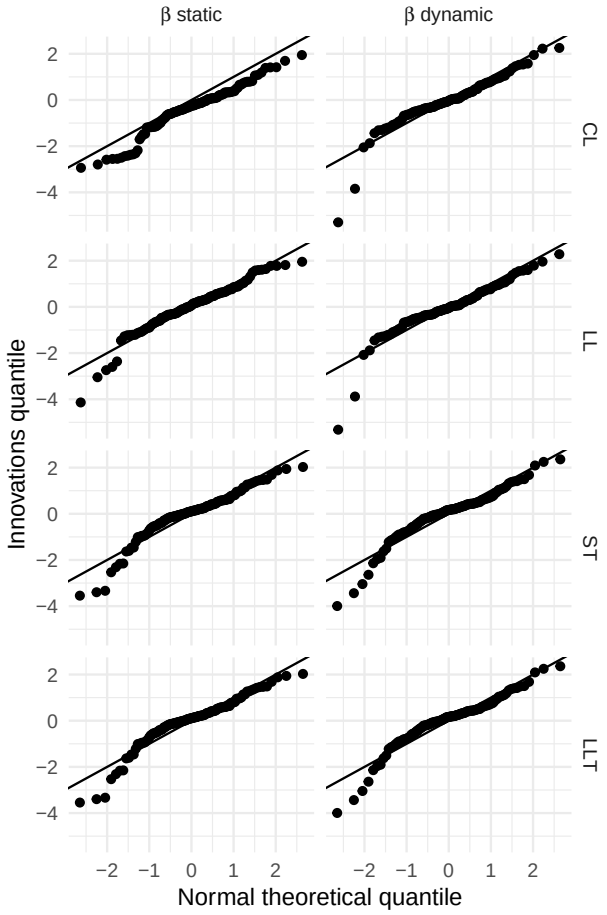


Figure E.3 Histogram of standardized innovations, by trend model (row panels) and time-dependence of the regression components (column panels). Dutch data.

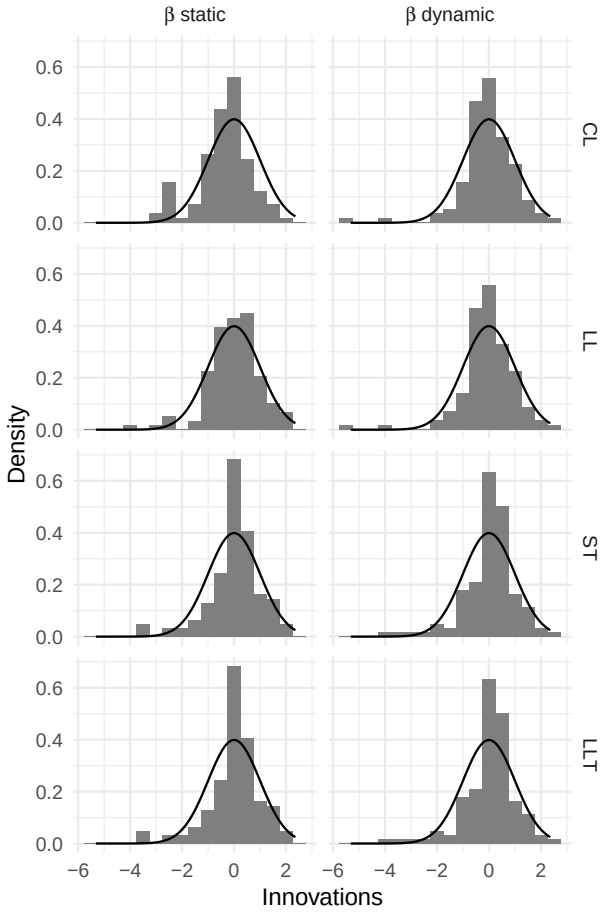
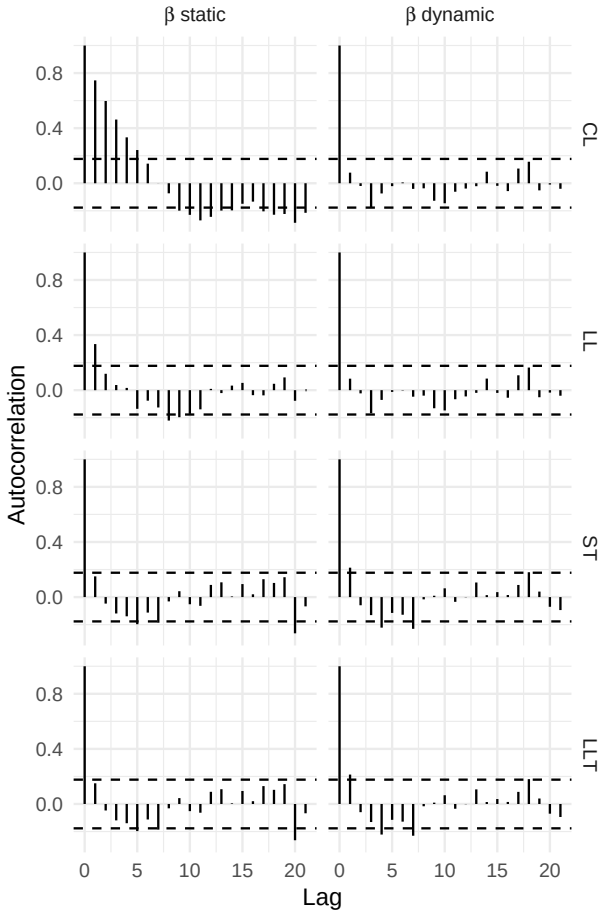


