TI 2024-037/III
Tinbergen Institute Discussion Paper

Statistical Early Warning Models with Applications

Lucas P. Harlaar[1]
Jacques J.F. Commandeur[2]
Jan A. van den Brakel[3]
Siem Jan Koopman[4]
Niels Bos[5]
Frits D. Bijleveld[6]

1 Statistics Netherlands, Maastricht University and Vrije Universiteit Amsterdam

2 SWOV Institute for Road Safety Research and Vrije Universiteit Amsterdam

3 Statistics Netherlands and Maastricht University

4 Vrije Universiteit Amsterdam and Tinbergen Institute

5 SWOV Institute for Road Safety Research

6 SWOV Institute for Road Safety Research and Vrije Universiteit Amsterdam

Statistical Early Warning Models with Applications

Lucas P. Harlaar*,
Statistics Netherlands, Maastricht University, Vrije Universiteit Amsterdam
Jacques J.F. Commandeur,
SWOV Institute for Road Safety Research, Vrije Universiteit Amsterdam
Jan A. van den Brakel,
Statistics Netherlands, Maastricht University
Siem Jan Koopman,
Vrije Universiteit Amsterdam, Tinbergen Institute
Niels Bos,
SWOV Institute for Road Safety Research
Frits D. Bijleveld,
SWOV Institute for Road Safety Research, Vrije Universiteit Amsterdam

May 24, 2024

Abstract

This paper investigates the feasibility of using earlier provisional data to improve the now- and forecasting accuracy of final and official statistics. We propose the use of a multivariate structural time series model which includes common trends and seasonal components to combine official statistics series with related auxiliary series. In this way, more precise and more timely nowcasts and forecasts of the official statistics can be obtained by exploiting the higher frequency and/or the more timely availability of the auxiliary series. The proposed method can be applied to different data sources consisting of any number of missing observations both at the beginning and at the end of the series simultaneously. Two empirical applications are presented. The first one focuses on fatal traffic accidents and the second one on labour force participation at the municipal level. The results demonstrate the effectiveness of our proposed approach in improving forecasting performance for the target series and providing early warnings to policy-makers.

*Corresponding author: Department of Quantitative Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. E-mail: lucas.harlaar@maastrichtuniversity.nl

1 Introduction

When realizations of a dynamic process suffer from long delays, it can potentially cause a variety of challenges in the practice of their statistical analyses, especially in situations where timely and accurate information is critical for decision-making and policy-making. Although high-frequency time series become increasingly available, many traditional low-frequency time series, as published by government and statistical offices, still suffer from publication lags of some length. A growing literature has emerged on methods that can incorporate provisional or more timely auxiliary data to make more accurate predictions for the current and future state of the unknown dynamic process; see, for example, Grybauskas, Pilinkienė, Lukauskas, Stundžienė, and Bruneckienė (2023) for an extensive literature overview. Many of these contributions rely on the use of multivariate frameworks that combine past data with present (preliminary) information, by jointly modelling the time series of interest with related auxiliary series; see, for example, Harvey (2006), Doz, Giannone, and Reichlin (2011), and Blasques, Koopman, Mallee, and Zhang (2016). The statistical analyses used in these contributions rely on state space methods; see Harvey (1989) and Durbin and Koopman (2012). Other approaches include the Bayesian analysis in Carriero, Clark, and Marcellino (2015) and the mixed-frequency analysis in Kuzin, Marcellino, and Schumacher (2011) where realizations can be sampled at different periodic frequencies (for example, monthly, quarterly and yearly).

In this paper we develop a general state space framework for the specific task of creating early warning indicators for variables of key interest. We illustrate the importance of this development by two applications of societal relevance. First, in road traffic safety studies, the reported number of traffic fatalities is of interest to the general public but also to policy makers in national governments. There are often multiple sources that issue data on the number of road traffic fatalities. For example, in many countries, the national statistical office provides the official monthly and yearly numbers of road fatalities. But government agencies for roads and traffic usually also provide information on the number of road fatalities. The construction of these statistics typically rely on different resources (for example, police and/or hospital records) and on different counting methods (which can also be based on different definitions). In this first empirical application, we consider the monthly number of road traffic fatalities in the Netherlands which are provided by Statistics Netherlands (SN) and the government agency "Rijkswaterstaat (RWS)". The publication frequency of their official statistics is yearly, while SN provides early provisional data on a monthly basis (flash-estimate). The provisional data is subject to revision throughout the year. The adopted definitions of the data from SN and RWS are such that the data can be regarded as realizations from the same underlying dynamic process for road safety.

Second, another variable of key importance, especially for economic policy makers and the general public, is the labour force participation (LFP) statistic. A recent literature overview of nowcasting unemployment using data from a Labour Force Survey (LFS) and other external resources is provided in Grybauskas, Pilinkienė, Lukauskas, Stundžienė, and Bruneckienė (2023). In case of the Netherlands, SN publishes labour force participation figures which are derived from the Tax Administration and Employee Insurance Agency registers at a quarterly frequency, at the municipal level (or even at the neighborhood level). This information, however, comes with a delay of more than one year. SN also conducts the Dutch LFS on a monthly basis and these estimates become available within the first two

weeks of the subsequent month. Due to sampling error and the higher frequency, this could be considered as a preliminary noisy indicator (flash-estimate) for the final figure retrieved directly from the Tax Administration and Employee Insurance Agency. In van den Brakel and Michiels (2021) these quarterly tax register figures are combined with *quarterly* LFS figures in a bivariate state space model to nowcast low regional figures on labour force participation derived from the tax register. In our second empirical application, we investigate the feasibility of using these earlier provisional *monthly* LFS data to improve the forecasting accuracy of the final LFP figure provided by SN.

For the two empirical studies, our general state space framework is based on a multivariate unobserved components time series model with trend and seasonal components as proposed and described in detail by Harvey (1989). The trend and seasonal components can capture common structures within the multiple set of realized time series observations. The overall methodology is general and can be applied to many different data sources which may be subject to missing observations, both at the beginning and at the end of the series. In the first empirical application it is shown that our proposed method of using earlier provisional data leads to more precise forecasts of the final and official monthly number of road fatalities in the Netherlands. In the second empirical application it is shown how the method achieves more accurate nowcasts of quarterly LFP rates of the municipality of Amsterdam. The two applications indicate that our proposed method can be applied in settings where the availability of auxiliary data is important for providing timely information to policy makers. The multivariate framework can facilitate an early warning system, allowing to detect whether the upcoming final figures are expected to substantially deviate from the past. Another important feature of the method is its robustness to unexpected turning points which are encountered, for example, during the Covid crisis. Both applications are based on time series data that include the Covid crisis period.

The remainder of this paper is organized as follows. Section 2 introduces the multivariate unobserved components time series model in general terms and it describes the estimation method and diagnostic checks for assessing the validity of the model. In Section 3 and Section 4 we present and discuss the results of the empirical studies for Road Safety and Labour Force processes, respectively. In both sections we start with outlining the available data, then we adjust the general model for the application at hand and we finalize both studies with an extensive in-sample assessment and forecasting study in which different model adjustments and forecasting scenarios are compared over an expanding window between 2011-2022. Section 5 concludes.

2 Methodology

2.1 Statistical Early Warning Model

Our statistical framework is based on the linear Gaussian state space model for the analysis of time series data, see Harvey (1989), Durbin and Koopman (2012), and Commandeur and Koopman (2007); see also Appendix A. An important motivation for opting for this methodology is that it easily handles missing data in time series, and transparently generalizes to the joint analysis of multivariate time series. Define the $N \times 1$ observation vector $\boldsymbol{y}_t = (y_{1,t}, \ldots, y_{N,t})'$ which represents the observations at time t of N time series variables

$y_{i,t}$, for $i = 1, \ldots, N$. The statistical early warning model belongs to the general class of *seemingly unrelated time series equations* (SUTSE) models, which can be considered as a multivariate extension of the univariate structural time series model of Harvey (1989). We can represent it as a multivariate unobserved components time series model given by

$$\boldsymbol{y}_t = \boldsymbol{\mu}_t + \boldsymbol{\gamma}_t + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \text{NID}(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon), \tag{1}$$

for $t = 1, \ldots, T$, with T as the length of the time series, where $\boldsymbol{\mu}_t$ is a $N \times 1$ vector of unobserved long-term stochastic trends, $\boldsymbol{\gamma}_t$ is a $N \times 1$ vector of unobserved seasonal stochastic effects and $\boldsymbol{\varepsilon}_t$ is a $N \times 1$ vector of the irregular, disturbance or noise component. The trend can be defined as the *local linear trend* model, where the trend consists of a stochastic level $\boldsymbol{\mu}_t$ and slope $\boldsymbol{\nu}_t$ components, and is given by

$$\boldsymbol{\mu}_{t+1} = \boldsymbol{\mu}_t + \boldsymbol{\nu}_t + \boldsymbol{\xi}_t, \qquad \boldsymbol{\xi}_t \sim \text{NID}(\mathbf{0}, \boldsymbol{\Sigma}_\xi), \tag{2}$$

$$\boldsymbol{\nu}_{t+1} = \boldsymbol{\nu}_t + \boldsymbol{\zeta}_t, \qquad \boldsymbol{\zeta}_t \sim \text{NID}(\mathbf{0}, \boldsymbol{\Sigma}_\zeta), \tag{3}$$

where the disturbance vectors $\boldsymbol{\xi}_t$ and $\boldsymbol{\zeta}_t$ are mutually independent (also at all lags) and both independent of $\boldsymbol{\varepsilon}_t$. Specific trend specifications can be obtained by introducing restrictions on the $N \times N$ covariance matrices $\boldsymbol{\Sigma}_\xi$ or $\boldsymbol{\Sigma}_\zeta$ by enforcing sparseness (zero coefficients), by having rank conditions or even by setting them equal to $\mathbf{0}$. Such specifications also lead to a more parsimonious model. The seasonal process can be modeled as a *dummy seasonal* model,

$$\boldsymbol{\gamma}_{t+1} = -\sum_{j=1}^{s-1} \boldsymbol{\gamma}_{t+1-j} + \boldsymbol{\omega}_t, \qquad \boldsymbol{\omega}_t \sim \text{NID}(\mathbf{0}, \boldsymbol{\Sigma}_\omega), \tag{4}$$

where s equals the number of months or quarters per year, depending on the frequency of the time series, and the disturbance vector $\boldsymbol{\omega}_t$ is independent of all other disturbance vectors. Alternatively, a *trigonometric seasonal* model can be used, see Appendix A. The model is defined as a SUTSE as it has for each equation its own trend, seasonal and irregular components, but at the same time it allows for cross-sectional correlations in the disturbances within these components. Their subsequent disturbance covariance matrices $\boldsymbol{\Sigma}_\zeta$, $\boldsymbol{\Sigma}_\omega$ and $\boldsymbol{\Sigma}_\xi$ can be diagonal, full rank or rank deficient (matrix with rank less than its dimension). A rank deficient covariance matrix implies the presence of common components or factors.

The benefit of common factor models is that the estimation procedure is more efficient because they are more parsimonious. Especially when multiple time series refer to the same dynamic process it seems reasonable to consider such model simplifications. When the timeliness of these time series differ, common factor models naturally allow for forecasting the less timely target variable(s), exploiting the timely information of the auxiliary variable(s). This is the main goal of our paper. Various common factor models are compared with univariate benchmarks in two empirical applications to show the flexibility of the methodology as it can be utilized in broad range of applications. As benchmarks a *basic structural time series* model with deterministic and stochastic slope and seasonal are used, introduced by Harvey (1989). Both models are univariate variations on the SUTSE models described in (1)-(4), where the first variant does not include a disturbance term for the slope and seasonal component. The preliminary estimate for the target series gives practitioners an early indication on the current state of the target series, before the latest observations have come in, hence we call the common factor models under consideration *statistical early warning* models.

2.2 State and parameter estimation

The class of SUTSE models outlined in the previous section contains unobserved components and covariance matrices for the stochastic disturbances in these components, which are both generally unknown. Estimates for the state variables, which include the unobserved components, can be retrieved with signal extraction algorithms such as the Kalman filter. The covariance matrices of the measurement disturbance terms and the state disturbance terms are often referred to as hyperparameters and can be estimated via the method of maximum likelihood.

In order to apply the Kalman filter, the SUTSE model is cast into the state space form which is relatively straightforward, see Harvey (1989, Chapter 8). Let the collection of observed information at time t be denoted by the $N \times t$ data matrix $\boldsymbol{Y}_t = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_t)'$. The unobserved components, such as trend $\boldsymbol{\mu}_t$, seasonal $\boldsymbol{\gamma}_t$ and associated components, are gathered in the state vector $\boldsymbol{\alpha}_t$, for which optimal estimates are retrieved by the Kalman filter. This recursive algorithm obtains the first two moments given all observed information, i.e. $\mathrm{E}(\boldsymbol{\alpha}_t|\boldsymbol{Y}_t)$ and $\mathrm{Var}(\boldsymbol{\alpha}_t|\boldsymbol{Y}_t)$. In general, the filtered estimate of the states at time t can be modified using the information that comes available after time t. This can be done by a backward recursive algorithm which we usually refer to as the Kalman smoother. Note that at time T, both methods obtain the same result, see Durbin and Koopman (2012, Chapter 4) for a discussions and derivations of these techniques. A convenient property of Kalman filtering and smoothing is that the recursions can be adjusted straightforwardly when missing observations in the time series need to be handled, see (Durbin and Koopman 2012, Section 4.10). At time $t = 1$, the Kalman filter needs to be initialised. For nonstationary processes and regression effects in the state vector, a diffuse initialization can be used, see Durbin and Koopman (2012, Chapter 5).

Both signal extraction techniques assume however that the system matrices and hyperparameters of the state space model are known. This is generally not the case, therefore the unknown values are replaced by their maximum likelihood estimates obtained by a numerical optimization procedure. The standard errors of the Kalman filter and smoother estimates do not account for the uncertainty of replacing the true hyperparameter values for their maximum likelihood estimates. Several parametric and nonparametric bootstrap methods exist for estimating this extra uncertainty, see Pfeffermann and Tiller (2005). The state space models discussed in this paper are implemented and fitted in OxMetrics (Doornik 2021), with the extensive use of SsfPack (Koopman, Shephard, and Doornik 2008).

2.3 Diagnostic checking

The model assumptions underlying the linear Gaussian models from Section 2.1 boil down to the assumption that the one-step-ahead-prediction errors are normally and independently distributed. Basic diagnostic tests can be performed on the standardised one-step ahead prediction errors to detect significant departures from the model assumptions. These prediction errors are often referred to as *residuals* or *innovations*, since they reflect the part of $\boldsymbol{y}_t$ that can not be predicted from the available information up until t. Note that the standardised innovations are transformed in order to apply these basic diagnostic tests on the individual elements separately, see Durbin and Koopman (2012, Chapter 7). Hence all N mutually independent standardised innovations are normally and serially independently

distributed with constant unit variance in a correctly specified state space model. Model misspecification could lead to biased estimates and incorrect inference, hence we propose several diagnostic checks here to detect this.

A standard test for serial correlation in the first P lags of the residuals is proposed by Ljung and Box (1978). This test statistic is asymptotically χ^2 distributed with $P-H$ degrees of freedom for structural models, under the null hypothesis that the residuals are serially independent and where H refers to the number of hyperparameters. To test whether the variance of the innovations is constant over time, we propose a basic heteroscedasticity test by dividing the sums of two exclusive squared residual subsets of size h. This test statistic should be $F_{h,h}$ distributed, under the null hypothesis of homoscedasticity. To challenge the normality assumption of these innovations, we obtain the skewness and kurtosis, which for a normal distribution should be centered around zero and three respectively. They can be combined into a test statistic that is asymptotically χ^2 with 2 degrees of freedom, under the null hypothesis that the innovations are normally distributed.

In statistical quality control, the Cusum (or cumulative sum control chart) is a sequential analysis technique originally developed by Page (1954). It is typically used for monitoring *change detection*. As its name implies, Cusum involves the calculation of a cumulative sum of the standardised residuals to detect possible breaks in the dynamic process of the time series. For a correctly specified model this cumulative sum is expected to sum to zero and when several consecutive prediction errors have the same sign (i.e., are all positive, or are all negative) than this implies that a structural break is occurring in the development of the series. Visually, a structural break can be spotted when this cumulative sum starts drifting away from the horizontal time axis. When the sum becomes more and more negative then the observed counts are getting smaller and smaller than what we would expect. On the other hand, when the sum becomes more and more positive then the observed counts are getting larger and larger than what should be expected, and this is reason to raise the alarm.

2.4 Goodness-of-fit: in-sample and out-of-sample

The in-sample fit of various competing models are evaluated based on the Akaike information criterion (AIC). This allows for a fair comparison of the log-likelihood values of competing models with different numbers of parameters. In general, models with more parameters will have a larger log-likelihood by default. The AIC corrects for this by penalizing the log-likelihood based on the number of parameters.

The prediction accuracy of the competing models are determined on the basis of three performance measures: the mean absolute percentage error (MAPE), the mean squared error (MSE) and the mean absolute error (MAE) which are defined as

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_{t|\Omega_t}}{y_t}\right| \times 100, \quad \text{MSE} = \frac{1}{n}\sum_{t=1}^{n}\left(y_t - \hat{y}_{t|\Omega_t}\right)^2, \quad \text{MAE} = \frac{1}{n}\sum_{t=1}^{n}\left|y_t - \hat{y}_{t|\Omega_t}\right|,$$

where $\hat{y}_{t|\Omega_t}$ is the nowcast or forecast (in the latter case, Ω_t is replaced by Ω_{t-1}) of the univariate target series y_t, given all available information in data set Ω_t at time t and n is the number of nowcasts in the forecast window. In the case of MAPE, the prediction error is expressed as a percentage relative to the actual observation.

3 Empirical Study I: Traffic Safety Monitor

3.1 Data

Both Statistics Netherlands (SN) and the government agency "Rijkswaterstaat (RWS)" are collecting statistics on road fatalities on a monthly frequency, which are being published on a monthly or yearly frequency. The data consist of four time series on the number of road fatalities, see Figure 1. These series are:

1. final number of road fatalities (FINRF) in the Netherlands, provided by SN;
2. registered number of road fatalities (REGRF) in the Netherlands, issued by RWS;
3. preliminary number of road fatalities amongst Dutch citizens (PRERFDC), from SN;
4. final number of road fatalities amongst Dutch citizens (FINRFDC), also from SN.

Once issued, the observations in the final series remain fixed and are not revised or adapted over time.

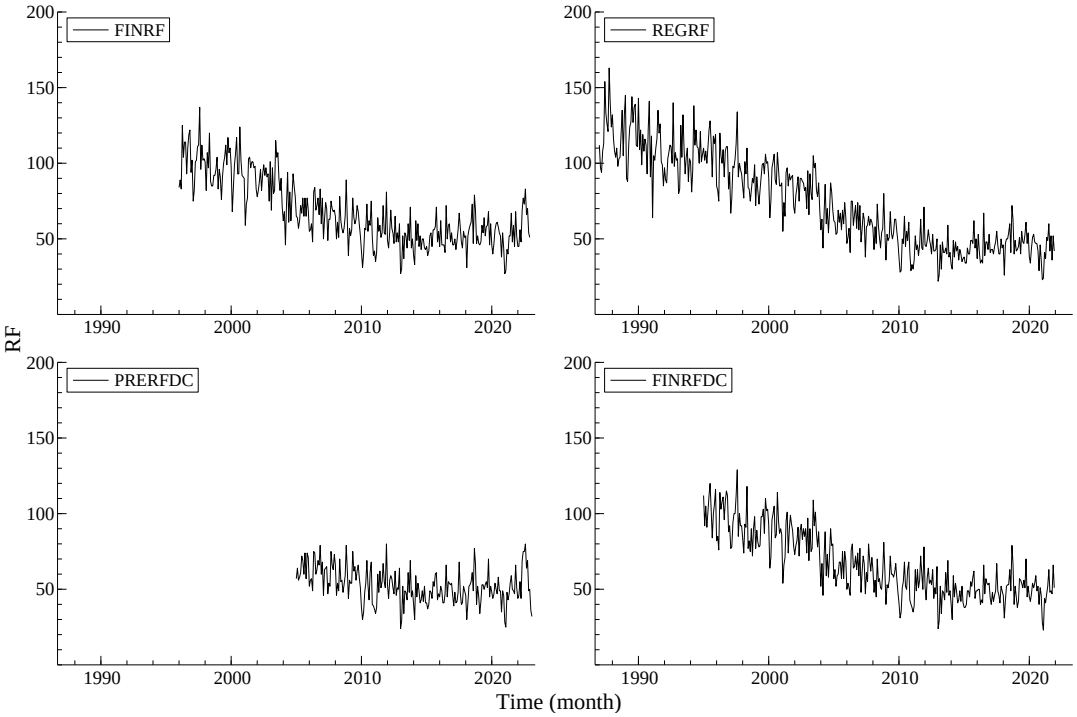

Figure 1: Available monthly observations of the time series on the number of road fatalities.

The first data that become available in the Netherlands are those on the preliminary number of road fatalities among Dutch citizens (PRERFDC), as they are updated every throughout the year. This is followed by the registered number of road fatalities (REGRF) a couple of months later, and then a couple of weeks later by the final number of road fatalities (FINRF). The final number of road fatalities among Dutch citizens (FINRFDC) are then the last to become available. The general aim of this empirical study is to investigate how to use

the preliminary data in the PRERFDC and in the REGRF series to improve the precision of what to expect for the final FINRF data before they are officially issued by Statistics Netherlands.

The data used for this empirical study are shown in Figure 1 and cover the months of 1996(1)-2022(12) for FINFR, 1987(1)-2021(12) for REGRF, 2005(1)-2023(2) for PRERFDC, and 1995(1)-2021(12) for FINRFDC.

3.2 Road Safety model

Since the four monthly time series all refer to the same process, that is, the underlying latent aggregate road safety process of the Netherlands, it seems reasonable to assume that the state disturbances of these four series are driven by the same dynamic process. We therefore impose the restriction that the level disturbances of the four series exhibit perfect linear dependence, and we apply the same restriction to the slope disturbances and the seasonal disturbances. This results in a SUTSE model with common levels, common slopes (and therefore common trends), and common seasonal components, which is obtained by restricting the disturbance covariance matrices Σ_ξ, Σ_ζ and Σ_ω to be all of *rank one*. As a result, the road safety model can be described as follows:

$$y_{i,t} = \theta_{\mu,i}\,\mu_t^{\dagger} + \theta_{\nu,i}\mu_t^{\dagger\dagger} + \theta_{\gamma,i}\,\gamma_t + \varepsilon_{i,t}, \qquad \varepsilon_{i,t} \sim \text{NID}\left(0, \sigma^2_{\varepsilon_i}\right), \quad i = 1, \ldots, 4, \tag{5}$$

where $y_{1,t}$ is the time series variable of interest (FINRF), $y_{2,t}$, $y_{3,t}$ and $y_{4,t}$ are the auxiliary time series of REGRF, PRERFDC and FINRFDC, respectively, $\mu_t^{\dagger}$ is the common level, $\mu_t^{\dagger\dagger}$ is the common trend, γ_t is the common seasonal and $\varepsilon_{i,t}$ is the irregular component which is assumed to be contemporaneously correlated across $i = 1, \ldots, 4$, the loading coefficients $\theta_{\mu,i}$, $\theta_{\nu,i}$ and $\theta_{\gamma,i}$ are fixed, and variance $\sigma^2_{\varepsilon_i}$ is strictly positive. Typical additional vectors in the SUTSE model, containing fixed intercepts, linear trends and seasonal effects to capture deterministic deviations from a common component, are not present in the Road Safety model because all four time series are based on the same well-defined definition of "road fatalities". The purpose of this model is to produce timely nowcasts and forecasts for the variable FINRF.

The common level $\mu_t^{\dagger}$ is a dynamic stochastic level and specified as the local level model of Harvey (1989),

$$\mu_{t+1}^{\dagger} = \mu_t^{\dagger} + \xi_t, \qquad \xi_t \sim \text{NID}\left(0, \sigma^2_{\xi}\right), \tag{6}$$

where the disturbance ξ_t is mutually independent (at all lags) of all other disturbances in the model. The common trend $\mu_t^{\dagger\dagger}$ contains a dynamic stochastic slope ν_t and is specified as the smooth trend model as given by

$$\begin{aligned} \mu_{t+1}^{\dagger\dagger} &= \mu_t^{\dagger\dagger} + \nu_t, \\ \nu_{t+1} &= \nu_t + \zeta_t, \qquad \zeta_t \sim \text{NID}\left(0, \sigma^2_{\zeta}\right), \end{aligned} \tag{7}$$

where disturbance ζ_t is mutually and serially independent of all other disturbances. The common seasonal is specified in trigonometric form, also see Harvey (1989),

$$\gamma_t = \sum_{j=1}^{(s/2)} \gamma_{j,t}, \tag{8}$$

where

$$\begin{aligned} \gamma_{j,t+1} &= \quad \cos(\lambda_j)\gamma_{j,t} + \sin(\lambda_j)\gamma^*_{j,t} + \omega_{j,t}, \qquad \lambda_j = 2\pi\, j/s, \\ \gamma^*_{j,t+1} &= -\sin(\lambda_j)\gamma_{j,t} + \cos(\lambda_j)\gamma^*_{j,t} + \omega^*_{jt}, \qquad j = 1, \ldots, (s/2), \end{aligned} \tag{9}$$

in which $\omega_{j,t}$ and $\omega^*_{j,t}$ are independent NID $(0, \sigma^2_\omega)$ disturbance terms (mutually and serially independent, also of all other disturbances) and $s = 12$ since we have monthly data. The irregular vector $(\varepsilon_{1,t}, \ldots, \varepsilon_{4,t})'$ is NID with a full non-singular covariance matrix which can be estimated by the method of maximum likelihood, together with the other parameters. The loadings or weights for the level, trend and seasonal components in each series are $\theta_{\mu,i}$, $\theta_{\nu,i}$ and $\theta_{\gamma,i}$, respectively, with $\theta_{\mu,1} = \theta_{\nu,1} = \theta_{\gamma,1} = 1$. The default option for the Road Safety model is to set $\theta_{\mu,i} = \theta_{\nu,i} = \theta_{\gamma,i} = 1$, for $i = 2, 3, 4$. Alternative option is to estimate the loadings of the auxiliary series $\theta_{\mu,i}$, $\theta_{\nu,i}$ and $\theta_{\gamma,i}$, for $i = 2, 3, 4$, by maximum likelihood, together with the variances for the level, trend, and seasonal, plus the irregular covariance matrix. A SUTSE model with stochastic level and with deterministic slope and seasonal effects, delivers the most parsimonious version and is obtained by restricting $\sigma^2_\zeta = \sigma^2_\gamma = 0$. The performance of these models is evaluated both in-sample and in terms of forecasting precision in the following subsections. A discussion on how to cast the Road Safety model into state space form can be found in Appendix B.1.

3.3 In-sample fit comparisons

Six variations on the general statistical early warning model are selected by applying different restrictions on the covariance matrix structure of the three unobserved components. In this section, they are compared based on the entire sample, before testing their performance in terms of forecasting precision. The covariance matrix of the irregular $\varepsilon_{i,t}$ for $i = 1, \ldots, 4$, in (5) is a full matrix in all models:

1. The first most basic model specifies one common stochastic level, while the slope and seasonal components are deterministic. The level loadings $\theta_{\mu,i} = 1$ for all i.

2. The second model introduces a stochastic common slope and trigonometric seasonal component to the model, where all loadings are fixed to be equal to one.

3. The third model estimates the loadings of the auxiliary series for the slope and seasonal components freely by maximum likelihood, instead of fixing them to one.

4. The fourth model resembles model 1, but the level loadings of the auxiliary series are estimated by maximum likelihood instead of fixed to one.

5. The fifth model extends model 4 by reintroducing stochastic common slope and seasonal components, while their auxiliary loadings are fixed to one.

6. The sixth model is the most extensive model which includes all common unobserved components and where the loadings of the auxiliary series are freely estimated for all.

Table 1 evaluates the in-sample performance of the models. Model 6 has the best model fit since it has the lowest AIC. The model fit of models 3, 4 and 5 are almost as good as model 6 since their AIC values are only slightly bigger. Based on the AIC of Models 1 and 2, they are clearly the worst fitting models. The model diagnostics for the standardized innovations

are presented per series and test for autocorrelation, heteroscedasticity and normality. These tests would all be insignificant under correct model specification. Overall, there seems to be no serial correlation present, but the homoscedasticity and normality hypotheses seem violated for the first and second series (FINRF and REGRF) for some of the models.

Table 1: Model diagnostics for Road Safety monitor series: Log likelihood, AIC, and tests on normality, autocorrelation and heteroscedasticity for FINRF ($y_{1,t}$), REGRF ($y_{2,t}$), PRERFDC ($y_{3,t}$) and FINRFDC ($y_{4,t}$).

Model		1	2	3	4	5	6
Series	Statistic						
	Log L	1212.121	1214.601	1230.019	1232.626	1235.745	1237.712
	AIC	-7.789	-7.792	-7.891	-7.920	-7.928	-7.940
$y_{1,t}$	Ljung Box	32.011	31.503	32.269	30.111	31.061	32.158
	H(h)	1.497*	1.457*	1.468*	1.353	1.303	1.318
	Skewness	-0.176	-0.126	-0.168	-0.108	-0.079	-0.063
	Kurtosis	3.771**	3.616*	3.789**	3.813**	3.733**	3.695**
	Normality	9.313**	5.738	9.525**	9.169*	7.283*	6.466*
$y_{2,t}$	Ljung Box	35.462	36.568	30.785	29.617	31.716	29.703
	H(h)	2.221**	2.150**	2.088**	2.008**	1.925**	1.912**
	Skewness	-0.292*	-0.250*	-0.283*	-0.239*	-0.200	-0.179
	Kurtosis	3.916**	3.742**	4.032**	3.978**	3.876**	3.880**
	Normality	20.026**	13.575**	23.489**	20.094**	15.717**	15.308**
$y_{3,t}$	Ljung Box	38.912	39.618	38.296	36.158	39.944	37.518
	H(h)	1.315	1.279	1.323	1.185	1.165	1.112
	Skewness	-0.237	-0.184	-0.244	-0.148	-0.113	-0.133
	Kurtosis	3.520	3.322	3.526	3.483	3.342	3.306
	Normality	4.219	2.039	4.399	2.742	1.434	1.400
$y_{4,t}$	Ljung Box	31.202	35.616	30.615	27.135	33.013	34.679
	H(h)	1.218	1.176	1.231	1.115	1.055	1.059
	Skewness	-0.285*	-0.229	-0.300*	-0.207	-0.185	-0.188
	Kurtosis	3.745**	3.476	3.751**	3.724**	3.534	3.477
	Normality	11.398**	5.657	11.985**	9.011*	5.475	4.783

$^{**}p < 0.01$, $^{*}p < 0.05$

In order to explore this further, a visual representation of the residual based diagnostics of model 4 is given in Figure 2. While the figures for the second series REGRF seem to contradict the test results on normality from Table 1, this could be explained by the larger amount of observations for this time series. The power of the tests to detect deviations from the null hypothesis increases with the sample size of the observed time series. Visually, the diagnostic checks show nothing alarming in terms of normality violations, autocorrelation, heteroscedasticity or drifting CUSUM for any of the series. This holds for all six models, as they all resemble Figure 2. See Figure B1 and Figure B2 in Appendix B, which show the visual innovation diagnostics for models 2 and 6 as example.

An overview of the estimated hyperparameters based on the entire sample between 1987-2023 is given in Table B1. Since the covariance matrix of the irregular is a full matrix in all models, we present the individual irregular variances from (5) and the correlations of these idiosyncratic noise terms between each series. Given that the four series represent similar

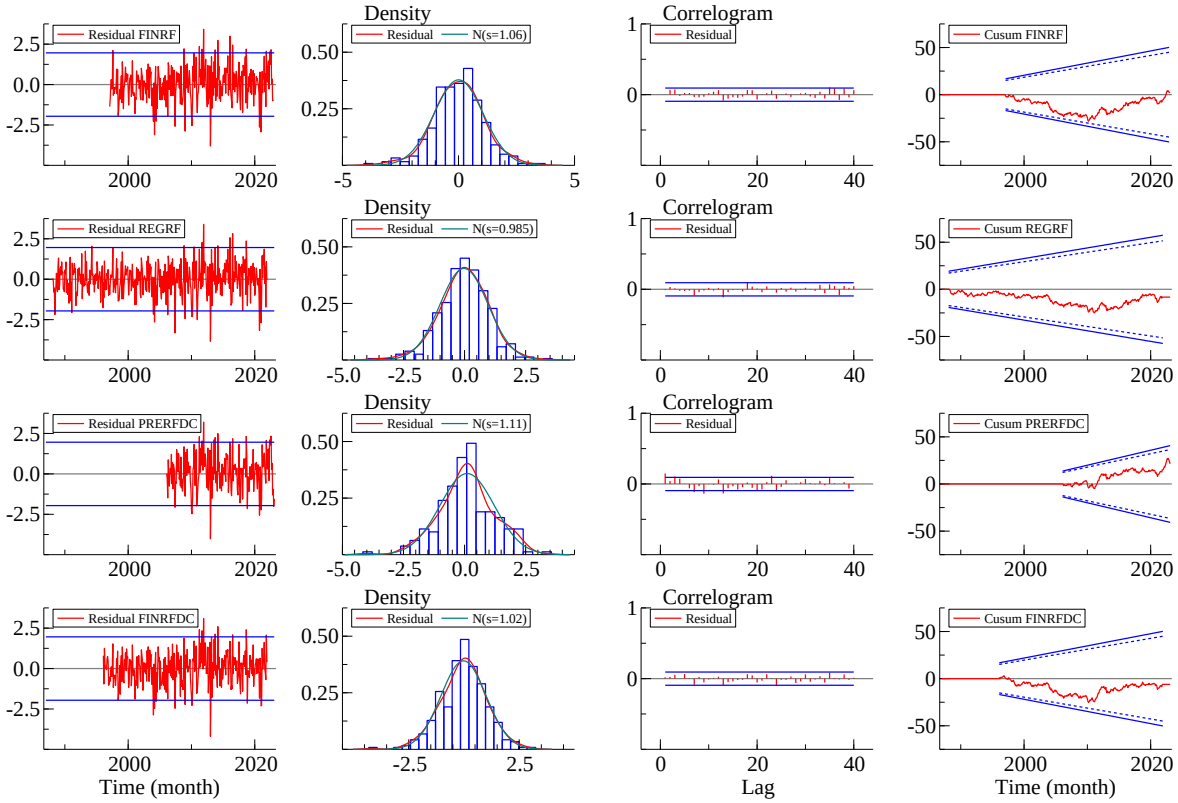

Figure 2: Innovations diagnostics of Model 4, obtained with Kalman smoother.

processes, it is expected to observe such high correlations. The high loading estimates in model 3 are the result of the small values for the corresponding variances of the common slope and seasonal disturbance terms. Further note that the parameter estimates for the variances of the common slope and seasonal disturbance terms in the more extensive models 5 and 6 tend towards zero as well. This implies that these stochastic extensions of model 4 do not provide much added benefit. Given that their AIC's were quite similar, we consider the more parsimonious model 4 as the best performing model in-sample.

Estimates of the unobserved components are presented in Figure 3. The estimates of the slope and seasonal components are time-invariant for this model specification. At the beginning of the series there are long periods with missing observations. As a result the standard errors of the smoothed trend are large or even missing at the beginning. Similar graphs for the other models can be found in Figures B3 - B7. It shows overall that multivariate models with common factor restrictions are able to decompose and capture the unobserved components that drive the dynamic progression of these time series. We notice that when loadings of a common factor are fixed to 1, the four component estimates should resemble each other, which can be seen in the plots of Model 2 and 5.

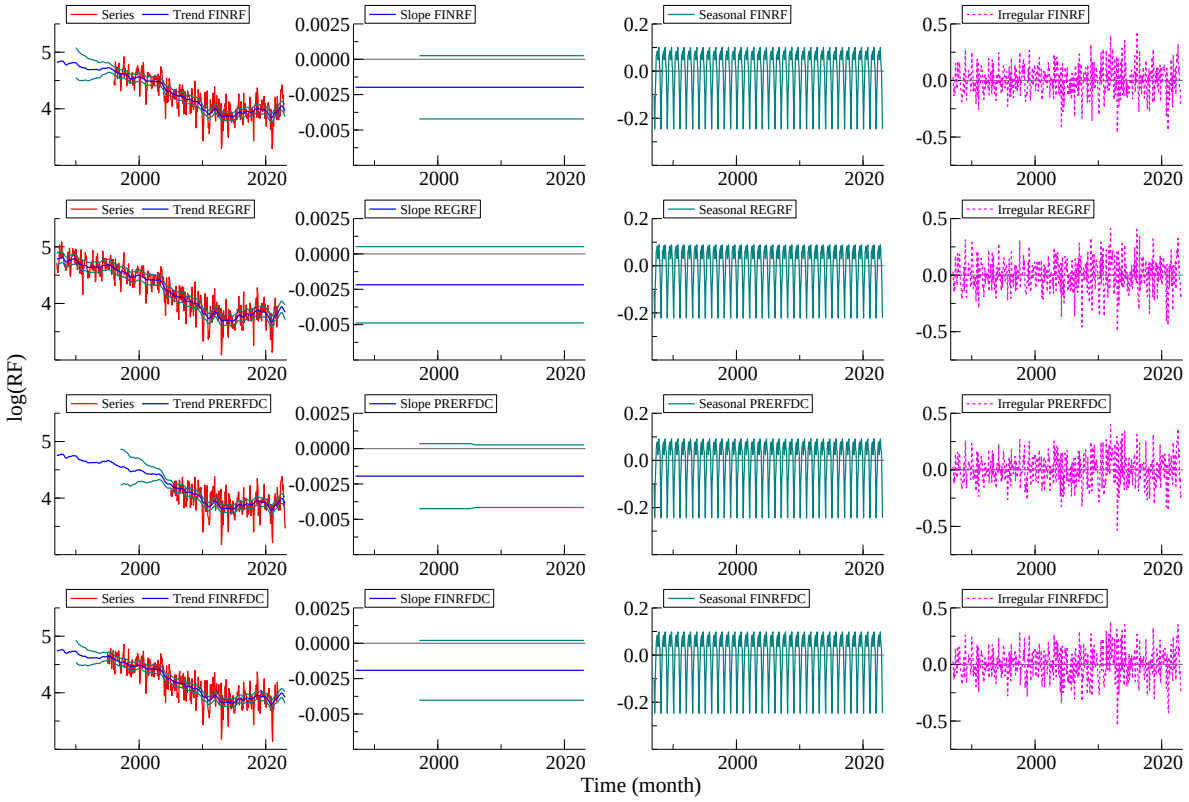

Figure 3: State vector estimates in Model 4, obtained with Kalman smoother.

3.4 Forecasting evidence

In order to investigate the relative precision of the forecasts of the final monthly number of fatalities (FINRF) in a certain year, say year X, we fitted the six SUTSE models with common trends and seasonal components presented in the previous section using five different scenarios.

In each scenario the forecasts for FINRF are calculated for the twelve months of year X, and then these forecasts are compared with the actual data for FINRF in year X, which had not been used when fitting the model. The first scenario is a benchmark to investigate whether the availability of more and more recent data for the *other series than FINFR* would indeed result in more precise forecasts. Thus for the benchmark the 1- to 12-step ahead forecasts for FINFR are calculated for the 12 months of year X using a univariate basic structural time series model with deterministic (DBSTSM) and stochastic (SBSTSM) slope and seasonal components, fitted to the FINFR monthly data observed up to and including year $X-1$.

All scenarios are meant to mimic different moments throughout the year at which SN and RWS publish new data. Only PRERFDC is updated on a monthly basis, with a publication lag of three months. FINRF, REGRF and FINRFDC are updated yearly. For example, our target series FINRF is updated around the end of April, hence at $X(4)$ all twelve monthly observations of $X-1$ become available. Scenarios 2, 3, 4 and 5 forecast FINRF using our proposed multivariate SUTSE models at $X(5)$, $X(10)$, $X{+}1(1)$ and $X{+}1(4)$ respectively and Table 2 shows the most recent observation of the auxiliary series at those moments in time.

Table 2: Time series availability (✓) in the forecast scenarios.*

Scenario	Forecast moment	Series	Month								
			$X-2(12)$	$X-1(3)$	$X-1(6)$	$X-1(9)$	$X-1(12)$	$X(3)$	$X(6)$	$X(9)$	$X(12)$
1	$X(4)$	FINRF	✓	✓	✓	✓	✓				
2	$X(5)$	FINRF	✓	✓	✓	✓	✓				
		REGRF	✓	✓	✓	✓	✓				
		PRERFDC	✓	✓	✓	✓	✓				
		FINRFDC	✓								
3	$X(10)$	FINRF	✓	✓	✓	✓	✓				
		REGRF	✓	✓	✓	✓	✓				
		PRERFDC	✓	✓	✓	✓	✓	✓	✓		
		FINRFDC	✓								
4	$X+1(1)$	FINRF	✓	✓	✓	✓	✓				
		REGRF	✓	✓	✓	✓	✓				
		PRERFDC	✓	✓	✓	✓	✓	✓	✓	✓	
		FINRFDC	✓								
5	$X+1(4)$	FINRF	✓	✓	✓	✓	✓				
		REGRF	✓	✓	✓	✓	✓	✓	✓	✓	✓
		PRERFDC	✓	✓	✓	✓	✓	✓	✓	✓	✓
		FINRFDC	✓	✓	✓	✓	✓				

*The year number is X, the previous year is $X-1$, and the number between brackets refers to the month.

Recall from the notation introduced at the end of Section 2 that Ω_t contains the data available at period t and that $Y_t = (y_1, \ldots, y_t)'$. In scenario 1, $\Omega_t = (Y_{1,t-j})$ for $j = 1, \ldots, 12$ depending on the month of forecast year X. In scenario 2, $\Omega_t = (Y_{1,t-j}, Y_{2,t-j}, Y_{3,t-j}, Y_{4,t-j-12})$, in scenario 3, $\Omega_t = (Y_{1,t-j}, Y_{2,t-j}, Y_{3,t-j+6}, Y_{4,t-j-12})$, in scenario 4, $\Omega_t = (Y_{1,t-j}, Y_{2,t-j}, Y_{3,t-j+9}, Y_{4,t-j-12})$ and in scenario 5, $\Omega_t = (Y_{1,t-j}, Y_{2,t-j+12}, Y_{3,t-j+12}, Y_{4,t-j})$.

All analyses were applied to the logarithms of the observations. To assess the forecast precision of the proposed models we applied a real-time forecasting study to the years $X = 2011, \ldots, 2022$ and computed the well-known criteria for precision: the mean squared error (MSE), the mean absolute error (MAE) and the mean absolute percentage error (MAPE). The resulting MAPE for each model in the five scenarios are presented in Table 3, while the other criteria can be found in Table B2 and Table B4. Each column in Table 3 contains the MAPE averaged over the 12 months of a particular year. The final column presents the average of all relative forecast errors over the entire period of the real time analysis. Note that the model parameters are re-estimated each year, as there are more observations available to fit our models on when we progress through the forecasting years. How the hyperparameter estimates of model 4 in scenario 3 develop over the years can be found in Table B5 as an example.

When comparing the forecasting precision of the univariate models in scenario 1 with the multivariate models of scenario 2 no clear increase in performance is found. However, scenario 2 does not utilize the timeliness of the auxiliary series yet, this is what scenarios 3, 4 and 5 are designed for. The multivariate SUTSE models in the final three scenarios clearly outperform the univariate structural time series models.

Table 3 averaged the MAPE for the h-step ahead forecasts for $h = 1, \ldots, 12$ within each year. Table 4 presents the MAPE forecasts for the 12 months or the 12 different h-step ahead forecasts, averaged over the years. This provides a comprehensible illustration of the

Table 3: MAPE forecasting performance based on the 1- to 12-step ahead forecasts per forecasting year and in total.