Figure E.4 Autocorrelation of standardized innovations, by trend model (row panels) and time-dependence of the regression components (column panels). Dutch data.



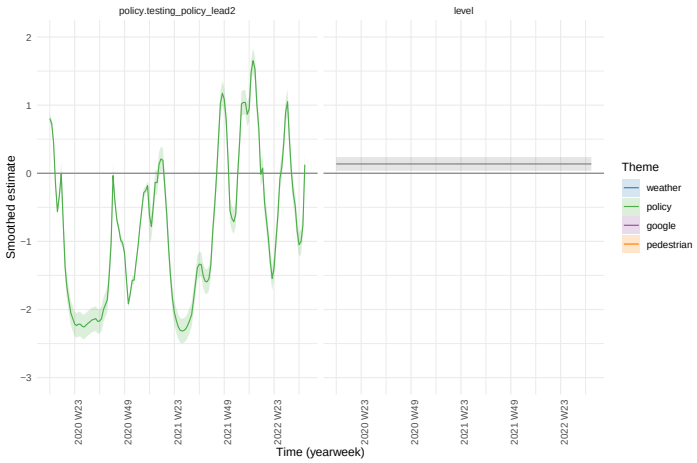
F Smoothed estimates of state variables

Figure F.1 Smooth estimates of state variables with 95% confidence intervals, by trend model (row panels) and time-dependency of the regression components (column panels). German data.