Scenario	Model	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Total
1	DBSTSM	3.444	3.092	5.083	2.806	3.827	3.721	3.098	4.829	2.362	2.368	4.393	5.563	3.716
	SBSTSM	3.423	3.092	5.083	2.806	3.827	3.721	3.098	4.829	2.362	2.368	4.393	5.563	3.714
2	1	3.542	3.147	5.512	3.082	3.475	3.509	2.895	4.751	2.483	2.347	4.505	5.380	3.719
	2	3.495	3.078	5.319	2.824	3.857	3.662	3.169	4.794	2.446	2.405	4.460	5.427	3.745
	3	3.614	3.046	5.018	2.939	3.753	3.773	3.667	4.794	2.545	2.340	4.595	5.463	3.796
	4	3.619	3.074	5.127	2.920	3.903	3.644	2.930	4.552	2.276	2.530	4.517	5.780	3.739
	5	3.494	3.184	5.139	2.799	3.821	3.761	3.071	4.254	2.116	3.104	5.081	4.990	3.735
	6	3.515	3.782	4.835	2.786	4.102	4.177	3.882	5.025	3.298	2.607	4.369	6.753	4.094
3	1	2.294	2.967	2.308	2.046	2.598	2.728	2.708	3.453	2.771	2.886	3.434	2.155	2.696
	2	2.349	2.978	2.104	2.359	2.587	2.607	2.742	3.412	2.684	3.013	3.404	2.027	2.689
	3	2.586	2.761	1.927	2.317	2.754	2.515	2.888	3.135	2.518	3.357	3.523	2.271	2.713
	4	2.322	2.723	2.153	1.898	3.060	2.875	2.814	3.342	2.608	3.202	3.505	2.163	2.722
	5	2.460	2.654	2.119	2.029	2.859	2.698	2.750	3.410	2.752	3.052	3.432	2.442	2.721
	6	2.436	2.943	1.983	2.254	3.172	2.421	2.814	3.010	2.422	3.962	3.660	2.628	2.809
4	1	2.272	2.051	1.860	2.088	1.551	1.947	2.202	2.309	2.695	2.775	3.176	1.403	2.194
	2	2.305	2.030	1.726	2.135	1.493	1.793	2.228	2.300	2.621	2.882	3.196	1.403	2.176
	3	2.565	1.828	1.542	2.168	1.448	1.673	2.487	2.137	2.589	3.259	3.287	1.483	2.206
	4	2.210	1.700	1.658	1.955	1.887	2.147	2.335	2.241	2.619	3.172	3.199	1.445	2.214
	5	2.345	1.858	1.649	2.088	1.592	1.952	2.324	2.463	2.676	2.875	3.193	1.090	2.176
	6	2.599	1.751	1.672	2.009	2.743	2.741	3.513	3.688	3.714	2.758	3.210	3.355	2.813
5	1	1.052	1.295	0.816	1.269	0.691	1.431	0.961	0.950	1.188	0.886	1.757	0.649	1.079
	2	1.075	1.286	0.786	1.250	0.695	1.409	0.961	0.940	1.192	0.883	1.763	0.646	1.074
	3	0.995	1.694	0.890	1.463	1.283	1.677	1.710	1.557	1.771	1.058	2.163	1.977	1.520
	4	0.725	1.217	0.835	1.236	1.047	1.638	1.252	1.415	1.811	1.739	1.726	2.335	1.415
	5	0.739	1.379	0.783	1.256	0.820	1.492	0.989	0.931	1.167	0.821	1.760	0.615	1.063
	6	1.218	1.451	0.752	1.177	0.734	1.365	1.093	0.987	1.202	1.037	1.891	0.801	1.142

Table 4: MAPE forecasting performance for 1- to 12-steps ahead forecasts per month.

Scenario	Model	Jan $t+1$	Feb $t+2$	Mar $t+3$	Apr $t+4$	May $t+5$	June $t+6$	July $t+7$	Aug $t+8$	Sept $t+9$	Oct $t+10$	Nov $t+11$	Dec $t+12$
1	DBSTSM	4.202	5.135	3.246	3.992	2.148	3.010	4.789	3.076	3.788	4.928	2.701	3.572
	SBSTSM	4.193	5.124	3.234	4.004	2.135	3.025	4.772	3.094	3.805	4.907	2.679	3.592
2	1	4.364	5.078	3.083	3.889	2.154	3.213	4.812	2.910	3.921	5.017	2.741	3.444
	2	4.404	5.204	2.931	4.039	2.171	3.243	4.868	3.068	3.877	4.993	2.651	3.487
	3	4.359	5.158	3.129	3.971	2.341	3.248	4.965	3.101	3.833	4.944	2.876	3.621
	4	4.096	5.008	3.257	3.904	2.405	2.975	4.664	3.267	4.020	4.839	2.636	3.800
	5	4.633	5.533	3.330	3.869	2.171	2.774	4.634	3.365	3.683	5.070	2.551	3.203
	6	4.502	5.315	4.354	4.307	2.480	3.395	4.965	3.878	3.830	5.068	2.558	4.479
3	1	1.817	2.279	1.515	1.172	1.347	2.425	4.328	2.816	3.549	4.860	2.696	3.543
	2	1.673	2.302	1.499	1.184	1.295	2.431	4.300	2.922	3.523	4.889	2.733	3.514
	3	1.576	2.461	1.639	1.248	1.083	2.369	4.472	3.161	3.342	4.887	2.790	3.526
	4	1.593	2.526	1.593	1.345	1.441	2.373	4.271	3.058	3.405	4.671	2.544	3.845
	5	1.510	2.299	1.741	1.359	1.302	2.740	4.255	3.157	3.477	5.023	2.477	3.316
	6	1.481	2.658	1.887	1.536	1.210	2.622	4.390	3.546	3.470	4.647	2.640	3.616
4	1	1.828	2.299	1.496	1.175	1.316	2.395	1.989	1.607	1.366	4.343	3.247	3.267
	2	1.675	2.325	1.496	1.200	1.267	2.397	1.971	1.609	1.354	4.405	3.122	3.293
	3	1.605	2.471	1.631	1.143	1.137	2.342	2.019	2.006	1.166	4.319	3.184	3.444
	4	1.632	2.510	1.599	1.377	1.478	2.321	1.854	1.734	1.281	4.206	2.985	3.592
	5	1.488	2.302	1.663	1.272	1.300	2.673	1.815	1.832	1.287	4.867	2.554	3.054
	6	1.761	2.696	2.443	2.079	1.782	3.596	2.329	2.665	2.455	4.961	2.984	4.003
5	1	1.066	0.921	1.055	1.393	1.233	1.451	1.111	1.110	0.522	1.044	1.350	0.690
	2	1.002	0.915	1.067	1.385	1.230	1.453	1.120	1.108	0.527	1.033	1.343	0.702
	3	1.407	1.197	1.443	1.803	1.628	2.048	1.426	1.632	0.908	1.357	2.043	1.344
	4	0.976	1.416	1.338	1.547	1.746	1.803	1.159	1.526	1.369	1.266	1.560	1.272
	5	0.879	1.116	1.174	1.221	1.379	1.418	1.062	1.183	0.638	0.896	1.169	0.618
	6	1.019	0.917	1.203	1.376	1.344	1.662	1.159	1.231	0.650	0.992	1.380	0.774

advantage of a multivariate extension for the 12 different forecast windows. The results in forecast scenario 3 and 4 clearly show a decrease in forecast accuracy for the largest horizons, for which there are no provisional observations for PRERFDC yet. However, the accuracy of those months still improves upon the univariate benchmarks for some of the models. For the shorter horizons it is clear that a multivariate model utilizes the more timely available provisional data. As a result, multivariate models provide much more accurate forecasts than univariate models.

Figure 4 illustrates these findings by comparing the forecasts of model 4 in the different scenarios with the realized target series observations. Moreover, it clearly shows that scenario 5 outperforms the rest in the final months of each year, which holds for the forecasts of all models as shown in Figures B8 - B12. This is in line with the construction of the scenarios, since it is the only scenario that uses auxiliary information of REGRF and PRERFDC up until the last quarter of X. The first graph of Figure B13 presents the corresponding forecast uncertainty. These forecast standard errors also follow a similar pattern. The standard errors in scenarios 1 and 2 are very similar and the highest overall. In scenario 3 and 4 a jump is visible at the 7- and 10-step ahead predictions, which are caused by the design of these scenarios since they are the subsequent months of the most recent monthly observation of PRERFDC. Naturally, the errors become slightly higher within each year, because a 1-step ahead prediction is made with more precision than a 12-step ahead prediction.

If we compare the performance between the models within a scenario in Table 3, we find that model 4 is actually one of the worst performing models in terms of forecasting precision. Model 2 outperforms the rest in scenario 3, while model 5 does best in scenario 5. Both

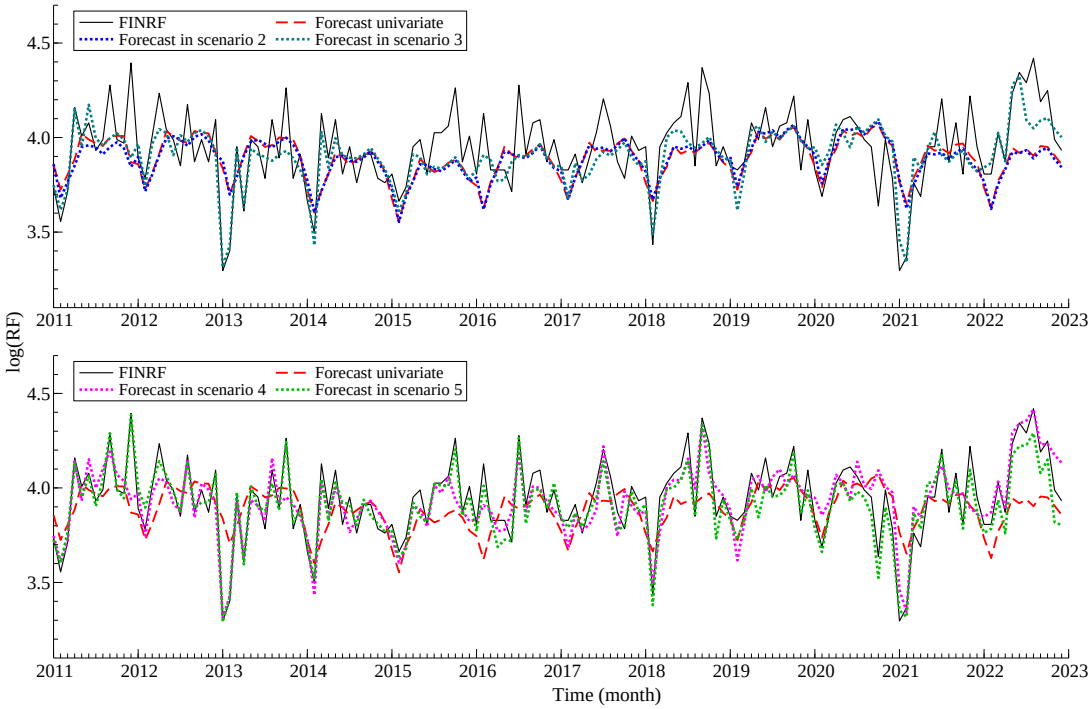

Figure 4: 1- to 12-step ahead forecasts of the FINRF series of Model 4, in the different forecasting scenarios over the forecasting years 2011-2022.

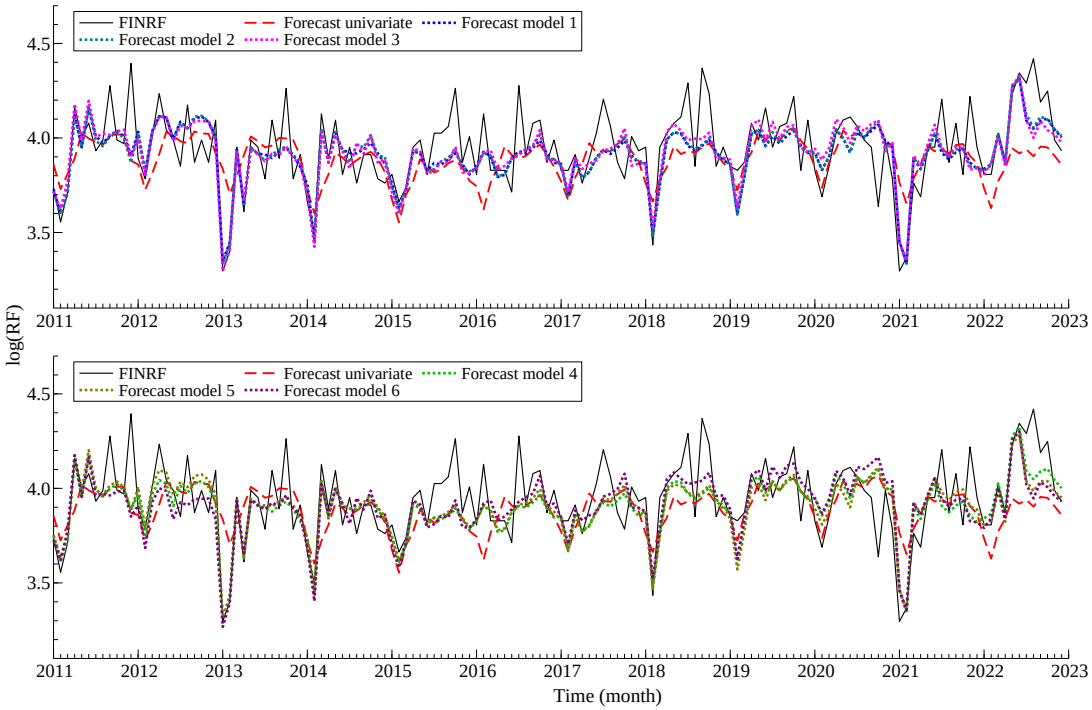

Figure 5: 1- to 12-step ahead forecasts of the FINRF series of the different models in scenario 3, over the forecasting years 2011-2022.

models present similar results in scenario 4. However, note that the MAPE is expressed as a percentage and the performance differences between models are much smaller than the differences between the scenarios. When considering the MSE and MAE in Table B2 and Table B4, it is shown that no model outperforms another model convincingly. Scenario 3 can arguably be considered as the most realistic one, when the forecasts are made in the month of October, in the year of interest. Figure 5 confirms the minor differences in the point estimates of each model visually for this scenario, but it holds in all multivariate scenarios, see Figures B14 - B16. Nevertheless, the second graph of Figure B13 shows that model 5 has slightly less uncertainty in scenario 3, therefore it could be considered as the preferred model when forecasting the target series FINRF. The combined forecasts and uncertainty of model 5 in scenario 3 and 5 are presented in Figure 6, it shows the gain in forecast accuracy with models that use information from correlated auxiliary series.

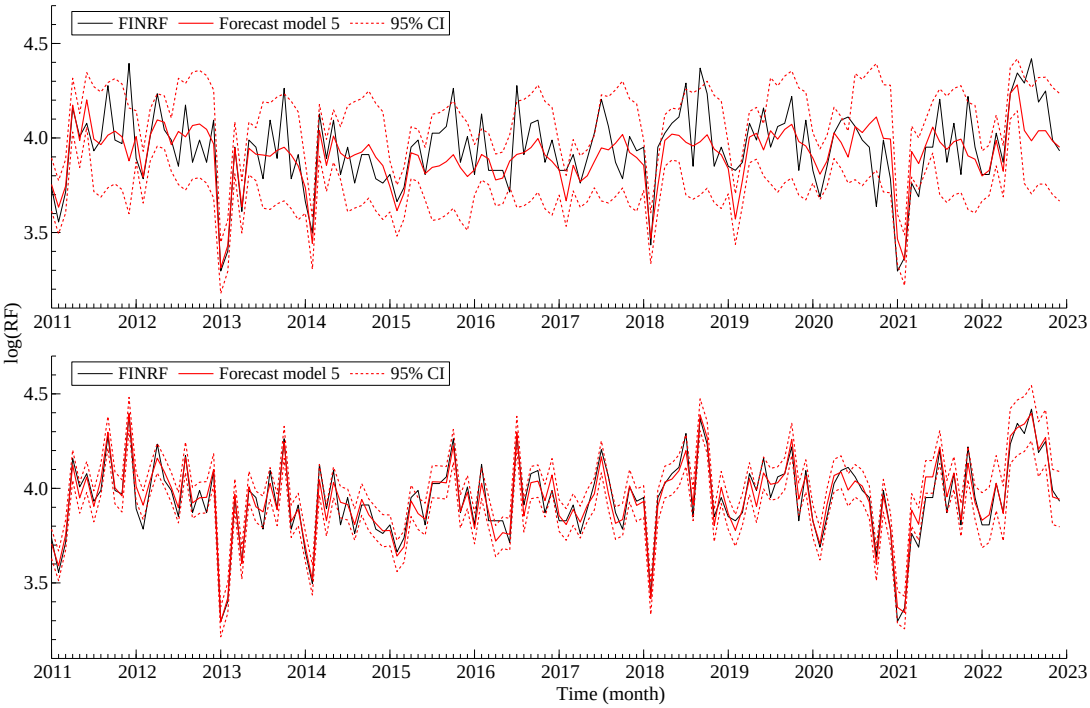

Figure 6: 1- to 12-step ahead forecasts with uncertainty of the FINal Road Fatalities series on a log scale, obtained with the best performing model 5 in scenario 3 (top) and scenario 5 (bottom).

4 Empirical Study II: Labour Force participation

4.1 Data

The data of this study consists of two time series with different observation frequencies. Both series capture the labour force participation (LFP), see Figure 7:

1. the LFP rate in the municipality of Amsterdam derived from the Tax Administration and Employee Insurance Agency registers (Register), on a quarterly frequency.

2. the LFP rate in the Netherlands estimated from the Labour Force Survey (LFS) issued by Statistics Netherlands, on a monthly frequency.

Typically, data published by the Tax Administration suffer from long publication delays. Self-employed taxpayers can opt for a delayed tax declaration, hence the final tax assessment can take several years. It takes approximately nine months after the reference year has ended before the first preliminary data is published, while it can take up to another year thereafter before the final figures are determined. On the contrary, the LFS is conducted every month, so the LFP rate can be obtained much faster from that source, see the discussion in van den Brakel and Michiels (2021).

The benefit from determining the LFP rate on the basis of the registers however is that this data cover the entire target population and can be used to derive very detailed

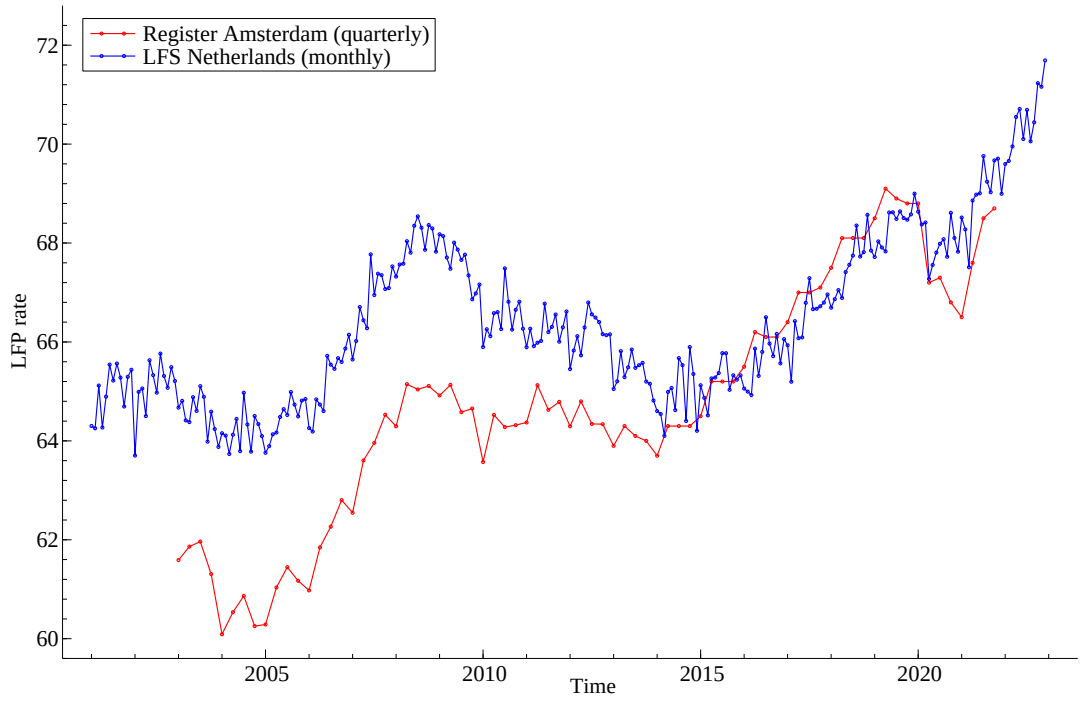

Figure 7: Available time series of quarterly labour force participation (LFP) rate in Amsterdam derived from the tax register and monthly LFP rate in the Netherlands obtained from the labour force survey.

regional figures. The LFS, on the other hand, is a sample which implies that LFP estimates contain sampling error. As a result, reliable monthly indicators can only be produced at high aggregation levels. van den Brakel and Michiels (2021) discuss inference methods that are used to generalize these sample findings to the target population. The general aim of this empirical study is to investigate how the noisy fast higher-frequent LFS series can be used to improve the precision of what to expect of the more precise slow lower-frequent LFP rate series from the tax register before their rates are officially issued. Figure 7 shows the data set of quarterly LFP rates for period 2003(1)-2021(4) from the "Register" and of monthly LFP rates for period 2001(1)-2022(12) from the LFS of SN.

4.2 Labour Force model

Let y_t denote the LFP rate for Amsterdam for quarter t, derived from the Tax Register and $x_{[\tau]}$ the estimate for the LFP rate at the national level of the Netherlands for month τ obtained from the LFS. The target variable of interest is the quarterly time series y_t, which will be combined in a model with the auxiliary monthly time series $x_{[\tau]}$ with the purpose to produce more timely nowcasts for y_t. The LFS is based on a rotating panel design. Each month a stratified simple random sample of persons enter the panel. This panel is observed five times at quarterly intervals. A sample that enters the panel in month τ is observed at months $\tau, \tau+3, \tau+6, \tau+9$ and $\tau+12$. After the fifth interview at $\tau+12$, the sample leaves the panel. This creates serial correlation between $x_{[\tau]}$ and $x_{[\tau+3j]}$, for $j = 1, \ldots, 4$.

The publication frequency difference of both time series restricts us from fitting a bivariate SUTSE model to the data straightforwardly. Many approaches exist in the literature for handling mixed frequency data. We take the approach of stacking the monthly observations in a quarterly 3×1 vector series, and incorporate this vector within a multivariate dynamic process of a quarterly frequency, see Blasques, Koopman, Mallee, and Zhang (2016) for the details. It is shown that low- or high-frequency representations of any linear dynamic process lead to equivalent maximized log-likelihood values and parameter estimates. The following quarterly notation for the monthly LFS series is adopted:

$$\boldsymbol{x}_t = \begin{pmatrix} x_{t,1} \\ x_{t,2} \\ x_{t,3} \end{pmatrix} = \begin{pmatrix} x_{[\tau]} \\ x_{[\tau+1]} \\ x_{[\tau+2]} \end{pmatrix}, \tag{10}$$

where $x_{t,i}$ refers to the i-th month within quarter t and τ is the corresponding monthly index. Then the Labour Force model can be written as a SUTSE model, that is

$$\boldsymbol{z}_t = \boldsymbol{\Theta}_\mu \boldsymbol{\mu}_t^\dagger + \boldsymbol{\mu}_\theta + \boldsymbol{\Theta}_\nu \boldsymbol{\mu}_t^{\dagger\dagger} + \boldsymbol{\mu}_{\theta,t} + \boldsymbol{\Theta}_\gamma \boldsymbol{\gamma}_t^\dagger + \boldsymbol{\gamma}_{\theta,t} + \boldsymbol{S}_t \boldsymbol{u}_t + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \text{NID}(\boldsymbol{0}, \boldsymbol{\Sigma}_\varepsilon), \tag{11}$$

where $\boldsymbol{z}_t = (y_t, \boldsymbol{x}_t')'$ is the 4×1 vector of observations, the components level $\boldsymbol{\mu}_t^\dagger$, trend $\boldsymbol{\mu}_t^{\dagger\dagger}$, seasonal $\boldsymbol{\gamma}_t^\dagger$ and irregular $\boldsymbol{\varepsilon}_t$ are defined similarly as in the model equation (5), but are now in bold as they are vectors, this also applies to the loading matrices $\boldsymbol{\Theta}_\mu$, $\boldsymbol{\Theta}_\nu$ and $\boldsymbol{\Theta}_\gamma$, vectors $\boldsymbol{\mu}_\theta$, $\boldsymbol{\mu}_{\theta,t}$ and $\boldsymbol{\gamma}_{\theta,t}$ capture the deterministic deviations from the common components for level, trend and seasonal, respectively, vector $\boldsymbol{S}_t \boldsymbol{u}_t$ accounts for the survey sampling structure in the data, and $\boldsymbol{\Sigma}_\varepsilon$ is the irregular covariance matrix. The loading matrices are imposed by a Cholesky decomposition for a rank deficient covariance matrix of the disturbance term, see Section 2.1 for further details. In particular, for a two factor model we have

$$\boldsymbol{\Sigma}_j = \boldsymbol{\Theta}_k \boldsymbol{\Sigma}_j^\dagger \boldsymbol{\Theta}_k' = \begin{bmatrix} 1 & 0 \\ \theta_{k,2,1} & 1 \\ \theta_{k,3,1} & \theta_{k,3,2} \\ \theta_{k,4,1} & \theta_{k,4,2} \end{bmatrix} \begin{bmatrix} \sigma_{j_1}^2 & 0 \\ 0 & \sigma_{j_2}^2 \end{bmatrix} \begin{bmatrix} 1 & \theta_{k,2,1} & \theta_{k,3,1} & \theta_{k,4,1} \\ 0 & 1 & \theta_{k,3,2} & \theta_{k,4,2} \end{bmatrix}, \tag{12}$$

with $(j,k) \in \{(\xi,\mu),(\zeta,\nu),(\omega,\gamma)\}$. In a similar way, the one factor model can be obtained straightforwardly by specifying $\boldsymbol{\Theta}_k = (1, \theta_{k,2,1}, \theta_{k,3,1}, \theta_{k,4,1})'$ and $\boldsymbol{\Sigma}_j^\dagger = \sigma_j^2$, where $\theta_{k,2,1}$, $\theta_{k,3,1}$ and $\theta_{k,4,1}$ can be interpreted as the loadings in (5) for $x_{t,1}$, $x_{t,2}$ and $x_{t,3}$, respectively, while σ_j^2 reflects the variance of the disturbance term. These unknown parameters can be estimated by the method of maximum likelihood or they can partly be restricted to zero or one. Finally, the dynamic specifications for the elements of the trend vectors $\boldsymbol{\mu}_t^\dagger$ and $\boldsymbol{\mu}_t^{\dagger\dagger}$ are given by (6) and (7), respectively. The dynamic specification for the elements of the seasonal component vector $\boldsymbol{\gamma}_t$ are specified as a dummy seasonal component, see Appendix C.1.

The sampling error effect is represented by the vector $\boldsymbol{S}_t \boldsymbol{u}_t$. The vector $\boldsymbol{u}_t = (u_{t,1}, u_{t,2}, u_{t,3})'$ contains the sampling errors of the monthly LFS estimates of $\boldsymbol{x}_t$. Note that the register variable y_t is based on a complete enumeration and is therefore observed without sampling error. The sampling errors of the monthly LFS estimates account for heteroscedasticity that arise from time-varying sample variance and autocorrelation induced by the rotating panel design of the LFS. This is achieved by scaling the sampling errors $u_{t,i}$ with the standard errors of $x_{t,i}$, that is $k_{t,i} = \sqrt{\text{Var}(x_{t,i})}$, hence $\boldsymbol{S}_t$ is a 4×3 scaling matrix where the first row contains zeros followed by $\text{diag}(k_{t,1}, k_{t,2}, k_{t,3})$. To account for serial correlation due to the sample

overlap of the rotating panel design, each $u_{t,i}$ is on its turn modelled as an AR(3) process with zero autocorrelation coefficients for the first two lags, i.e. $u_{t,i} = \phi u_{t-1,i} + e_{t,i}$. Note that ϕ can be estimated from the survey data, see also van den Brakel and Krieg (2015) or van den Brakel and Michiels (2021). They found that the AR(3) coefficient $\hat{\phi} = 0.59$, which is treated as known here. The 4×1 vector $\boldsymbol{\varepsilon}_t$ contains the idiosyncratic noise of the quarterly register series and monthly LFS series, where the latter is again stacked according to (10). Both are considered NID sequences, hence $\boldsymbol{\Sigma}_\varepsilon = \text{diag}(\sigma^2_{\varepsilon_y}, \sigma^2_{\varepsilon_x}, \sigma^2_{\varepsilon_x}, \sigma^2_{\varepsilon_x})$. Further details and a discussion on how the Labour Force model can be cast into state space form are provided in Appendix C.1.

4.3 Covid correction

The Covid-19 crisis had strong effects on the labour force figures. The start of the Covid crisis is indeed marked with sharp turning points in the monthly LFS figures. The crisis also resulted in a sudden increase of the monthly changes in the labour force figures (van den Brakel, Souren, and Krieg 2022). The flexibility of the time series model to pick up a sudden increase in the monthly changes is mainly determined by the variance of the level and slope disturbance terms. At the start of the Covid crisis, the values of these variances is based on the volatility of the period-to -period changes observed in the past. As a result the time series model will not be able to follow a sudden increase in this volatility, resulting in a temporarily misspecification of the time series model. The time series model can be adapted by temporarily increasing the flexibility of the trend. This can be achieved by increasing the variances of the level and slope disturbance terms of the trend by multiplying the maximum likelihood estimates of these variances with a time dependent factor which values are used as a-priory known values in the model. See van den Brakel, Souren, and Krieg (2022) and Gonçalves, Hidalgo, Silva, and van den Brakel (2022) for a detailed discussion.

The following time-varying covariance structures for the level and slope disturbance terms are considered:

$$\boldsymbol{\xi}_t \sim \text{NID}(\mathbf{0}, \sigma^2_{\xi,t}\boldsymbol{\Sigma}_\xi), \qquad \boldsymbol{\zeta}_t \sim \text{NID}(\mathbf{0}, \sigma^2_{\zeta,t}\boldsymbol{\Sigma}_\zeta), \tag{13}$$

where the variance factors $\sigma^2_{\xi,t}$ and $\sigma^2_{\zeta,t}$ are assumed to be known and are equal to one throughout the length of the sample, except for the period surrounding the outbreak at the start of 2020. To minimize the manual adjustments, the time-varying factors are quickly returning their values to one, typically after a few quarters.

4.4 In-sample fit comparisons

In this applications four variants of the Labour Force model are tested. First they are compared based on their in-sample performance. The variants are obtained by applying restrictions on the covariance matrix structures of the different unobserved components. The sampling error and irregular structures are kept diagonal in all models:

1. In the first model, both trend components and the dummy seasonal component are specified by one common factor. This is enforced by $\boldsymbol{\Sigma}_\xi$, $\boldsymbol{\Sigma}_\zeta$ and $\boldsymbol{\Sigma}_\omega$ all being rank one matrices where the loadings are freely estimated by maximum likelihood.

2. In the second model, both trend components are still specified by one common factor, but the seasonal component now has four independent factors. This is enforced by a diagonal $\mathbf{\Sigma}_\omega$ of rank four.

3. In the third model, the trend components are modelled by two common factors, while the seasonal component is again specified by one common factor.

4. In the fourth model, the trend components are specified by two common factors. The seasonal component is modeled by four independent factors.

Table 5 shows the model diagnostics for the four different models, which is used to evaluate the in-sample model performance. Model 3 has the lowest AIC, with and without increasing the flexibility of the trend during the Covid crisis. This indicates that Model 3 is the best fitting model to our data. Various tests on the standardized innovations are presented. Increasing the flexibility of the trends clearly leads to a general improvement when comparing the diagnostics in the table. No evidence can be found for the presence of heteroscedasticity anymore, while there is less evidence for violations of the normality assumption.

Table 5: Model diagnostics of LFS models with constant variance (left panel) and time-varying variances (right panel) for trend disturbance terms.

		constant trend variance				time-varying trend variance			
	Model	1	2	3	4	1	2	3	4
Series	Statistic								
	Log L	-135.173	-147.909	-126.531	-127.469	-118.166	-132.159	-110.772	-112.105
	AIC	3.920	4.279	3.677	3.703	3.441	3.835	3.233	3.271
y_t	Ljung Box	9.429	21.158	9.392	9.372	15.095	10.152	10.110	10.346
	H(h)	1.785	2.570*	2.211*	2.182*	0.589	1.636	0.751	0.811
	Skewness	-2.069**	-1.616**	-1.033**	-1.043**	0.485	0.391	0.631*	0.678*
	Kurtosis	14.795**	10.916**	10.370**	10.354**	3.825	5.619**	3.948	4.032
	Normality	462.184**	216.312**	173.312**	172.857**	4.798	22.100**	7.371*	8.586*
$x_{t,1}$	Ljung Box	18.935	17.592	20.862	20.202	23.666	27.310*	28.229*	27.564*
	H(h)	2.479*	2.919**	2.414*	2.413*	1.666	1.452	1.512	1.431
	Skewness	-0.762**	-0.454	-0.660*	-0.667*	-0.452	-0.199	-0.214	-0.131
	Kurtosis	5.664**	4.681**	5.007**	4.986**	3.755	3.333	3.153	3.207
	Normality	32.569**	12.623**	19.965**	19.797**	4.801	0.930	0.711	0.385
$x_{t,2}$	Ljung Box	12.577	10.971	11.201	11.402	16.246	8.448	14.596	10.651
	H(h)	1.867	2.554**	2.403*	2.356*	1.445	1.457	1.521	1.404
	Skewness	-0.581*	-0.162	-0.471	-0.506	-0.209	-0.049	-0.177	-0.239
	Kurtosis	4.250*	3.824	4.185*	4.231*	2.882	2.831	3.110	3.367
	Normality	10.063**	2.714	7.918*	8.775*	0.652	0.132	0.478	1.254
$x_{t,3}$	Ljung Box	60.199**	30.718*	32.893**	32.484**	65.889**	33.144**	36.945**	36.714**
	H(h)	1.591	2.699**	2.909**	2.827**	1.234	1.608	1.706	1.683
	Skewness	-0.219	0.171	0.084	0.053	-0.011	-0.088	0.207	0.210
	Kurtosis	3.839	4.077*	4.302*	4.294*	3.113	3.515	3.441	3.512
	Normality	3.103	4.417	5.960	5.827	0.046	1.025	1.266	1.515

$^{**}p < 0.01$, $^{*}p < 0.05$

A visual representation of the innovations from Model 3 for all four series is found in Figure 8, which clearly shows the large negative one-step ahead prediction errors when Covid started spreading. The largest drop in the labour force participation took place in April of 2020, but recovered quickly in the following months, which is also visible in the monthly LFS figures in Figure 7. This clarifies why the negative prediction in 2020Q2 is less extreme for the M2 and M3 series compared to the M1 series.

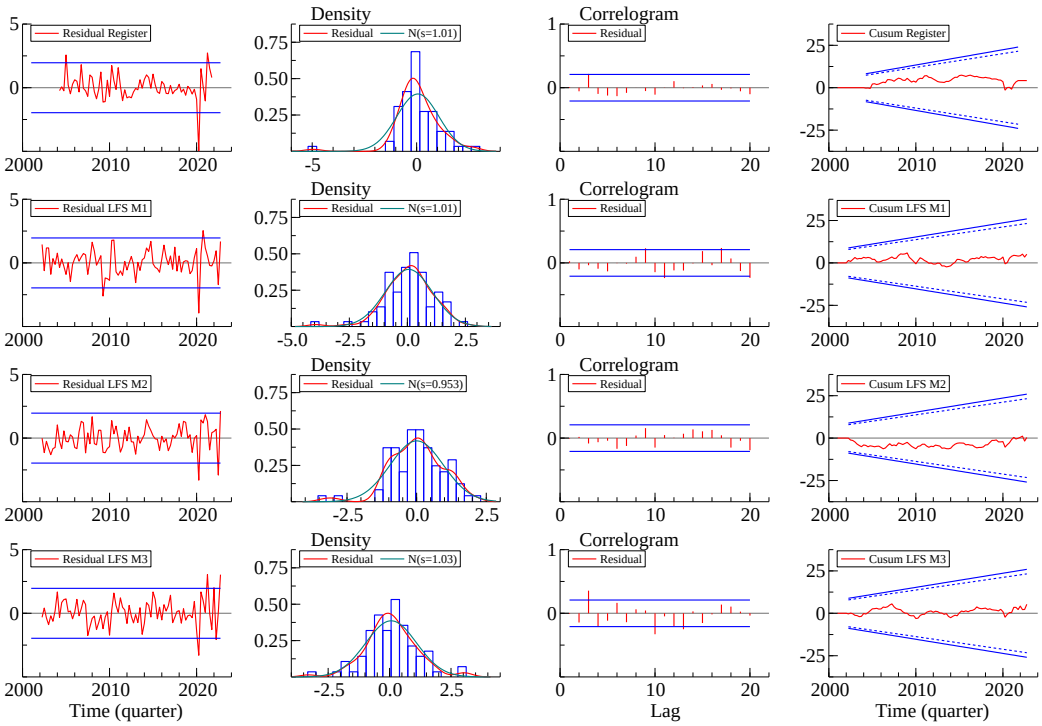

Figure 8: Innovations diagnostics of Model 3, obtained with Kalman smoother, with constant variances for trend disturbance terms.

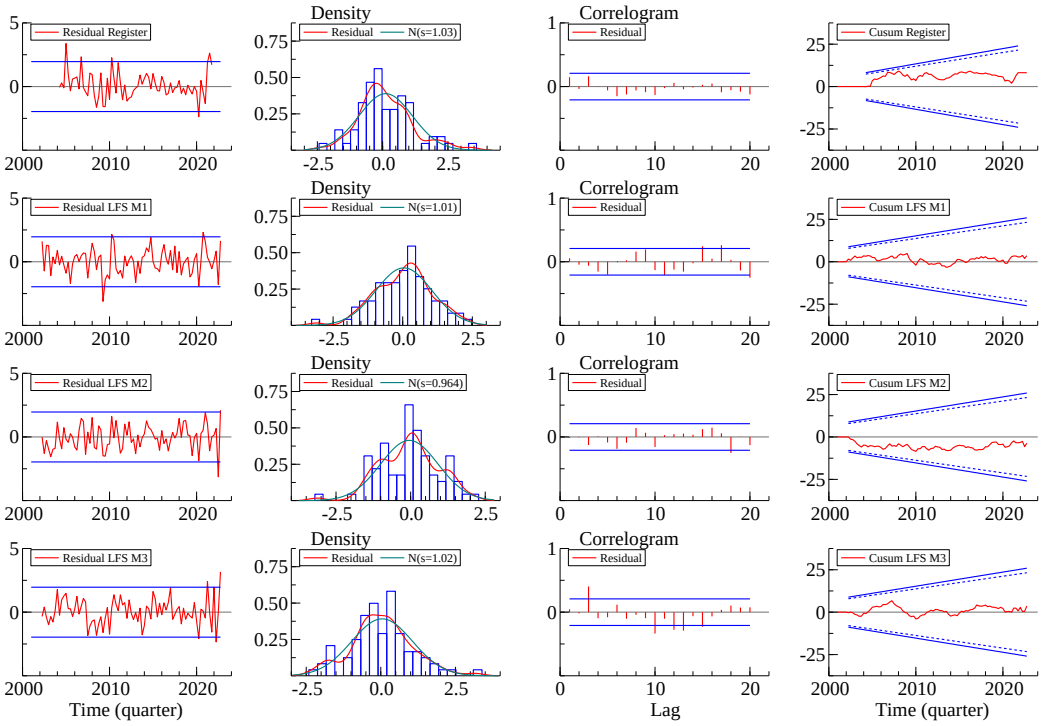

Figure 9: Innovations diagnostics of Model 3, obtained with Kalman smoother, with time-varying variances for trend disturbance terms.

Figure 9 shows the improvement in the innovation diagnostics when the flexibility of the trends is increased at the start of the Covid crisis. In general, these visual diagnostic checks show nothing alarming in terms of normality violations, autocorrelation, heteroscedasticity or drifting CUSUM. The diagnostics of Model 1 show a similar improvement in the prediction error, see Figures C1 and C2, however the CUSUM starts to drift outside the confidence bounds. This is not the case for Models 2 and 4. Their diagnostics resemble those of Model 3, see Figures C3 - C6.

The maximum likelihood procedure is based on all observations between 2001-2022. The resulting parameter estimates are presented in Table C1. Table C2 shows the factors from (13) used to increase the variances of the level and slope disturbance terms. They are responsible for the parameter estimate difference in Table C1. The factors are multiplied with the entire (rank deficient) level and slope covariance matrices, and Table C1 presents the resulting variance and loading parameters after decomposing these covariance matrices in (12). Similar to findings in van den Brakel and Michiels (2021), we observe a small variance for the irregular compared to the variance of the sampling error, indicating the latter captures most of this idiosyncratic variation. Model 2 without increasing the flexibility of the trend during the Covid period is estimated as a common smooth trend model, as the disturbance terms of the level and seasonal get variance parameters not far from zero. It seems that specifying a model with four independent seasonal components is redundant, because again most variance parameters are near zero. The hyperparameter estimates of Model 3 confirm the conclusion based on the AIC that a model with two common factors for both trend components and one common seasonal factor is the best fitting model to our data.

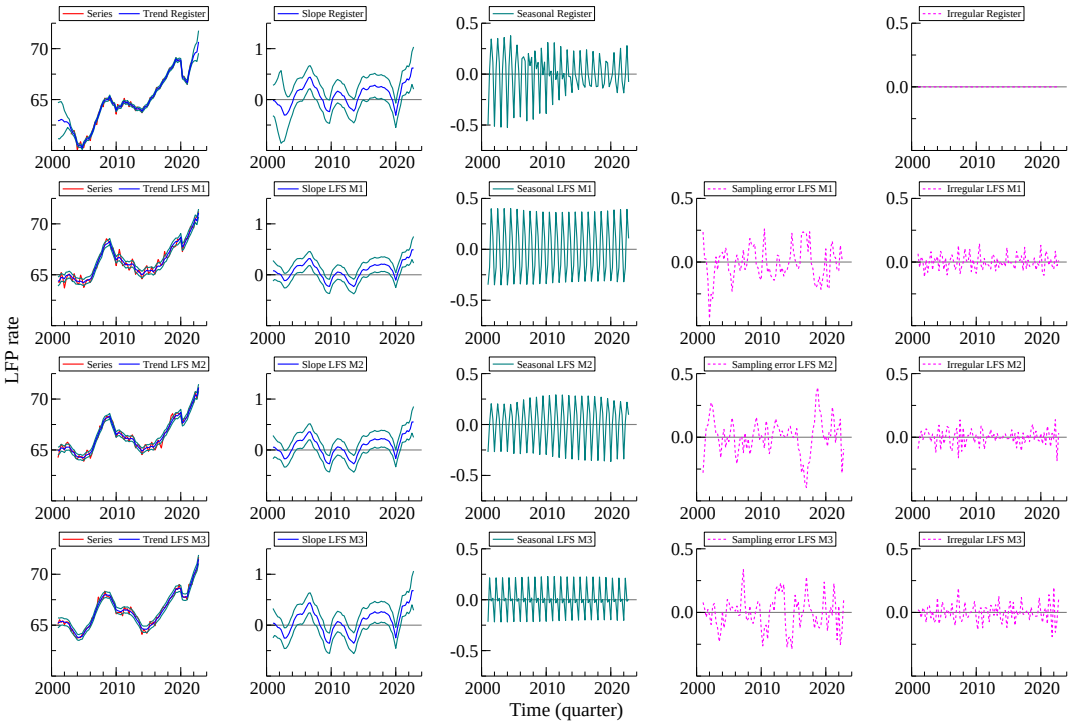

Figure 10: State vector estimates in Model 3, obtained with Kalman smoother.

Figure 10 shows the resulting smoothed estimates of the unobserved components when the estimated hyperparameters of Model 3 are used in the Kalman filter and smoother recursions. Figure 11 shows the same with the time-varying trend disturbance term variances and especially in the figures of the slope one can see the confidence bound becoming temporarily wider because of it. It also shows that the sampling error and irregular are closely related. For some models, the estimated irregular becomes very small, but the sampling error becomes larger and vice versa, see Figures C7 - C12. If either the irregular or sampling error becomes too small it is left out of the figure to avoid confusion. In general, the proposed multivariate models are able to identify the common unobserved components accurately.

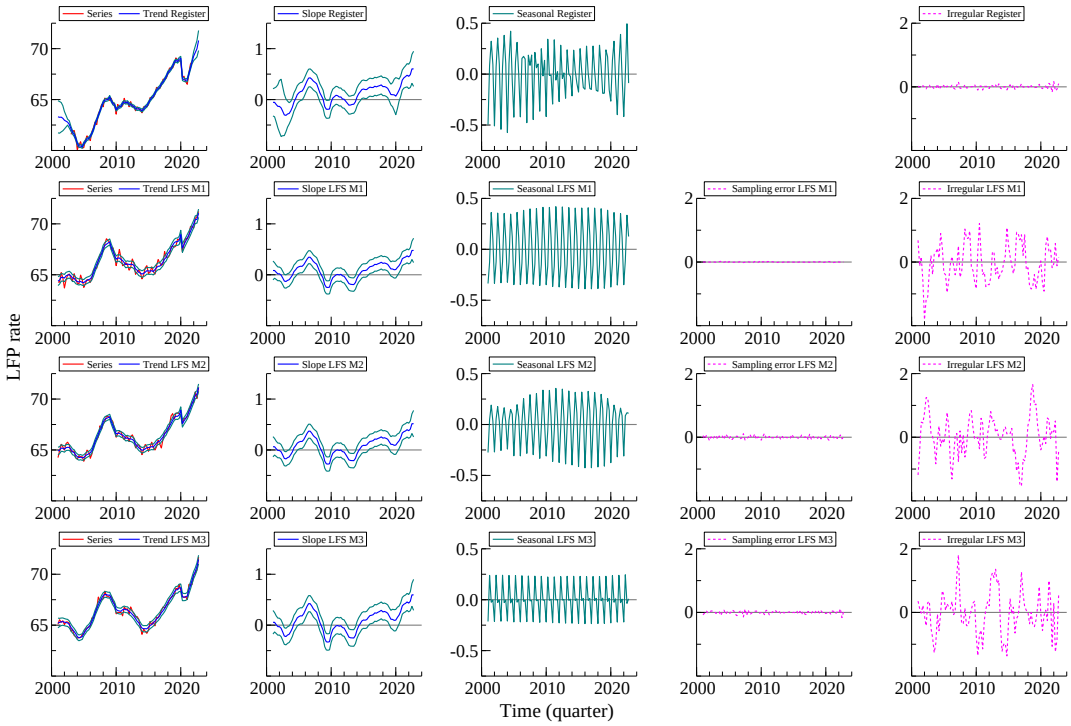

Figure 11: State vector estimates in Model 3 including time-varying variances for the trend disturbance terms at the start of the Covid outbreak.