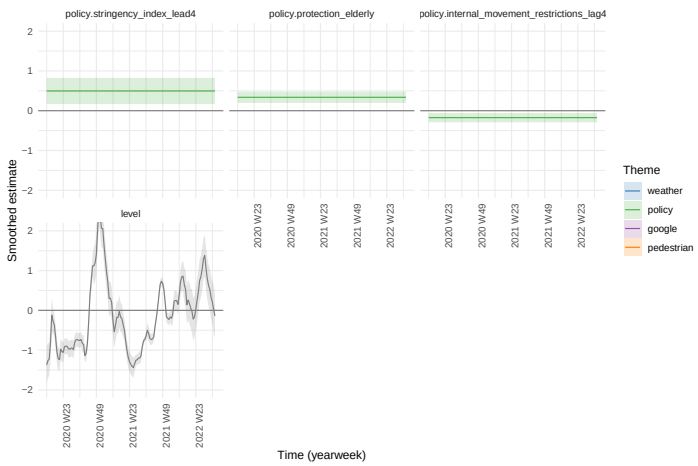
(a) β static

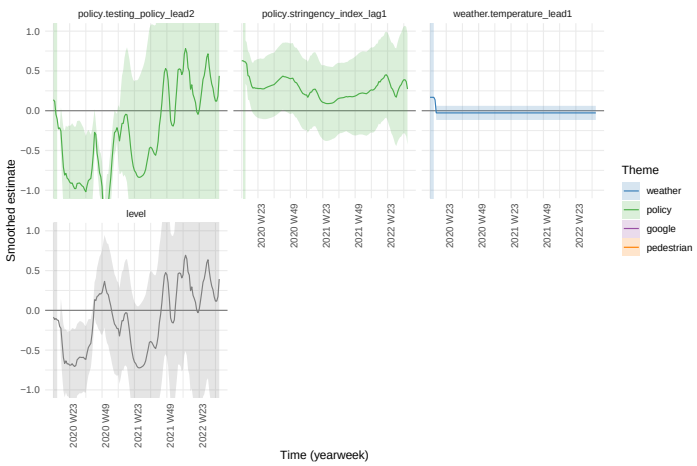
(b) β dynamic

CL

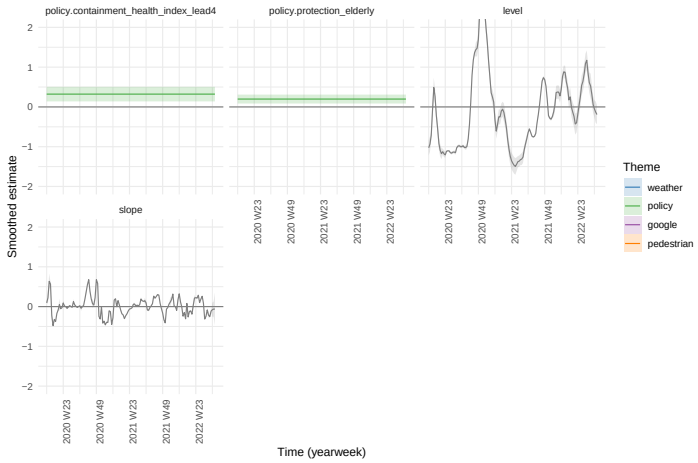


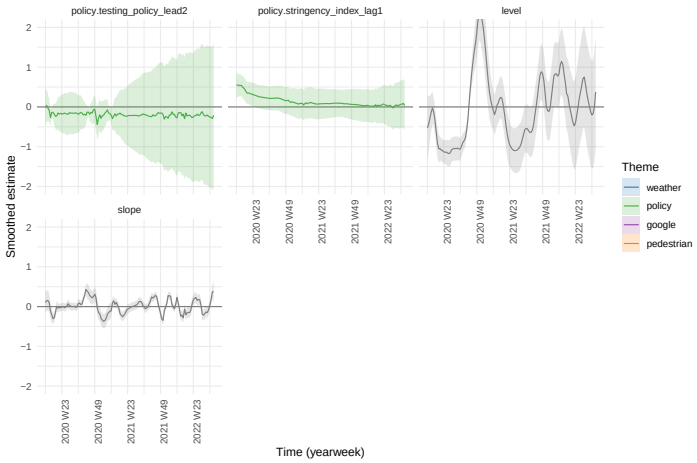
LL



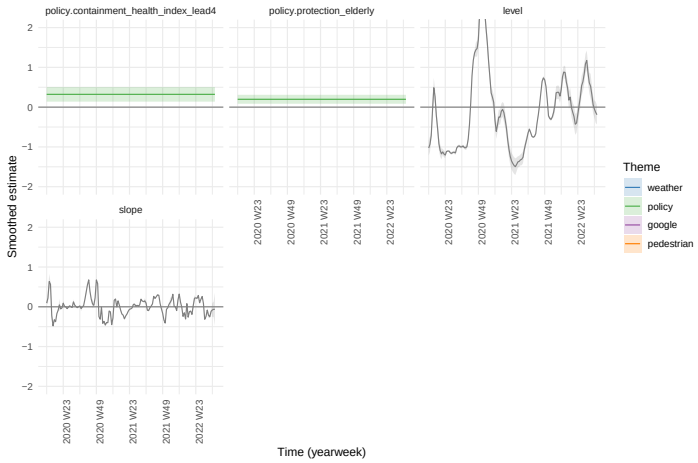


ST





LLT



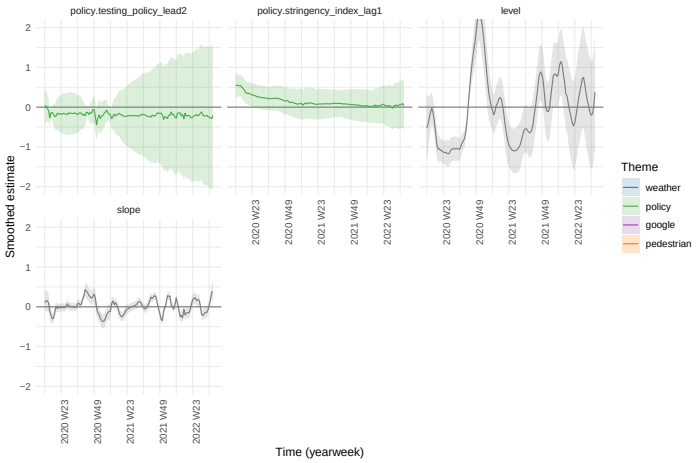
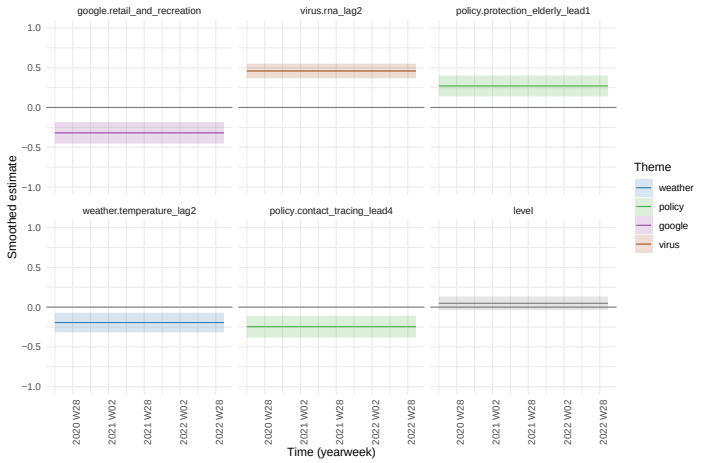


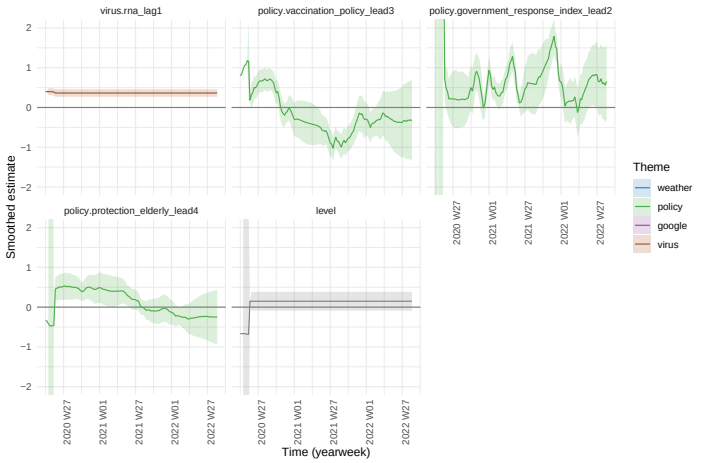
Figure F.2 Smooth estimates of state variables with 95% confidence intervals, by trend model (row panels) and time-dependency of the regression components (column panels). Dutch data.

(a) β static

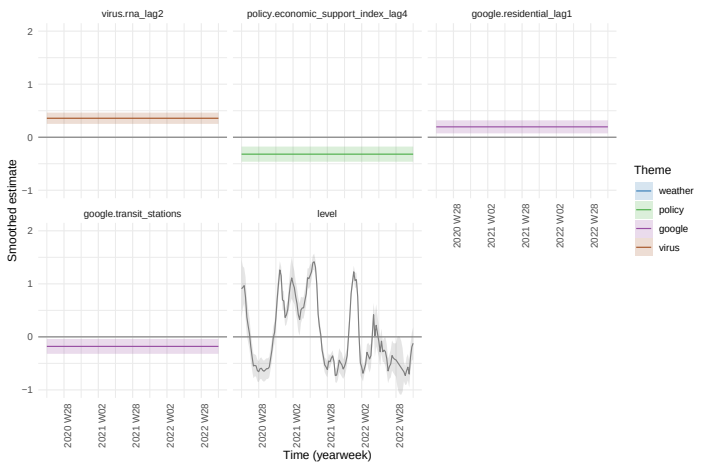
(b) β dynamic

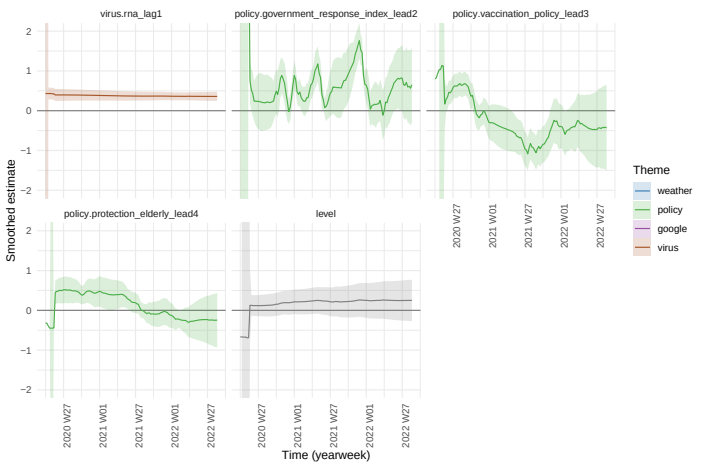
CL



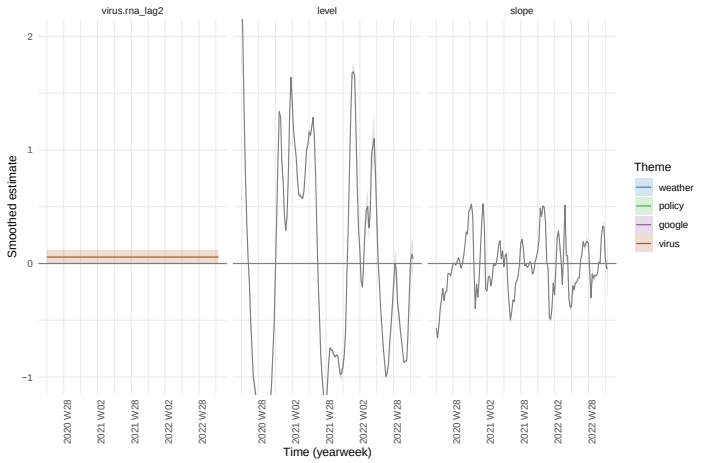