4.5 Forecasting evidence

In this section the forecasting performance of our SUTSE models are tested for the quarterly LPF rate figures that are derived from the tax register (y_t). The quarter of interest is denoted with Q. This forecasting study identifies four different scenarios in which the Labour Force models are fitted. They are presented in Table 6.

In each scenario the 5-step ahead forecast for the tax register LFP rate in quarter Q is calculated and compared with the actual observation. Since the tax register series y_t becomes available at a quarterly frequency with a delay of about five quarters, the forecast horizon in this application is 5 quarters as shown in Table 6. Under scenario 1 a univariate *basic structural time series* model is applied. This scenario serves as a benchmark to assess to which extend the availability of the faster and more recent data from the LFS would

Table 6: Time series availability (✓) in the forecast scenarios.*

Scenario	Forecast moment	Series	Quarter					
			$Q-5$	$Q-4$	$Q-3$	$Q-2$	$Q-1$	Q
1	$Q(1)$	Register	✓					
2	$Q(2)$	Register	✓					
		LFS M1	✓	✓	✓	✓	✓	✓
		LFS M2	✓	✓	✓	✓	✓	
		LFS M3	✓	✓	✓	✓	✓	
3	$Q(3)$	Register	✓					
		LFS M1	✓	✓	✓	✓	✓	✓
		LFS M2	✓	✓	✓	✓	✓	✓
		LFS M3	✓	✓	✓	✓	✓	
4	$Q+1(1)$	Register	✓					
		LFS M1	✓	✓	✓	✓	✓	✓
		LFS M2	✓	✓	✓	✓	✓	✓
		LFS M3	✓	✓	✓	✓	✓	✓

*The quarter is Q and the number between brackets refers to the month in Q, 1,2,3.

improve the forecasting precision in the other three scenarios. Scenario 2, 3 and 4 include auxiliary information already available up until Q, which indicates that the forecasts can be thought of as nowcasts. These scenarios resemble the moments when the monthly LFS figures for the first, second and third month of that quarter become available and a more precise nowcast for y_t can be obtained. This is about two weeks after the end of the reference month. Recall that Ω_t contains the data available at period t and that $Y_t = (y_1, \ldots, y_t)'$. In scenario 1, $\Omega_t = (Y_{t-5})$, in scenario 2, $\Omega_t = (Y_{t-5}, X_{t,1}, X_{t-1,2}, X_{t-1,3})$, in scenario 3, $\Omega_t = (Y_{t-5}, X_{t,1}, X_{t,2}, X_{t-1,3})$ and in scenario 4, $\Omega_t = (Y_{t-5}, X_{t,1}, X_{t,2}, X_{t,3})$.

The forecasting study was applied to the quarters $Q = 2011(1), \ldots, 2021(4)$. Table 7 shows the MAPE for the separate years and the average over all years. Results are presented for all models with and without increasing the flexibility of the trend during the Covid crisis. Table C3 and Table C4 present the MSE and MAE. The hyperparameters are re-estimated by maximum likelihood in each quarter, to take into account the most recent observations while progressing through the forecast window. Comparing scenario 1 with the other scenarios in the performance measure tables show that the additional value of the auxiliary information from the LFS is small during the period 2014 until 2019 where the LFP rate has a stable increasing trend. The contribution of the auxiliary information is substantial in periods around a turning point, i.e. 2012 and 2013 and the years of the Covid crisis. This indicates that the nowcast of a detailed, precise figure derived from a register that comes with a long publication delay of five reference periods can be improved successfully with a timely indicator that is obtained from a sample survey.

Increasing the variances of the level and slope disturbance terms clearly increases the prediction errors in 2020 and 2021. It shows that, while this procedure did have a positive effect on the one-step ahead prediction errors as is shown in Section 4.4, it negatively influences the five-step ahead prediction errors. Table 7 shows that it had the smallest impact when model 3 was used for forecasting. The influence of these correction factors on the

Table 7: MAPE forecasting performance based on the 5-step ahead forecasts per forecasting year and in total, for models with time constant and time-varying variances for trend disturbance terms.*

Scenario	Model	2011	2012	2013	2014	2015	2016	2017	2018	2019	(2019)	2020	(2020)	2021	(2021)	Total	(Total)
1	BSTSM	2.207	1.819	0.722	0.748	1.463	0.560	0.339	0.405	0.405	0.405	3.371	3.371	5.704	6.523	1.613	1.687
2	1	2.410	1.089	0.417	1.319	1.093	0.807	0.332	0.289	0.512	0.487	2.168	3.280	1.889	5.494	1.120	1.547
	2	2.061	0.882	0.485	1.148	0.994	0.785	0.308	0.216	0.413	0.406	2.104	3.418	1.939	5.880	1.030	1.508
	3	1.457	0.982	0.293	1.268	1.085	0.924	0.422	0.416	0.641	0.618	2.110	3.363	1.765	2.812	1.033	1.240
	4	1.833	1.138	0.659	1.487	0.467	0.795	0.801	0.289	0.333	0.322	1.723	2.039	1.904	4.190	1.039	1.275
3	1	2.405	1.090	0.443	1.240	1.126	0.872	0.389	0.427	0.446	0.422	2.083	3.195	1.995	5.608	1.138	1.565
	2	2.107	0.885	0.468	1.072	1.021	0.847	0.387	0.352	0.350	0.343	2.044	3.310	2.010	5.978	1.049	1.525
	3	1.191	1.546	0.635	1.662	0.731	0.492	1.125	0.633	0.263	0.247	1.715	2.079	1.979	2.063	1.088	1.128
	4	1.850	1.835	0.930	1.602	0.442	0.427	0.974	0.546	0.268	0.253	1.657	1.988	2.042	4.316	1.143	1.378
4	1	2.360	0.877	0.172	0.815	0.994	0.786	0.509	0.585	0.552	0.519	1.975	3.200	1.984	5.297	1.055	1.465
	2	1.854	0.880	0.528	1.268	0.916	0.783	0.276	0.311	0.397	0.423	1.927	2.913	1.950	6.018	1.008	1.470
	3	1.494	1.099	0.476	1.345	0.583	0.591	0.966	0.627	0.583	0.566	1.722	2.205	1.926	5.037	1.038	1.363
	4	1.801	1.063	0.628	1.445	0.484	0.458	0.765	0.964	0.441	0.417	1.694	2.412	1.987	5.242	1.066	1.425

*Brackets refer to results including time-varying trend disturbance term variances.

uncertainty of the forecasts is visualized in Figure C16. This figure also confirms that the increase in the forecast standard errors surrounding the Covid crisis is of a lesser magnitude for model 3, which could explain the smaller increase in prediction errors for this model. For all models and scenarios it follows that increasing the flexibility of the trend component during the Covid period deteriorates the forecast performance.

Figure 12 compares the nowcasts of model 3 in the different scenarios. The forecasts in

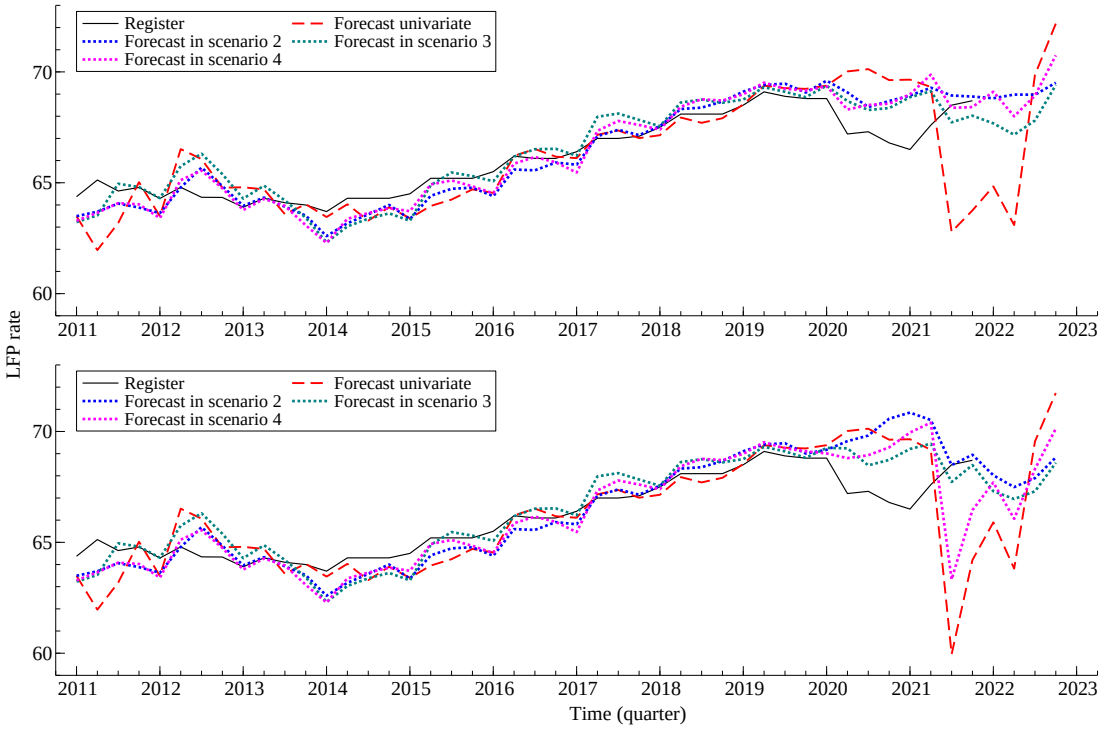

Figure 12: 5-step ahead forecasts of the register LPF rate series of Model 3, in the different forecast/nowcasting scenarios, constant (top) and time-varying (bottom) variances for trend disturbance terms.

the univariate scenario 1 seem to react late to changes in the trend, i.e. the lack of more timely information and the long publication lag that a univariate model has to deal with is clearly visible. When comparing scenario 4 with 2 or 3 in either Table 7 or Figure 12, it seems like waiting until the final monthly observation within the quarter of interest does not necessarily lead to a better forecast. This holds for all Models, see Figures C13 - C15. As expected the forecast uncertainty in Figure C16 is higher in scenario 1, but does not differ much in the other scenarios. Recall that the difference between them is the availability of just one extra monthly LFS observation when fitting the model and generating the forecasts, hence this is not that surprising.

As mentioned before, Figure 12 indicates that in the period between 2014 and 2020 with a stable upward trend in the LFP rate, all scenarios generate roughly equal nowcast errors. The advantage of a multivariate extension is clearly visible in times of changes in the trend. The figures also indicate that the models with a flexible trend component during the Covid period, result in larger nowcast errors. When considering their overall forecasting performance within a scenario in Table 7 instead, we already mentioned that model 3 outperforms the other models, which was mainly due to the fact that this model was less affected by the correction factors used to increase the variances of the level and slope disturbance terms. However, without this correction, model 2 is actually the best performing model. In general, model 1 is the worst performing model. Overall the largest difference is between the univariate model without auxiliary information versus the multivariate series that include auxiliary information. Differences between the three scenarios and multivariate models that include auxiliary information are smaller. When looking at the yearly disaggregation however, it

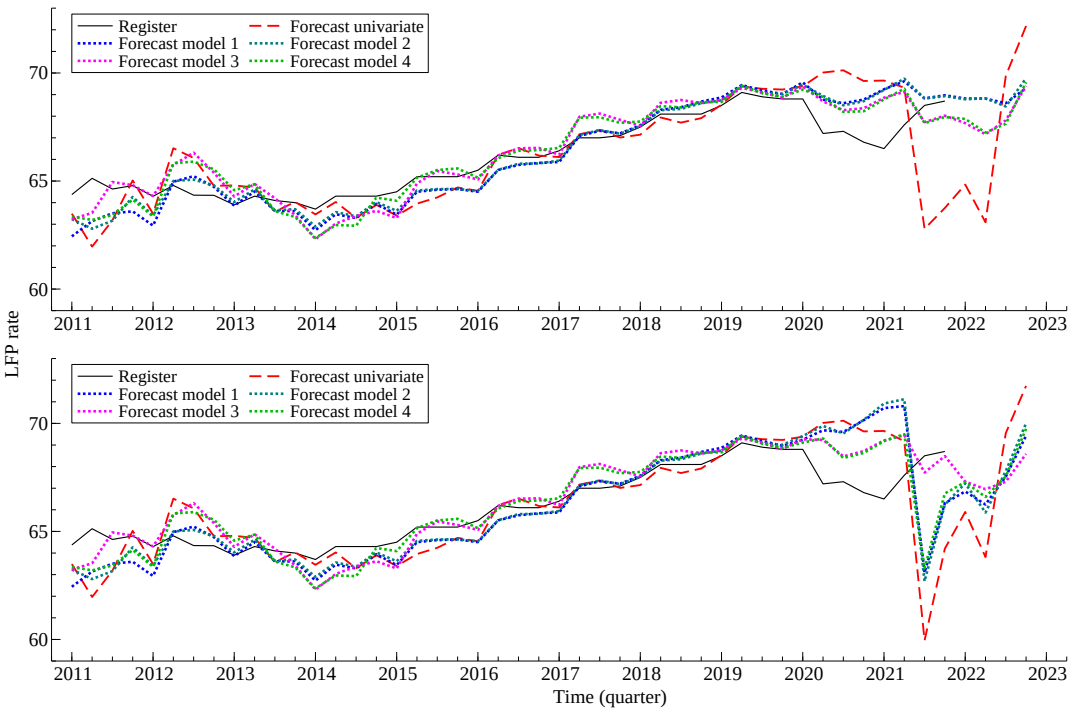

Figure 13: 5-step ahead forecasts of the register LPF rate series of the different models, in forecasting scenario 3, constant (top) and time-varying (bottom) variances for trend disturbance terms.

follows that this is again mainly due to years with changes in the trend. In stable years, this could be the other way around. Figure 13 shows the nowcasts of the different models in scenario 3 as example and it confirms this finding, but it holds for all multivariate scenarios, see Figures C17 and C18. Figure 14 shows the combined nowcasts and uncertainty of models 2 and 3 in scenario 3, since they can be considered the preferred models with constant or time-varying trend disturbance variances.

Figure 14: Forecasts with uncertainty of the Register LFP rate obtained with the best performing models 2 & 3 in scenario 3, with constant (top) and time-varying (bottom) variances for trend disturbance terms.

5 Conclusion

In this study we have developed a procedure for the early detection of the development of dynamic processes for which the corresponding observations exhibit long publication lags. The procedure exploit more timely information from auxiliary time series data. For this purpose, we have adopted a multivariate SUTSE model with common trends and common seasonal components, and have assumed that the different time series have similar underlying data generating processes. The proposed method is very general as it can be applied to a wide range of situations, in particular in official statistics, to construct timely early indicators. The recent Covid crisis emphasized the relevance of having timely indicators on a wide range of topics for policy and decision makers. In two empirical applications of key interest to the general public, we have validated the model assumptions and we have shown that the methods can handle different publication frequencies and can generate nowcasts and forecasts at different points throughout the year. A solution when mixed-frequencies exist is

proposed and it is shown how the methods perform in the case of unexpected circumstances such as the Covid outbreak.

In the first application, we have analyzed the development of the Dutch road safety process as observed in the official and final monthly number of road traffic fatalities in the Netherlands, issued once a year by Statistics Netherlands (SN). We proposed different variants of multivariate SUTSE models with common trends and seasonal components that can be treated as statistical early warning models, and that can detect potentially alarming developments in the official and final monthly number of road fatalities at an early stage. The models use three other proxies for the development of Dutch road safety, where two of these proxies contain observations that become available before the official and final statistics are issued by SN. The preliminary observations in these other sources clearly help to improve the forecasting precision of the official figures, compared to the forecasts obtained from univariate models. Particularly, a multivariate SUTSE model with one common level, one common slope and one common seasonal generates the best forecasts.

In the second application, we have shown how detailed indicators of the labour force participation (LFP) rate at the municipal level can be nowcasted using monthly estimates for the LFP rate at the national level. The detailed regional figures are derived from a tax register, which can be interpreted as a complete and accurate enumeration of quarterly LFP figures. These figures, however, come with a delay of five quarters which heavily compromise their relevance. The national figures on the other hand are obtained with the Labour Force Survey (LFS), which results in noisy LFP estimates that are available two weeks after a month has ended. Multivariate SUTSE models are proposed that combine the quarterly figures for the capital Amsterdam with monthly LFS figures. The model accounts for the sampling error of the LFS estimates and serial correlation induced by the rotating panel design of the LFS. The labour force figures are strongly affected by the Covid outbreak. The crisis induced turning points and strongly increased the period-to-period changes. To avoid model misspecification it was necessary to increase the flexibility of the trend by treating the variances of the level and slope disturbance terms as time-varying parameters. A forecasting study comparison has shown that this Covid modification deteriorates the multi-period forecast performance. Hence, while the one-step ahead prediction error/innovation diagnostics benefit from the increased flexibility, it does not lead to more precise five-step ahead predictions. It is found that the inclusion of more timely auxiliary information becomes beneficial for nowcasting the quarterly labour force rate, compared to predictions made by a univariate model, particularly for periods where the trend is not monotonically in- or decreasing. In general, a multivariate SUTSE model with one common level, one common slope and deterministic dummy seasonal component performs the best without the time-varying trend disturbance terms. With the increased trend flexibility, a model with two common levels, two common slopes and one common seasonal component achieves the best predictions.

Overall, this paper shows that the use of multivariate SUTSE models works well when there is a need for a preliminary estimate of a dynamic process for which the observations become available with potentially long publication lags. We have effectively developed a statistical tool that is able to provide an early warning on which policy-makers can act. The method is general, widely applicable and can deal with problems that arise in practice. More extensive higher dimensional applications could be explored further, as well as forecasting multiple target series simultaneously.

References

Blasques, F., S. J. Koopman, M. Mallee, and Z. Zhang (2016). Weighted maximum likelihood for dynamic factor analysis and forecasting with mixed frequency data. *Journal of Econometrics 193*(2), 405–417.

Carriero, A., T. E. Clark, and M. Marcellino (2015). Realtime nowcasting with a bayesian mixed frequency model with stochastic volatility. *Journal of the Royal Statistical Society Series A: Statistics in Society 178*(4), 837–862.

Commandeur, J. J. F. and S. J. Koopman (2007). *An Introduction to State Space Time Series Analysis*. Oxford: Oxford University Press.

Doornik, J. A. (2021). *Object-oriented matrix programming using Ox 9.0.* London: Timberlake Consultants Press.

Doz, C., D. Giannone, and L. Reichlin (2011). A two-step estimator for large approximate dynamic factor models based on kalman filtering. *Journal of Econometrics 164*(1), 188–205. Annals Issue on Forecasting.

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

Gonçalves, C., L. Hidalgo, D. B. N. Silva, and J. van den Brakel (2022). Single-month unemployment rate estimates for the Brazilian labour force survey using state-space models. *Journal of the Royal Statistical Society Series A (General) 185*(4), 1707–1732.

Grybauskas, A., V. Pilinkienė, M. Lukauskas, A. Stundžienė, and J. Bruneckienė (2023). Nowcasting unemployment using neural networks and multi-dimensional google trends data. *Economies 11*(5), 130.

Harvey, A. (2006). Chapter 7 forecasting with unobserved components time series models. Volume 1 of *Handbook of Economic Forecasting*, pp. 327–412. Elsevier.

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

Koopman, S. J., A. C. Harvey, J. A. Doornik, and N. Shephard (2009). *Stamp 8.2: Structural Time Series Analyser, Modeller and Predictor.* London: Timberlake Consultants.

Koopman, S. J., N. Shephard, and J. A. Doornik (2008). *Statistical Algorithms for Models in State Space Form: SsfPack 3.0.* London: Timberlake Consultants.

Kuzin, V., M. Marcellino, and C. Schumacher (2011). Midas vs. mixed-frequency var: Nowcasting gdp in the euro area. *International Journal of Forecasting 27*(2), 529–542.

Ljung, G. M. and G. E. P. Box (1978). On a measure of lack of fit in time series models. *Biometrika 65*(2), 297–303.

Page, E. S. (1954). Continuous inspection scheme. *Biometrika 41*, 100–115.

Pfeffermann, D. and R. Tiller (2005). Bootstrap approximation to prediction mse for state–space models with estimated parameters. *Journal of Time Series Analysis 26*(6), 893–916.

van den Brakel, J. and S. Krieg (2015). Dealing with small sample sizes, rotation group bias and discontinuities in a rotating panel design. *Survey Methodology 41*(2), 267–296.

van den Brakel, J. and J. Michiels (2021). Nowcasting register labour force participation rates in municipal districts using survey data. *Journal of Official Statistics 37*(4), 1009–1045.

van den Brakel, J., M. Souren, and S. Krieg (2022). Estimating monthly labour force figures during the covid-19 pandemic in the netherlands. *Journal of the Royal Statistical Society, Series A 185*(5), 1560–1583.

A Supplementary material Section 2

We start off from the state space approach for the analysis of time series data, see Harvey (1989), Durbin and Koopman (2012), and Commandeur and Koopman (2007). An important motivation for opting for this methodology is that it easily handles missing data in time series, and transparently generalizes to the analysis of multivariate time series. Let $\boldsymbol{y}_t$ denote the $N \times 1$ observation vector that contains the N observations at time t. The general linear Gaussian state space model for the T-dimensional observation sequence $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T$ is given by

$$\boldsymbol{y}_t = \boldsymbol{Z}_t \boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \text{NID}(\boldsymbol{0}, \boldsymbol{H}_t), \tag{A.1}$$

$$\boldsymbol{\alpha}_{t+1} = \boldsymbol{T}_t \boldsymbol{\alpha}_t + \boldsymbol{R}_t \boldsymbol{\eta}_t, \qquad \boldsymbol{\eta}_t \sim \text{NID}(\boldsymbol{0}, \boldsymbol{Q}_t), \qquad t = 1, \ldots, T, \tag{A.2}$$

where $\boldsymbol{\alpha}_t$ is the state vector, $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\eta}_t$ are disturbance vectors and the system matrices $\boldsymbol{Z}_t$, $\boldsymbol{T}_t$, $\boldsymbol{R}_t$, $\boldsymbol{H}_t$ and $\boldsymbol{Q}_t$ are fixed and known but a selection of elements may depend on an unknown parameter vector. Equation (A.1) is called the *observation* or *measurement equation*, while (A.2) is called the *state* or *transition equation*. The $M \times 1$ state vector $\boldsymbol{\alpha}_t$ is unobserved. The $N \times 1$ irregular vector $\boldsymbol{\varepsilon}_t$ has zero mean and $N \times N$ variance matrix $\boldsymbol{H}_t$.