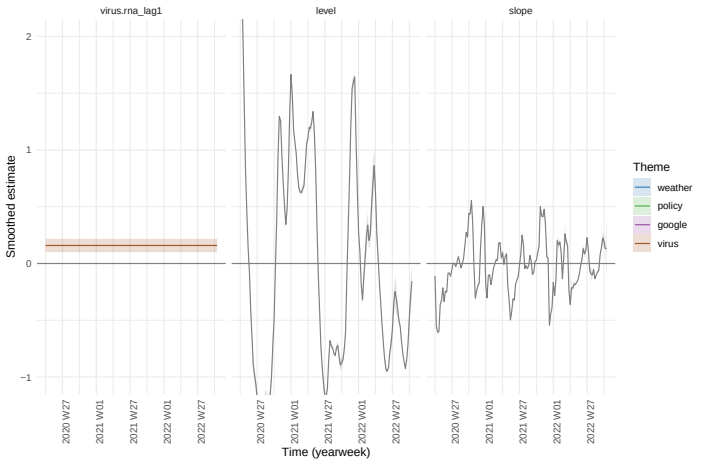
LL



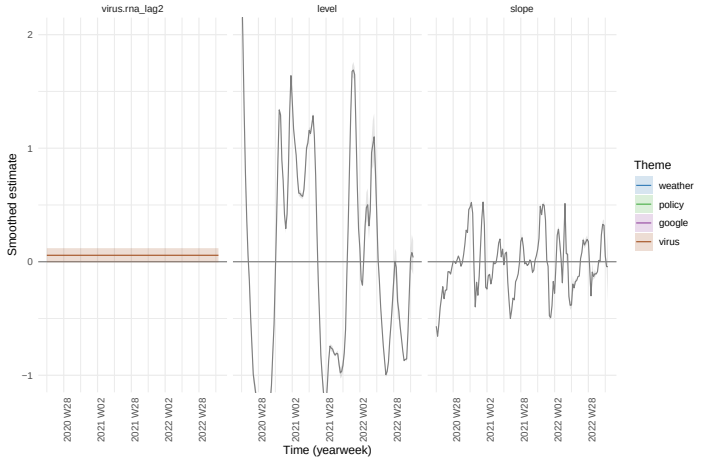


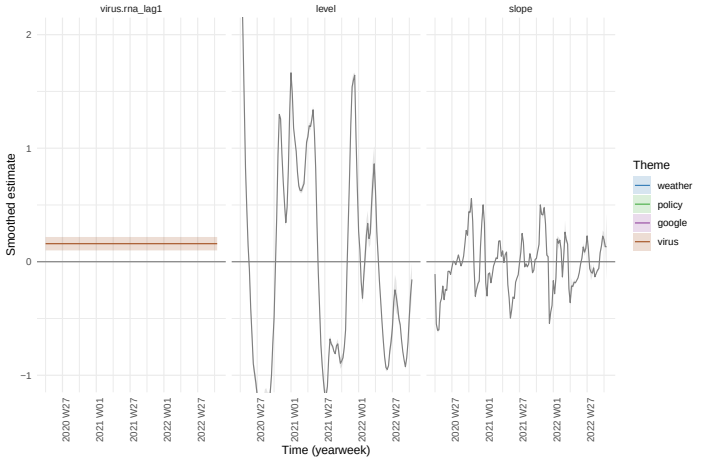
ST





LLT





Colophon

Publisher
Statistics Netherlands
Henri Faasdreef 312, 2492 JP The Hague
www.cbs.nl

Prepress
Statistics Netherlands, Grafimedia

Design
Edenspiekermann

Information
Telephone +31 88 570 70 70, fax +31 70 337 59 94
Via contact form: www.cbs.nl/information