The $N \times M$ matrix $\boldsymbol{Z}_t$ links the observation vector $\boldsymbol{y}_t$ with the unobservable state vector $\boldsymbol{\alpha}_t$. Besides the state variables that define the trend, seasonal and cycle, $\boldsymbol{\alpha}_t$ may also consist of regression variables. The $M \times M$ transition matrix $\boldsymbol{T}_t$ in (A.2) determines the dynamic evolution of the state vector. The $R \times 1$ disturbance vector $\boldsymbol{\eta}_t$ for the state vector update has zero mean and $R \times R$ variance matrix $\boldsymbol{Q}_t$. The observation and state disturbances $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\eta}_t$ are assumed to be serially independent and independent of each other at all time points. In many standard cases, $R = M$ and matrix $\boldsymbol{R}_t$ is the identity matrix I_M. In other cases, matrix $\boldsymbol{R}_t$ is an $M \times R$ selection matrix with $R < M$. Although matrix $\boldsymbol{R}_t$ can be specified freely, it is often composed of a selection from the first R columns of the identity matrix I_M.

By appropriate choices of the vectors $\boldsymbol{\alpha}_t$, $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\eta}_t$, and of the matrices $\boldsymbol{Z}_t$, $\boldsymbol{T}_t$, $\boldsymbol{H}_t$, $\boldsymbol{R}_t$ and $\boldsymbol{Q}_t$, a wide range of different time series models can be derived from (A.1) and (A.2). In this paper we discuss the class of *seemingly unrelated time series equations* (SUTSE) models, which can be considered a multivariate extension of structural time series models:

$$\boldsymbol{y}_t = \boldsymbol{\mu}_t + \boldsymbol{\gamma}_t + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \text{NID}(\boldsymbol{0}, \boldsymbol{\Sigma}_\varepsilon), \tag{A.3}$$

where $\boldsymbol{y}_t$ is a $N \times 1$ vector of time series, $\boldsymbol{\mu}_t$ is a $N \times 1$ vector of unobserved trends, $\boldsymbol{\gamma}_t$ is a $N \times 1$ vector of unobserved seasonal effects and $\boldsymbol{\varepsilon}_t$ is a $N \times 1$ vector of the irregular or noise component. The trend is defined as the *local linear trend* model, where the trend consists of a stochastic level and slope component,

$$\boldsymbol{\mu}_{t+1} = \boldsymbol{\mu}_t + \boldsymbol{\nu}_t + \boldsymbol{\xi}_t, \qquad \boldsymbol{\xi}_t \sim \text{NID}(\boldsymbol{0}, \boldsymbol{\Sigma}_\xi), \tag{A.4}$$

$$\boldsymbol{\nu}_{t+1} = \boldsymbol{\nu}_t + \boldsymbol{\zeta}_t, \qquad \boldsymbol{\zeta}_t \sim \text{NID}(\boldsymbol{0}, \boldsymbol{\Sigma}_\zeta). \tag{A.5}$$

More parsimonious trend models can be obtained straightforwardly by restricting the $N \times N$ covariance matrices $\boldsymbol{\Sigma}_\xi$ or $\boldsymbol{\Sigma}_\zeta$ to be equal to $\boldsymbol{0}$. The seasonal process can be modeled as a *dummy seasonal* model,

$$\boldsymbol{\gamma}_{t+1} = -\sum_{j=1}^{s-1} \boldsymbol{\gamma}_{t+1-j} + \boldsymbol{\omega}_t, \qquad \boldsymbol{\omega}_t \sim \text{NID}(\boldsymbol{0}, \boldsymbol{\Sigma}_\omega), \tag{A.6}$$

where s equals the number of months or quarters per year, depending on the frequency of the time series. Alternatively, a *trigonometric seasonal* model can be used,

$$\boldsymbol{\gamma}_t = \sum_{j=1}^{(s/2)} \boldsymbol{\gamma}_{j,t}, \tag{A.7}$$

where

$$\begin{aligned} \boldsymbol{\gamma}_{j,t+1} &= \cos\left(\frac{2\pi j}{s}\right)\boldsymbol{\gamma}_{j,t} + \sin\left(\frac{2\pi j}{s}\right)\boldsymbol{\gamma}^*_{j,t} + \boldsymbol{\omega}_{j,t}, \\ \boldsymbol{\gamma}^*_{j,t+1} &= -\sin\left(\frac{2\pi j}{s}\right)\boldsymbol{\gamma}_{j,t} + \cos\left(\frac{2\pi j}{s}\right)\boldsymbol{\gamma}^*_{j,t} + \boldsymbol{\omega}^*_{j,t}, \qquad j = 1, \ldots, (s/2), \end{aligned} \tag{A.8}$$

and where the trigonometric functions are scalars, $\boldsymbol{\gamma}_{j,t}$ and $\boldsymbol{\gamma}^*_{j,t}$ are $N \times 1$ vectors and $\boldsymbol{\omega}_{j,t}$ and $\boldsymbol{\omega}^*_{j,t}$ are independent NID$(\mathbf{0}, \boldsymbol{\Sigma}_\omega)$ variables. The disturbance terms of the different components in (A.3), (A.4), (A.5) and (A.6) or (A.8) are mutually independent. However, the model allows for cross-sectional correlations in the disturbances within these components. In other words, $\boldsymbol{\Sigma}_\xi$, $\boldsymbol{\Sigma}_\zeta$ and $\boldsymbol{\Sigma}_\omega$ can be diagonal, full rank or rank deficient (i.e., a matrix with rank less than full).

In case of rank deficient covariance matrices, the SUTSE model from (A.3) can be written as a common factor model, in which some or all of the components are driven by disturbance vectors with less than N elements:

$$\boldsymbol{y}_t = \boldsymbol{\Theta}_\mu \boldsymbol{\mu}^\dagger_t + \boldsymbol{\mu}_\theta + \boldsymbol{\Theta}_\nu \boldsymbol{\mu}^{\dagger\dagger}_t + \boldsymbol{\mu}_{\theta,t} + \boldsymbol{\Theta}_\gamma \boldsymbol{\gamma}^\dagger_t + \boldsymbol{\gamma}_{\theta,t} + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \text{NID}(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon), \tag{A.9}$$

$$\boldsymbol{\mu}^\dagger_{t+1} = \boldsymbol{\mu}^\dagger_t + \boldsymbol{\xi}^\dagger_t, \qquad \boldsymbol{\xi}^\dagger_t \sim \text{NID}(\mathbf{0}, \boldsymbol{\Sigma}^\dagger_\xi), \tag{A.10}$$

$$\boldsymbol{\mu}^{\dagger\dagger}_{t+1} = \boldsymbol{\mu}^{\dagger\dagger}_t + \boldsymbol{\nu}^{\dagger\dagger}_t, \tag{A.11}$$

$$\boldsymbol{\nu}^{\dagger\dagger}_{t+1} = \boldsymbol{\nu}^{\dagger\dagger}_t + \boldsymbol{\zeta}^{\dagger\dagger}_t, \qquad \boldsymbol{\zeta}^{\dagger\dagger}_t \sim \text{NID}(\mathbf{0}, \boldsymbol{\Sigma}^{\dagger\dagger}_\zeta), \tag{A.12}$$

where $\boldsymbol{\mu}^\dagger_t$ is a $K_\mu \times 1$ vector of common levels, modelled as in (A.4) with $\boldsymbol{\nu}^\dagger_t$ equal to $\mathbf{0}$ for all t and diagonal $\boldsymbol{\Sigma}^\dagger_\xi$, $\boldsymbol{\mu}^{\dagger\dagger}_t$ is a $K_\nu \times 1$ vector of common slopes, modelled as in (A.4) and (A.5) without a vector with level disturbance terms and diagonal $\boldsymbol{\Sigma}^{\dagger\dagger}_\zeta$ and $\boldsymbol{\gamma}^\dagger_t$ is a $K_\gamma \times 1$ vector of common seasonals, modelled as in (A.6) or (A.7) and (A.8) and diagonal $\boldsymbol{\Sigma}^\dagger_\omega$. The $N \times 1$ vector $\boldsymbol{\mu}_\theta$ consists of K_μ zeros as first elements followed by a $N - K_\mu$ vector $\bar{\boldsymbol{\mu}}$ of fixed levels as remainder. The $N \times 1$ vector $\boldsymbol{\mu}_{\theta,t}$ consists of K_ν zeros as first elements followed by a $N - K_\nu$ vector $\bar{\boldsymbol{\mu}} + \bar{\boldsymbol{\nu}}t$ of fixed linear trends as remainder. The $N \times 1$ vector $\gamma_{\boldsymbol{\theta},t}$ consists of K_γ zeros as first elements followed by a $N - K_\gamma$ vector of fixed seasonal effects as remainder, i.e. $\bar{\gamma}_t = \bar{\gamma}_j$ if t corresponds to seasonal period j, for $j = 1, \ldots, s$. The $N \times K_\mu$ factor loading matrix $\boldsymbol{\Theta}_\mu$ has the ij-th element θ_{ij} equal to zero for $j > i$ and θ_{ii} equal to one. Similar structures apply to $\boldsymbol{\Theta}_\nu$ and $\boldsymbol{\Theta}_\gamma$. Note that we can obtain the original SUTSE notation (A.3) by writing $\boldsymbol{\mu}_t = \boldsymbol{\Theta}_\mu \boldsymbol{\mu}^\dagger_t + \boldsymbol{\mu}_\theta + \boldsymbol{\Theta}_\nu \boldsymbol{\mu}^{\dagger\dagger}_t + \boldsymbol{\mu}_{\theta,t}$ and $\boldsymbol{\gamma}_t = \boldsymbol{\Theta}_\gamma \boldsymbol{\gamma}^\dagger_t + \boldsymbol{\gamma}_{\theta,t}$, while $\boldsymbol{\Sigma}_\xi = \boldsymbol{\Theta}_\mu \boldsymbol{\Sigma}^\dagger_\xi \boldsymbol{\Theta}'_\mu$, $\boldsymbol{\Sigma}_\zeta = \boldsymbol{\Theta}_\nu \boldsymbol{\Sigma}^{\dagger\dagger}_\zeta \boldsymbol{\Theta}'_\nu$ and $\boldsymbol{\Sigma}_\gamma = \boldsymbol{\Theta}_\gamma \boldsymbol{\Sigma}^\dagger_\omega \boldsymbol{\Theta}'_\gamma$ are singular matrices with rank K_μ, K_ν and K_γ respectively (Koopman, Harvey, Doornik, and Shephard 2009).

B Supplementary material Section 3

B.1 State space form

The Road Safety model can be written in linear state space form from (A.1)-(A.2). The measurement equation is:

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \\ y_{4,t} \end{pmatrix} = \left[\begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} \otimes \boldsymbol{I}_4 \right] \boldsymbol{\alpha}_t + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \\ \varepsilon_{4,t} \end{pmatrix}, \tag{B.1}$$

where the state vector $\boldsymbol{\alpha}_t$ is specified as follows:

$$\boldsymbol{\alpha}_t = \begin{pmatrix} \boldsymbol{\mu}_t' & \boldsymbol{\nu}_t' & \boldsymbol{\gamma}_t' \end{pmatrix}', \quad \boldsymbol{\mu}_t = \begin{pmatrix} \mu_{1,t} & \mu_{2,t} & \mu_{3,t} & \mu_{4,t} \end{pmatrix}', \quad \boldsymbol{\nu}_t = \begin{pmatrix} \nu_{1,t} & \nu_{2,t} & \nu_{3,t} & \nu_{4,t} \end{pmatrix}',$$

$$\boldsymbol{\gamma}_t = \begin{pmatrix} \boldsymbol{\gamma}_{1,t}' & \boldsymbol{\gamma}_{1,t}^{*\prime} & \boldsymbol{\gamma}_{2,t}' & \dots & \boldsymbol{\gamma}_{5,t}' & \boldsymbol{\gamma}_{5,t}^{*\prime} & \boldsymbol{\gamma}_{6,t}' \end{pmatrix}',$$

$$\boldsymbol{\gamma}_{j,t} = \begin{pmatrix} \gamma_{1,t} & \gamma_{2,t} & \gamma_{3,t} & \gamma_{4,t} \end{pmatrix}' \quad \forall\, j \in \{1, \dots, 6\},$$

$$\boldsymbol{\gamma}_{j,t}^{*} = \begin{pmatrix} \gamma_{1,t}^{*} & \gamma_{2,t}^{*} & \gamma_{3,t}^{*} & \gamma_{4,t}^{*} \end{pmatrix}' \quad \forall\, j \in \{1, \dots, 5\},$$

which contains the trend and trigonometric seasonal components for the four monthly time series. As discussed the idiosyncratic noise of the target and auxiliary series are Gaussian i.i.d. sequences with a full irregular covariance matrix:

$$\boldsymbol{\Sigma}_\varepsilon = \begin{bmatrix} \sigma_{\varepsilon_1}^2 & \sigma_{\varepsilon_{1,2}} & \sigma_{\varepsilon_{1,3}} & \sigma_{\varepsilon_{1,4}} \\ \sigma_{\varepsilon_{2,1}} & \sigma_{\varepsilon_2}^2 & \sigma_{\varepsilon_{2,3}} & \sigma_{\varepsilon_{2,4}} \\ \sigma_{\varepsilon_{3,1}} & \sigma_{\varepsilon_{3,2}} & \sigma_{\varepsilon_3}^2 & \sigma_{\varepsilon_{3,4}} \\ \sigma_{\varepsilon_{4,1}} & \sigma_{\varepsilon_{4,2}} & \sigma_{\varepsilon_{4,3}} & \sigma_{\varepsilon_4}^2 \end{bmatrix}. \tag{B.2}$$

The state vector's dynamic transitioning can be described as follows:

$$\boldsymbol{\alpha}_{t+1} = \begin{bmatrix} \boldsymbol{T}_\mu & \boldsymbol{0}_{8\times 44} \\ \boldsymbol{0}_{44\times 8} & \boldsymbol{T}_\gamma \end{bmatrix} \boldsymbol{\alpha}_t + \boldsymbol{\eta}_t, \qquad \boldsymbol{\eta}_t \sim \mathbf{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_\eta). \tag{B.3}$$

The trend is defined as local linear trend model and the seasonal component as stochastic trigonometric seasonal model, see Harvey (1989) and Durbin and Koopman (2012), by specifying:

$$\boldsymbol{T}_\mu = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \otimes \boldsymbol{I}_4, \quad \boldsymbol{T}_\gamma = \operatorname{diag}(\boldsymbol{C}_1, \boldsymbol{C}_2, \boldsymbol{C}_3, \boldsymbol{C}_4, \boldsymbol{C}_5, -\boldsymbol{I}_4),$$

$$\boldsymbol{C}_j = \begin{bmatrix} \cos\lambda_j & \sin\lambda_j \\ -\sin\lambda_j & \cos\lambda_j \end{bmatrix} \otimes \boldsymbol{I}_4, \quad \lambda_j = \frac{\pi j}{6}, \quad \forall\, j \in \{1, \dots, 5\},$$

$$\boldsymbol{\eta}_t = \begin{pmatrix} \boldsymbol{\xi}_t' & \boldsymbol{\zeta}_t' & \boldsymbol{\omega}_{1,t}' & \boldsymbol{\omega}_{1,t}^{*\prime} & \boldsymbol{\omega}_{2,t}' & \dots & \boldsymbol{\omega}_{5,t}' & \boldsymbol{\omega}_{5,t}^{*\prime} & \boldsymbol{\omega}_{6,t}' \end{pmatrix}',$$

$$\boldsymbol{\xi}_t = \begin{pmatrix} \xi_{1,t} & \xi_{2,t} & \xi_{3,t} & \xi_{4,t} \end{pmatrix}', \quad \boldsymbol{\zeta}_t = \begin{pmatrix} \zeta_{1,t} & \zeta_{2,t} & \zeta_{3,t} & \zeta_{4,t} \end{pmatrix}',$$

$$\boldsymbol{\omega}_{j,t} = \begin{pmatrix} \omega_{1,t} & \omega_{2,t} & \omega_{3,t} & \omega_{4,t} \end{pmatrix}', \quad \forall\, j \in \{1, \dots, 6\},$$

$$\boldsymbol{\omega}_{j,t}^{*} = \begin{pmatrix} \omega_{1,t}^{*} & \omega_{2,t}^{*} & \omega_{3,t}^{*} & \omega_{4,t}^{*} \end{pmatrix}', \quad \forall\, j \in \{1, \dots, 5\}.$$

The covariance matrix of the disturbance terms is a blockdiagonal matrix of 4×4 submatrices:

$$\boldsymbol{\Sigma}_\eta = \text{diag}(\boldsymbol{\Sigma}_\xi, \boldsymbol{\Sigma}_\zeta, \boldsymbol{I}_{11} \otimes \boldsymbol{\Sigma}_\omega), \tag{B.4}$$

where these disturbance term variance blocks are rank deficient in case of common factors for the corresponding component. Loadings are then retrieved from the lower triangular matrix of a Cholesky decomposition of such a block.

B.2 Tables

Table B1: Parameter estimates.

Model	1	2	3	4	5	6
$\hat{\sigma}^2_{\xi}$	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005
$\hat{\theta}_{\mu,2}$	1	1	1	1.2089	1.2747	1.2328
$\hat{\theta}_{\mu,3}$	1	1	1	0.9900	0.9861	1.0910
$\hat{\theta}_{\mu,4}$	1	1	1	0.9466	0.9358	0.9253
$\hat{\sigma}^2_{\zeta}$	-	2.574e-14	9.896e-13	-	1.754e-08	1.586e-12
$\hat{\theta}_{\nu,2}$	-	1	-180.2032	-	1	90.7448
$\hat{\theta}_{\nu,3}$	-	1	57.4598	-	1	-151.8539
$\hat{\theta}_{\nu,4}$	-	1	74.5109	-	1	-10.3725
$\hat{\sigma}^2_{\omega}$	-	2.519e-06	0	-	2.570e-06	2.822e-06
$\hat{\theta}_{\gamma,2}$	-	1	94.9402	-	1	0.9037
$\hat{\theta}_{\gamma,3}$	-	1	101.0762	-	1	0.9014
$\hat{\theta}_{\gamma,4}$	-	1	-38.0955	-	1	0.9857
$\hat{\sigma}^2_{\varepsilon_1}$	0.0199	0.0187	0.0198	0.0199	0.0189	0.0188
$\hat{\sigma}^2_{\varepsilon_2}$	0.0239	0.0227	0.0233	0.0226	0.0216	0.0215
$\hat{\sigma}^2_{\varepsilon_3}$	0.0220	0.0209	0.0220	0.0220	0.0211	0.0205
$\hat{\sigma}^2_{\varepsilon_4}$	0.0216	0.0204	0.0215	0.0217	0.0209	0.0206
$\hat{\rho}_{\varepsilon_{2,1}}$	0.9282	0.9244	0.9364	0.9349	0.9335	0.9345
$\hat{\rho}_{\varepsilon_{3,1}}$	0.8651	0.8577	0.8660	0.8651	0.8592	0.8617
$\hat{\rho}_{\varepsilon_{4,1}}$	0.9020	0.8963	0.9038	0.9038	0.8994	0.8991
$\hat{\rho}_{\varepsilon_{3,2}}$	0.8145	0.8050	0.8294	0.8226	0.8166	0.8190
$\hat{\rho}_{\varepsilon_{4,2}}$	0.8377	0.8292	0.8552	0.8512	0.8464	0.8479
$\hat{\rho}_{\varepsilon_{4,3}}$	0.9507	0.9480	0.9507	0.9510	0.9490	0.9509

Table B2: MSE forecasting performance based on the 1- to 12-step ahead forecasts per forecasting year and in total.

Scenario	Model	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Total
1	DBSTSM	0.040	0.024	0.056	0.022	0.033	0.044	0.019	0.049	0.013	0.021	0.046	0.076	0.037
	SBSTSM	0.041	0.024	0.056	0.022	0.033	0.044	0.019	0.049	0.013	0.021	0.046	0.076	0.037
2	1	0.038	0.022	0.066	0.023	0.028	0.041	0.017	0.049	0.014	0.020	0.050	0.074	0.037
	2	0.041	0.022	0.060	0.022	0.032	0.045	0.019	0.049	0.014	0.021	0.048	0.072	0.037
	3	0.040	0.021	0.054	0.024	0.030	0.047	0.025	0.050	0.015	0.020	0.048	0.072	0.037
	4	0.047	0.024	0.058	0.023	0.033	0.044	0.018	0.045	0.011	0.023	0.049	0.081	0.038
	5	0.041	0.025	0.057	0.022	0.032	0.048	0.019	0.038	0.010	0.033	0.057	0.064	0.037
	6	0.046	0.037	0.050	0.024	0.036	0.052	0.032	0.055	0.024	0.018	0.047	0.102	0.044
3	1	0.028	0.019	0.016	0.008	0.021	0.021	0.016	0.035	0.017	0.025	0.027	0.018	0.021
	2	0.031	0.020	0.014	0.011	0.020	0.021	0.017	0.033	0.016	0.027	0.027	0.017	0.021
	3	0.029	0.017	0.013	0.011	0.021	0.020	0.020	0.029	0.014	0.029	0.029	0.023	0.021
	4	0.029	0.016	0.016	0.006	0.026	0.023	0.018	0.032	0.015	0.029	0.030	0.019	0.022
	5	0.030	0.016	0.015	0.007	0.023	0.021	0.018	0.032	0.017	0.030	0.025	0.027	0.022
	6	0.035	0.021	0.013	0.010	0.025	0.019	0.019	0.024	0.014	0.040	0.031	0.032	0.024
4	1	0.021	0.009	0.011	0.008	0.010	0.009	0.010	0.012	0.016	0.024	0.021	0.007	0.013
	2	0.024	0.009	0.010	0.008	0.010	0.009	0.011	0.012	0.016	0.026	0.022	0.007	0.014
	3	0.025	0.008	0.009	0.009	0.009	0.008	0.014	0.011	0.015	0.028	0.024	0.008	0.014
	4	0.021	0.008	0.011	0.008	0.014	0.010	0.012	0.012	0.015	0.028	0.021	0.007	0.014
	5	0.026	0.008	0.010	0.008	0.011	0.009	0.013	0.014	0.017	0.030	0.022	0.004	0.014
	6	0.026	0.009	0.012	0.009	0.019	0.017	0.024	0.032	0.036	0.019	0.023	0.025	0.021
5	1	0.003	0.004	0.002	0.003	0.002	0.004	0.002	0.002	0.004	0.002	0.006	0.001	0.003
	2	0.003	0.004	0.002	0.003	0.001	0.004	0.002	0.002	0.004	0.002	0.006	0.001	0.003
	3	0.003	0.007	0.002	0.004	0.004	0.006	0.006	0.006	0.008	0.002	0.009	0.010	0.006
	4	0.001	0.004	0.002	0.003	0.003	0.006	0.004	0.005	0.007	0.007	0.005	0.011	0.005
	5	0.001	0.004	0.002	0.003	0.002	0.004	0.002	0.002	0.003	0.002	0.006	0.001	0.003
	6	0.003	0.005	0.001	0.003	0.001	0.004	0.003	0.002	0.003	0.003	0.008	0.001	0.003

Table B3: MSE forecasting performance based on the 1- to 12-step ahead forecasts per month.

Scenario	Model	t+1	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec
1	DBSTSM	0.050	0.051	0.029	0.034	0.014	0.027	0.054	0.034	0.036	0.056	0.018	0.037
	SBSTSM	0.050	0.051	0.029	0.034	0.014	0.027	0.054	0.034	0.036	0.056	0.018	0.038
2	1	0.060	0.051	0.026	0.032	0.015	0.027	0.052	0.032	0.038	0.054	0.019	0.033
	2	0.056	0.054	0.026	0.034	0.015	0.028	0.052	0.033	0.037	0.055	0.019	0.036
	3	0.052	0.054	0.029	0.034	0.015	0.028	0.054	0.033	0.037	0.053	0.021	0.036
	4	0.053	0.049	0.030	0.034	0.016	0.026	0.053	0.037	0.041	0.056	0.019	0.042
	5	0.061	0.060	0.028	0.035	0.013	0.022	0.047	0.034	0.033	0.061	0.015	0.038
	6	0.048	0.055	0.045	0.040	0.018	0.034	0.065	0.050	0.042	0.059	0.019	0.051
3	1	0.007	0.012	0.007	0.004	0.003	0.014	0.044	0.023	0.033	0.051	0.020	0.033
	2	0.006	0.012	0.007	0.004	0.003	0.014	0.043	0.023	0.032	0.051	0.021	0.036
	3	0.006	0.013	0.007	0.005	0.003	0.013	0.044	0.028	0.030	0.052	0.022	0.034
	4	0.005	0.014	0.007	0.006	0.004	0.012	0.044	0.026	0.032	0.052	0.019	0.038
	5	0.006	0.014	0.008	0.005	0.004	0.017	0.042	0.029	0.031	0.056	0.017	0.035
	6	0.005	0.014	0.008	0.008	0.003	0.014	0.042	0.037	0.030	0.056	0.021	0.043
4	1	0.007	0.012	0.006	0.004	0.003	0.014	0.008	0.007	0.004	0.043	0.022	0.029
	2	0.006	0.012	0.006	0.004	0.003	0.014	0.008	0.007	0.004	0.044	0.022	0.033
	3	0.006	0.014	0.007	0.004	0.003	0.013	0.009	0.010	0.003	0.044	0.024	0.032
	4	0.006	0.014	0.007	0.006	0.004	0.012	0.008	0.007	0.004	0.045	0.021	0.033
	5	0.005	0.014	0.008	0.005	0.004	0.016	0.008	0.008	0.004	0.052	0.017	0.031
	6	0.007	0.019	0.015	0.010	0.007	0.028	0.016	0.016	0.014	0.055	0.023	0.041
5	1	0.003	0.002	0.003	0.005	0.004	0.005	0.003	0.003	0.001	0.002	0.004	0.002
	2	0.002	0.002	0.003	0.005	0.004	0.005	0.003	0.003	0.001	0.002	0.004	0.002
	3	0.005	0.003	0.005	0.007	0.005	0.009	0.005	0.008	0.002	0.004	0.009	0.005
	4	0.002	0.004	0.004	0.005	0.007	0.009	0.004	0.005	0.004	0.004	0.006	0.004
	5	0.002	0.003	0.003	0.004	0.004	0.005	0.003	0.003	0.001	0.001	0.003	0.001
	6	0.002	0.003	0.003	0.004	0.004	0.006	0.003	0.003	0.001	0.002	0.004	0.002

Table B4: MAE forecasting performance based on the 1- to 12-step ahead forecasts per forecasting year and in total.

Scenario	Model	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Total
1	DBSTSM	0.141	0.125	0.186	0.110	0.153	0.150	0.122	0.196	0.095	0.090	0.163	0.234	0.147
	SBSTSM	0.140	0.125	0.186	0.110	0.153	0.150	0.122	0.196	0.095	0.090	0.163	0.234	0.147
2	1	0.144	0.126	0.201	0.120	0.140	0.142	0.114	0.193	0.100	0.090	0.167	0.227	0.147
	2	0.143	0.124	0.194	0.111	0.154	0.148	0.125	0.195	0.099	0.092	0.166	0.228	0.148
	3	0.147	0.122	0.183	0.115	0.150	0.152	0.144	0.195	0.103	0.089	0.171	0.229	0.150
	4	0.149	0.124	0.187	0.114	0.157	0.147	0.116	0.185	0.092	0.096	0.169	0.243	0.148
	5	0.143	0.128	0.188	0.110	0.153	0.152	0.120	0.171	0.085	0.118	0.186	0.210	0.147
	6	0.144	0.153	0.179	0.110	0.164	0.169	0.154	0.205	0.133	0.101	0.164	0.283	0.163
3	1	0.096	0.117	0.091	0.078	0.105	0.110	0.107	0.143	0.111	0.110	0.134	0.091	0.108
	2	0.099	0.118	0.083	0.090	0.105	0.105	0.109	0.141	0.108	0.115	0.133	0.086	0.108
	3	0.107	0.109	0.077	0.089	0.111	0.102	0.114	0.129	0.101	0.128	0.138	0.097	0.108
	4	0.097	0.109	0.086	0.073	0.124	0.116	0.112	0.138	0.104	0.122	0.137	0.091	0.109
	5	0.103	0.106	0.084	0.078	0.116	0.109	0.109	0.141	0.110	0.117	0.133	0.104	0.109
	6	0.102	0.119	0.079	0.087	0.128	0.098	0.111	0.123	0.096	0.151	0.143	0.112	0.112
4	1	0.094	0.081	0.073	0.080	0.063	0.077	0.086	0.093	0.108	0.106	0.123	0.056	0.087
	2	0.096	0.081	0.068	0.082	0.061	0.071	0.087	0.092	0.105	0.110	0.124	0.057	0.086
	3	0.105	0.073	0.062	0.084	0.058	0.066	0.097	0.085	0.104	0.124	0.127	0.060	0.087
	4	0.092	0.068	0.066	0.075	0.076	0.085	0.091	0.090	0.105	0.121	0.124	0.058	0.088
	5	0.097	0.074	0.066	0.080	0.065	0.078	0.091	0.100	0.107	0.110	0.123	0.045	0.086
	6	0.107	0.070	0.067	0.078	0.110	0.110	0.138	0.151	0.150	0.108	0.126	0.139	0.113
5	1	0.042	0.051	0.032	0.049	0.027	0.056	0.037	0.038	0.047	0.035	0.067	0.027	0.042
	2	0.043	0.051	0.031	0.049	0.027	0.055	0.037	0.038	0.047	0.035	0.067	0.026	0.042
	3	0.040	0.066	0.034	0.056	0.051	0.065	0.066	0.062	0.070	0.042	0.083	0.083	0.060
	4	0.028	0.048	0.032	0.048	0.041	0.065	0.050	0.057	0.073	0.068	0.066	0.096	0.056
	5	0.029	0.054	0.030	0.049	0.032	0.059	0.039	0.037	0.046	0.032	0.067	0.025	0.042
	6	0.049	0.057	0.029	0.046	0.029	0.054	0.042	0.040	0.048	0.041	0.072	0.033	0.045

Table B5: Real-time hyperparameter estimates of model 4 in forecast scenario 3.

	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
$\hat{\sigma}^2_{\xi}$	0.0002	0.0002	0.0002	0.0001	0.0001	0.0002	0.0002	0.0002	0.0003	0.0003	0.0003	0.0005
$\hat{\theta}_{\mu,2}$	1.2646	1.3993	1.4644	1.4910	1.4967	1.3865	1.2426	1.2558	1.2419	1.2379	1.2159	1.2073
$\hat{\theta}_{\mu,3}$	1.9184	2.1501	1.6638	1.6446	1.5516	1.2046	1.1237	0.9525	0.9069	0.8985	0.9137	0.9676
$\hat{\theta}_{\mu,4}$	1.0381	1.0632	0.9874	0.9513	0.9646	0.9343	0.9055	0.8808	0.8841	0.8700	0.8951	0.9429
$\hat{\sigma}^2_{\varepsilon_1}$	0.0170	0.0177	0.0192	0.0198	0.0196	0.0194	0.0196	0.0196	0.0198	0.0194	0.0196	0.0199
$\hat{\sigma}^2_{\varepsilon_2}$	0.0197	0.0204	0.0215	0.0223	0.0217	0.0214	0.0223	0.0220	0.0224	0.0219	0.0225	0.0226
$\hat{\sigma}^2_{\varepsilon_3}$	0.0165	0.0174	0.0205	0.0216	0.0211	0.0209	0.0205	0.0213	0.0216	0.0214	0.0219	0.0217
$\hat{\sigma}^2_{\varepsilon_4}$	0.0197	0.0194	0.0216	0.0228	0.0225	0.0223	0.0219	0.0219	0.0218	0.0213	0.0217	0.0212
$\hat{\rho}_{\varepsilon_{2,1}}$	0.9463	0.9468	0.9486	0.9484	0.9456	0.9437	0.9398	0.9391	0.9409	0.9373	0.9376	0.9347
$\hat{\rho}_{\varepsilon_{3,1}}$	0.8674	0.8845	0.8925	0.9012	0.8848	0.8855	0.8809	0.8824	0.8798	0.8700	0.8678	0.8608
$\hat{\rho}_{\varepsilon_{4,1}}$	0.9303	0.9292	0.9310	0.9320	0.9287	0.9196	0.9179	0.9121	0.9125	0.9063	0.9053	0.9027
$\hat{\rho}_{\varepsilon_{3,2}}$	0.8232	0.8364	0.8492	0.8562	0.8404	0.8378	0.8395	0.8401	0.8424	0.8280	0.8258	0.8198
$\hat{\rho}_{\varepsilon_{4,2}}$	0.8944	0.8907	0.8938	0.8943	0.8887	0.8771	0.8751	0.8691	0.8725	0.8624	0.8584	0.8507
$\hat{\rho}_{\varepsilon_{4,3}}$	0.9330	0.9420	0.9506	0.9497	0.9494	0.9546	0.9538	0.9510	0.9497	0.9506	0.9524	0.9494

B.3 Figures

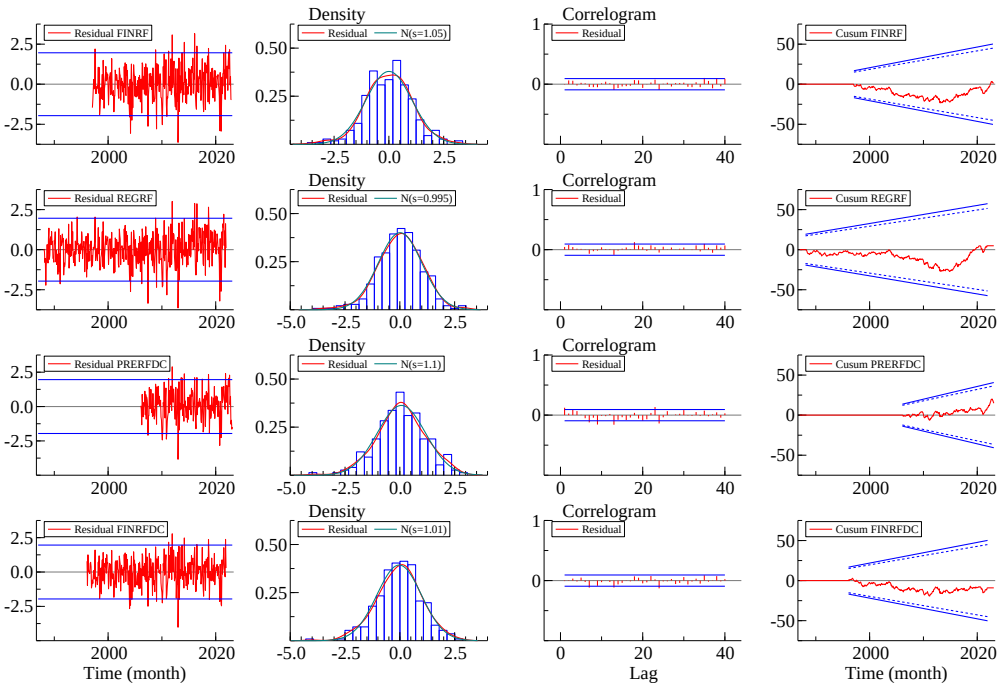

Figure B1: Innovations diagnostics of Model 2, obtained with Kalman smoother.

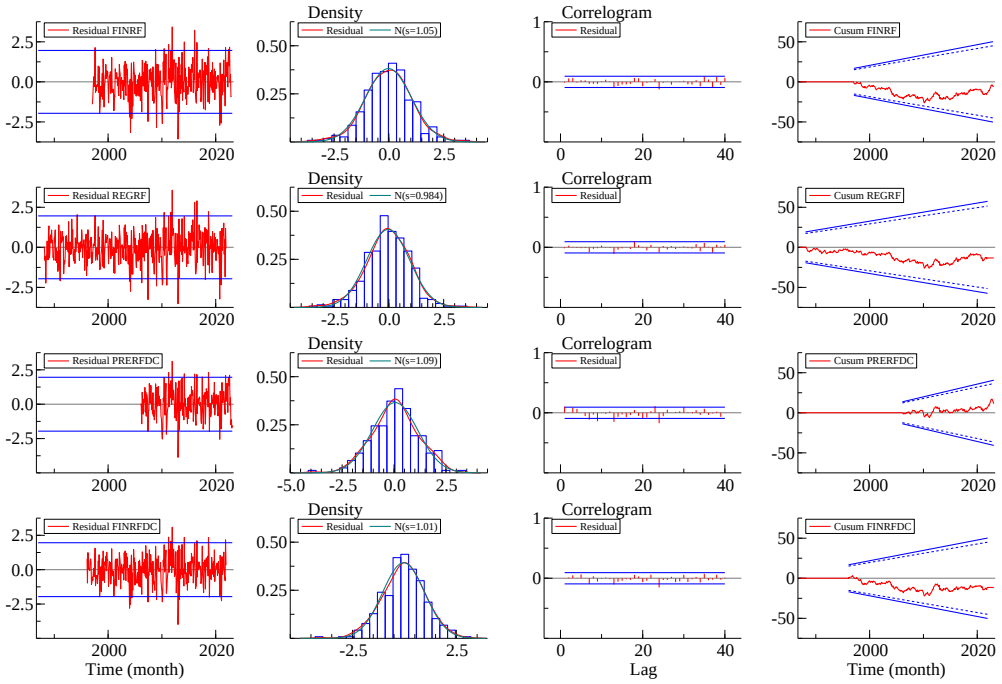

Figure B2: Innovations diagnostics of Model 6, obtained with Kalman smoother.

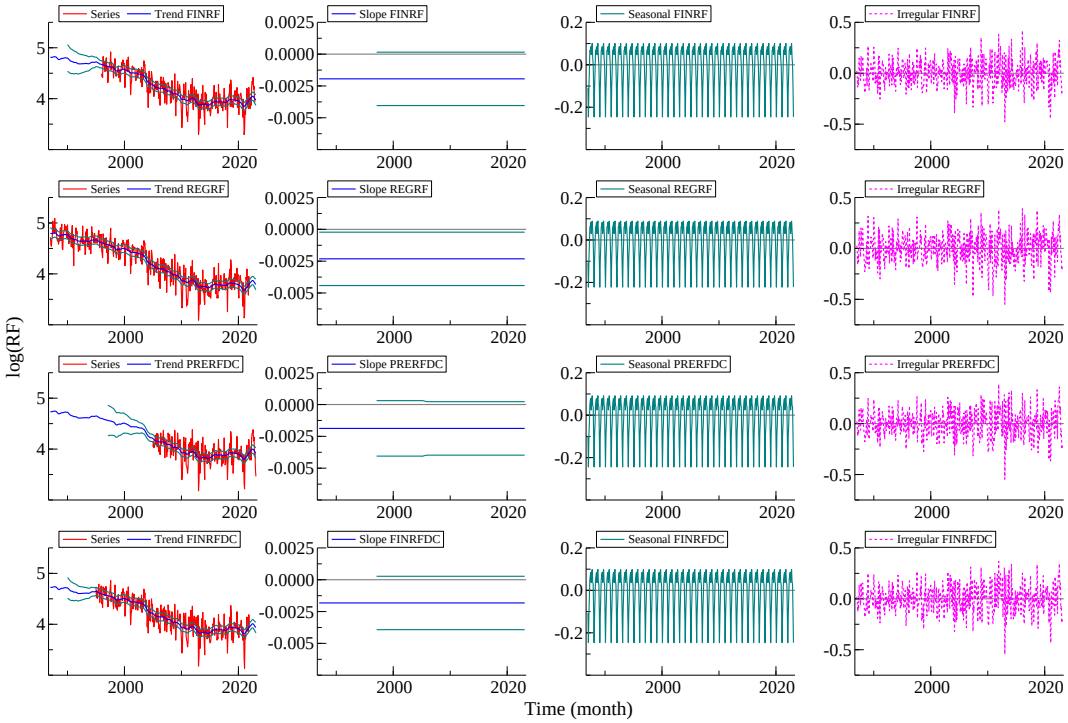

Figure B3: State vector estimates in Model 1, obtained with Kalman smoother.

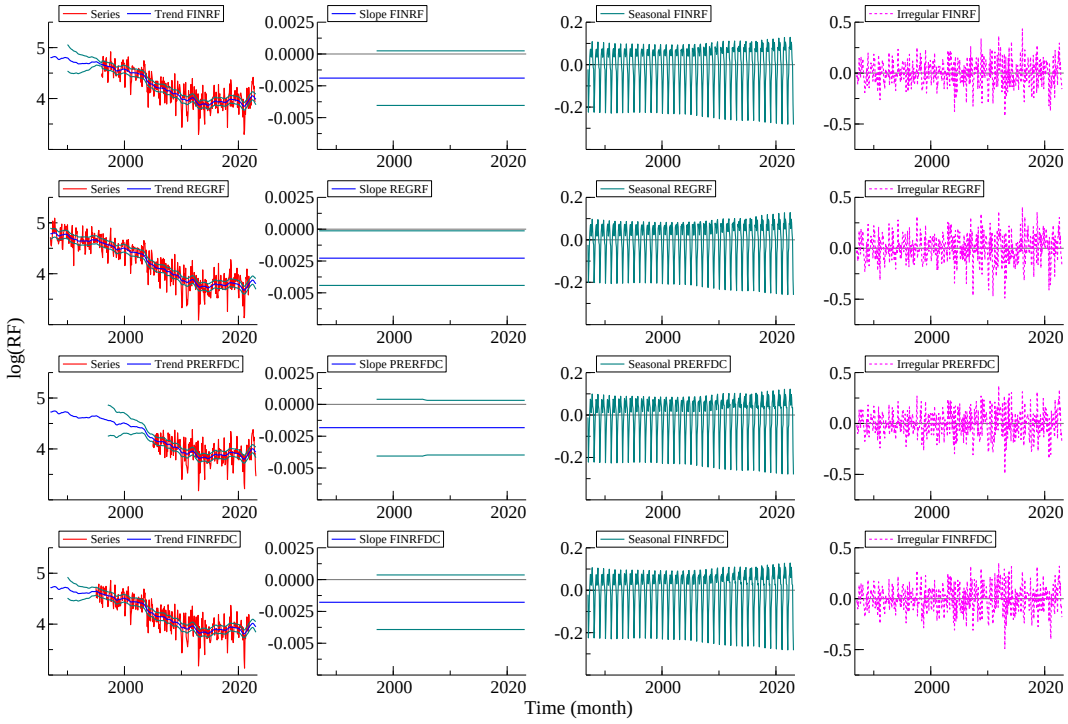

Figure B4: State vector estimates in Model 2, obtained with Kalman smoother.

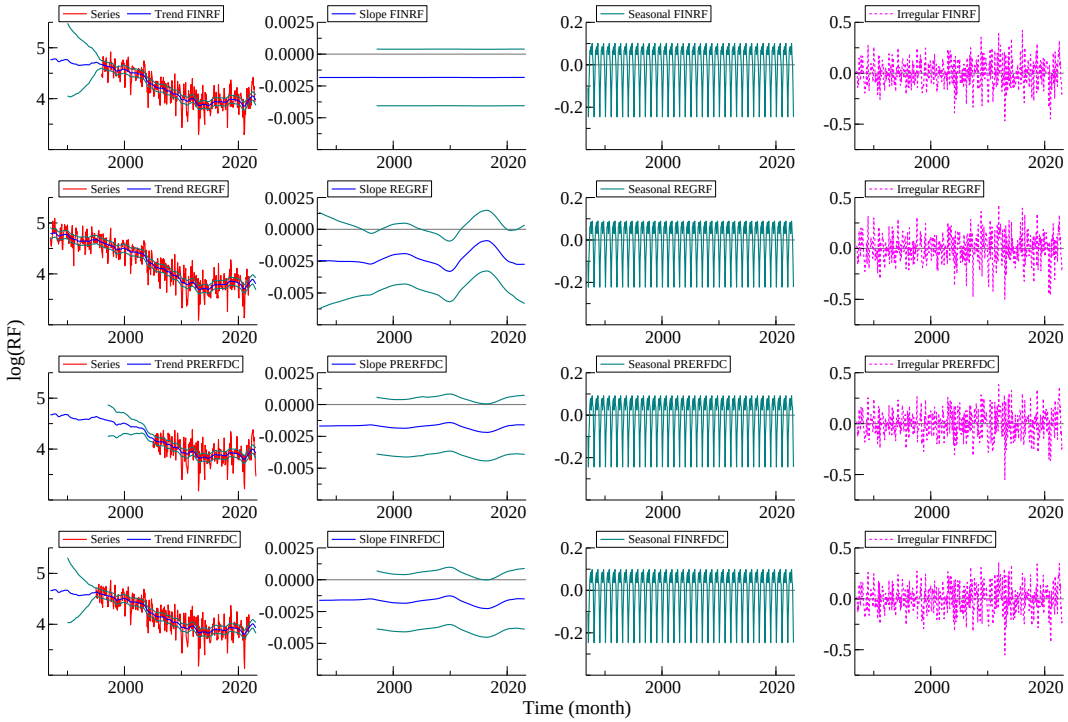

Figure B5: State vector estimates in Model 3, obtained with Kalman smoother.

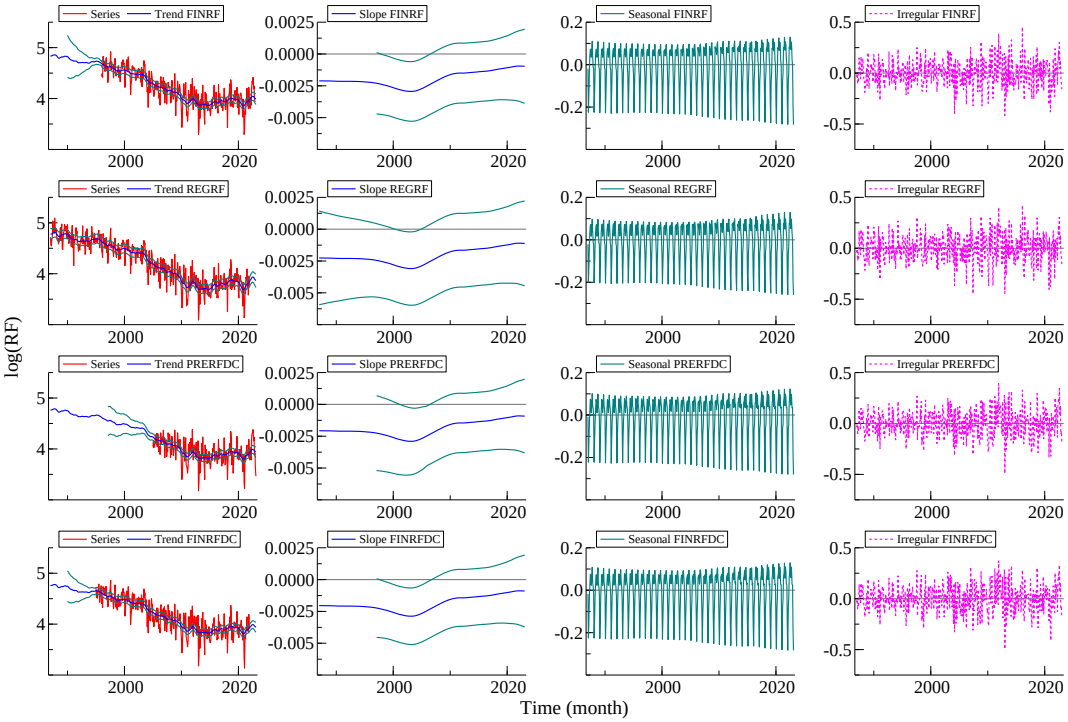

Figure B6: State vector estimates in Model 5, obtained with Kalman smoother.

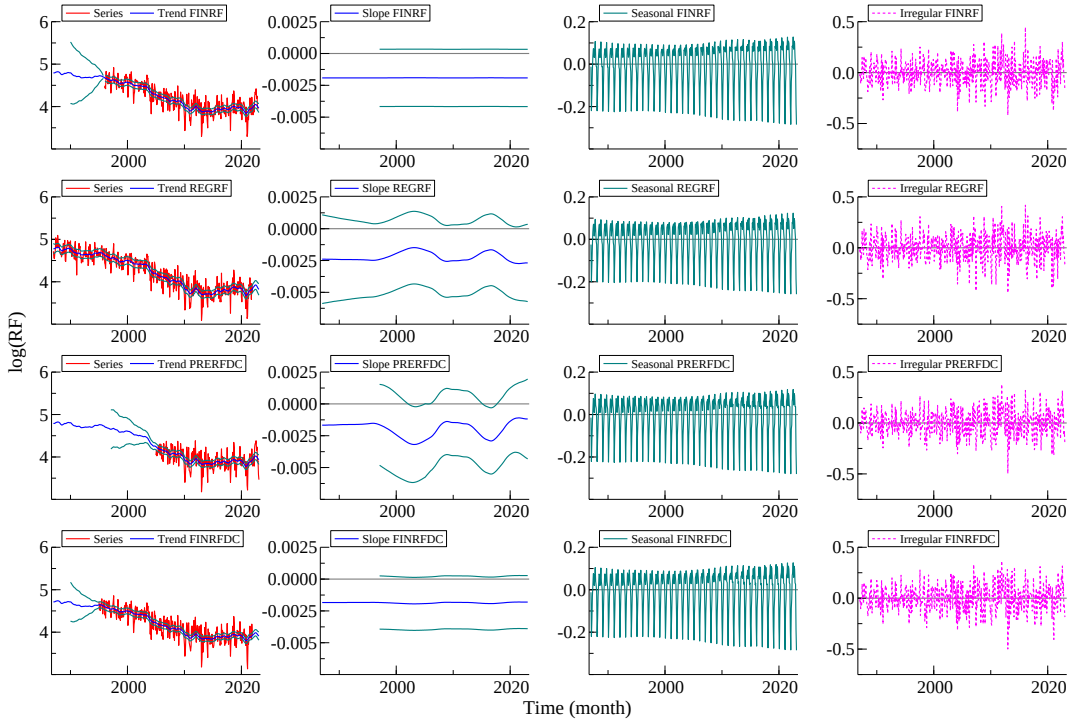

Figure B7: State vector estimates in Model 6, obtained with Kalman smoother.

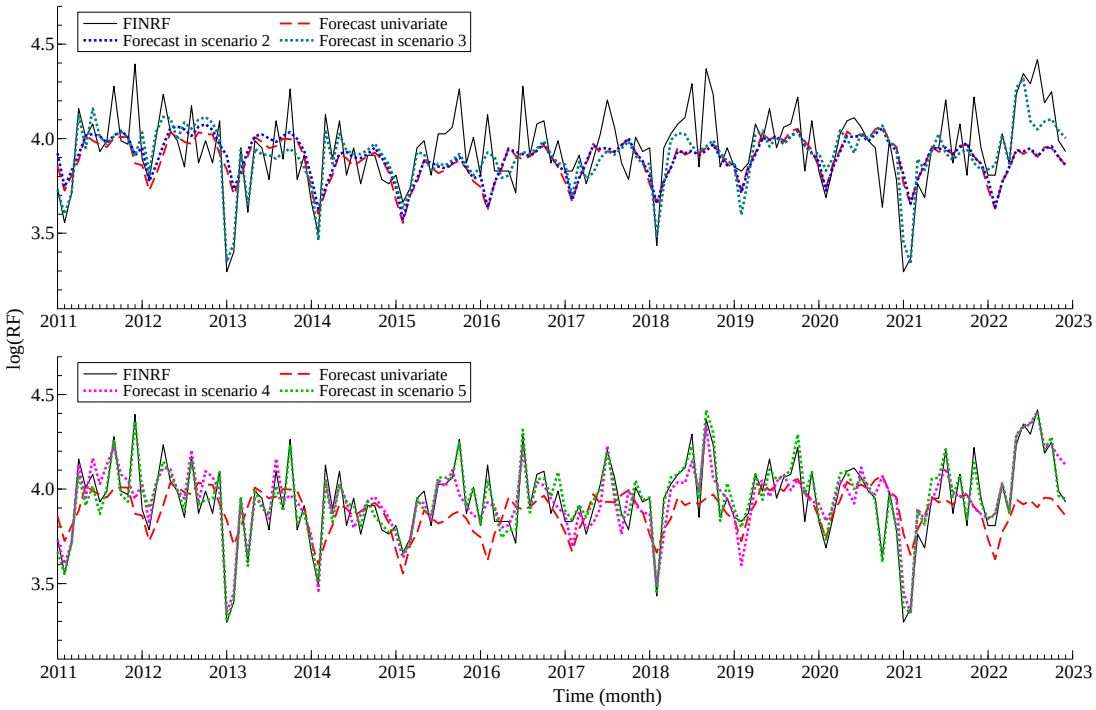

Figure B8: 1- to 12-step ahead forecasts of the FINRF series of Model 1,in the different forecasting scenarios over the forecasting years 2011-2022.

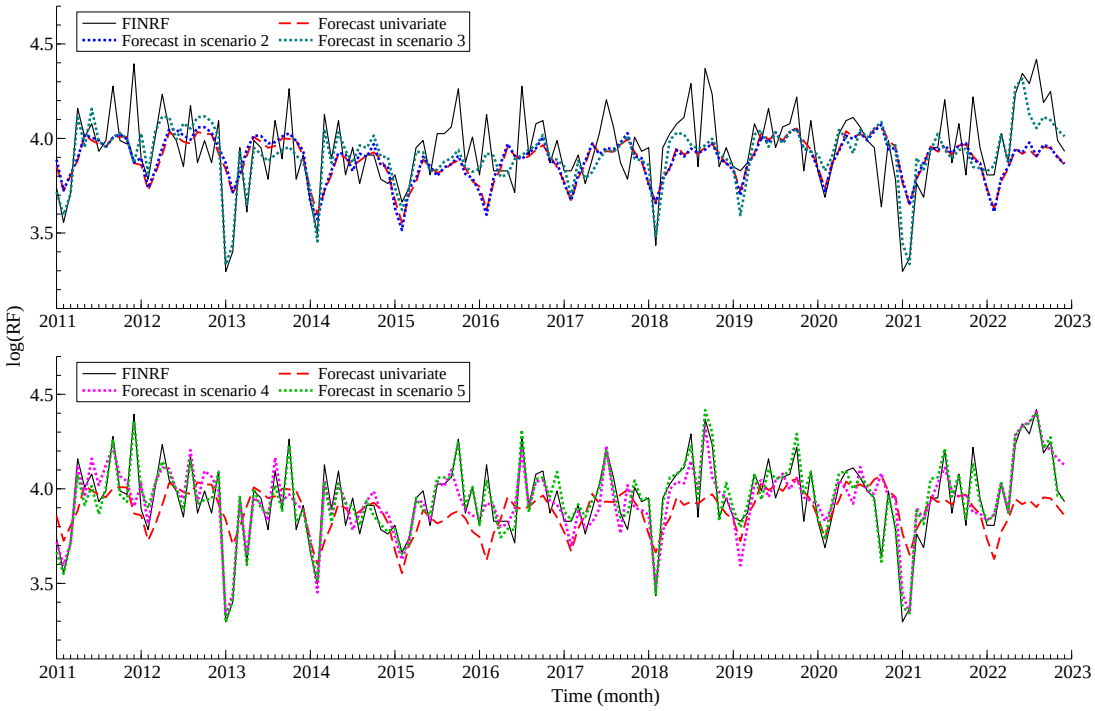

Figure B9: 1- to 12-step ahead forecasts of the FINRF series of Model 2,in the different forecasting scenarios over the forecasting years 2011-2022.

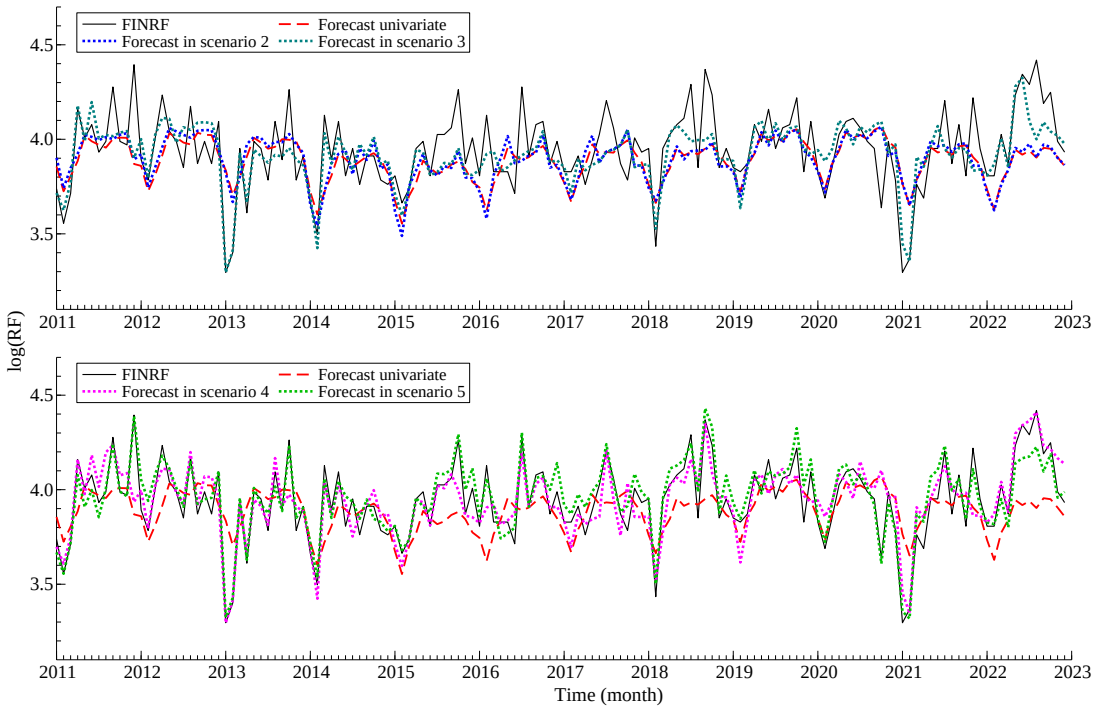

Figure B10: 1- to 12-step ahead forecasts of the FINRF series of Model 3,in the different forecasting scenarios over the forecasting years 2011-2022.

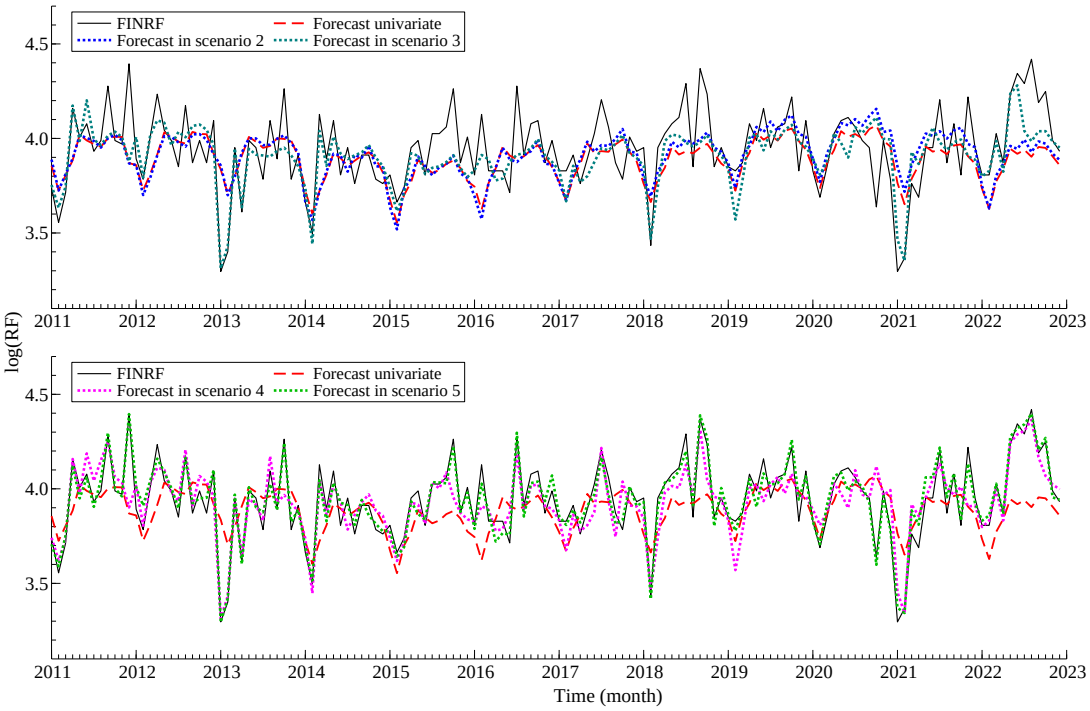

Figure B11: 1- to 12-step ahead forecasts of the FINRF series of Model 5,in the different forecasting scenarios over the forecasting years 2011-2022.

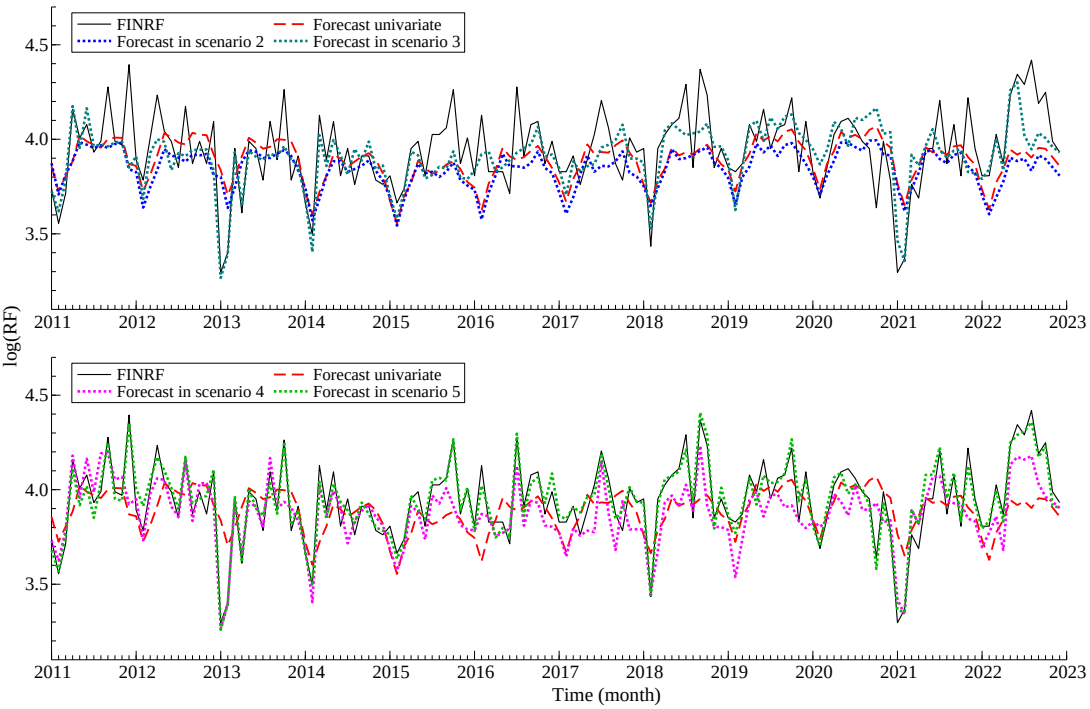

Figure B12: 1- to 12-step ahead forecasts of the FINRF series of Model 6,in the different forecasting scenarios over the forecasting years 2011-2022.

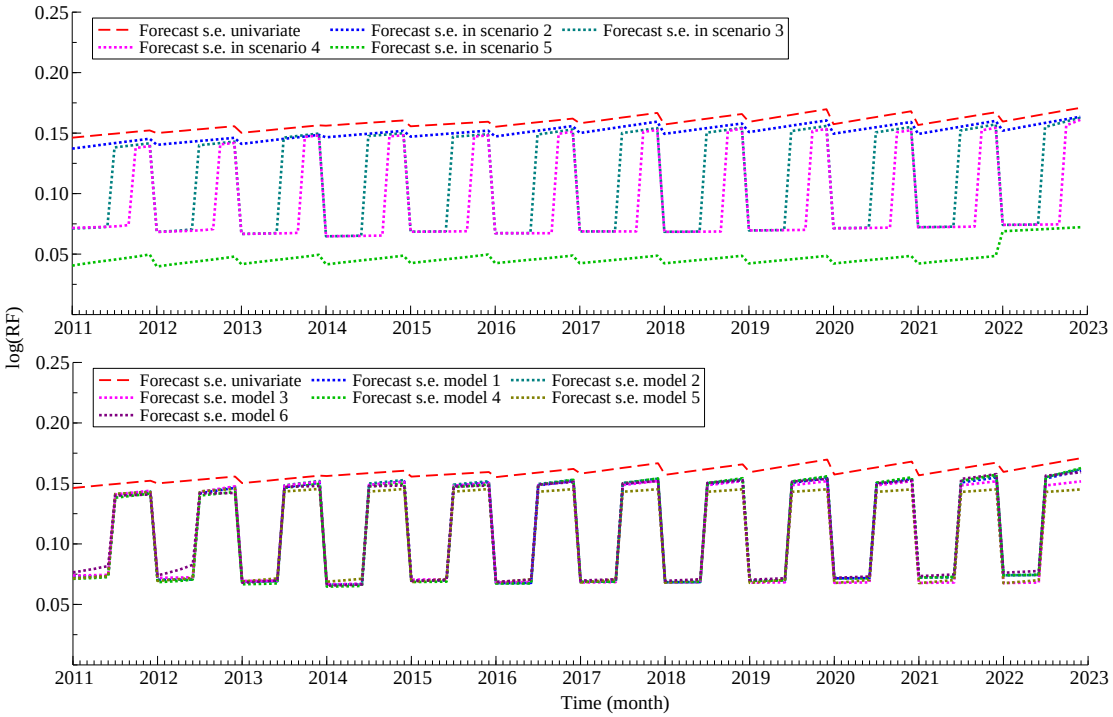

Figure B13: 1- to 12-step ahead forecast standard errors comparison from model 4 in the different scenarios and of the different models in scenario 3 respectively.

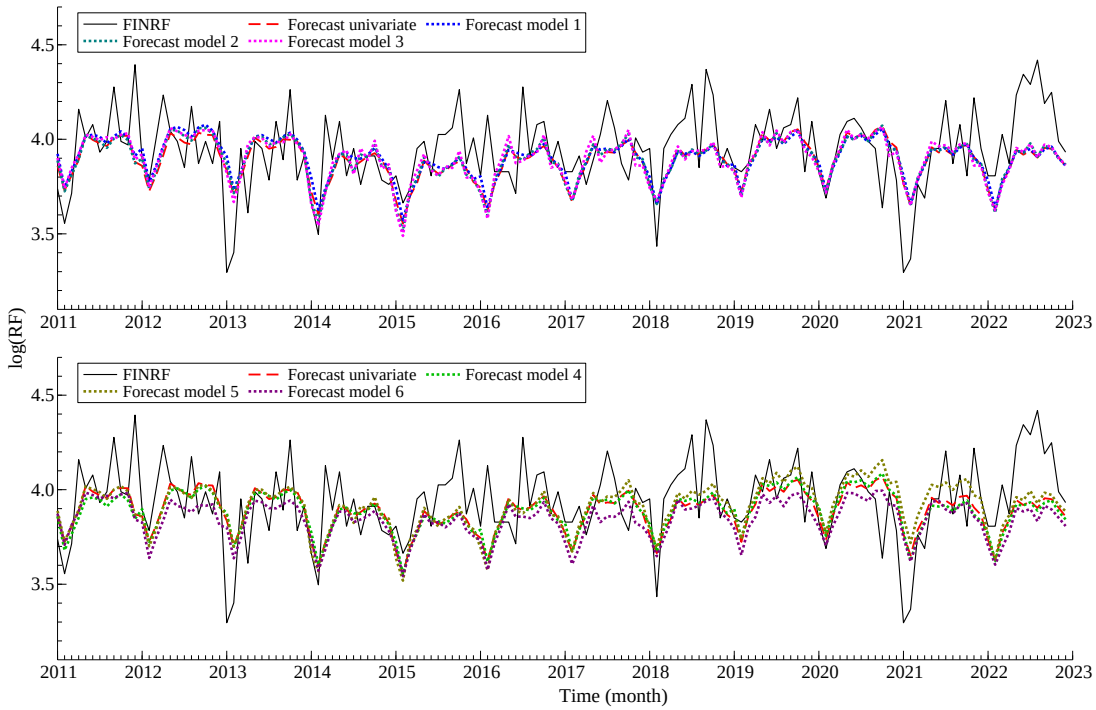

Figure B14: 1- to 12-step ahead forecasts of the FINRF series of the different models in scenario 2, over the forecasting years 2011-2022.

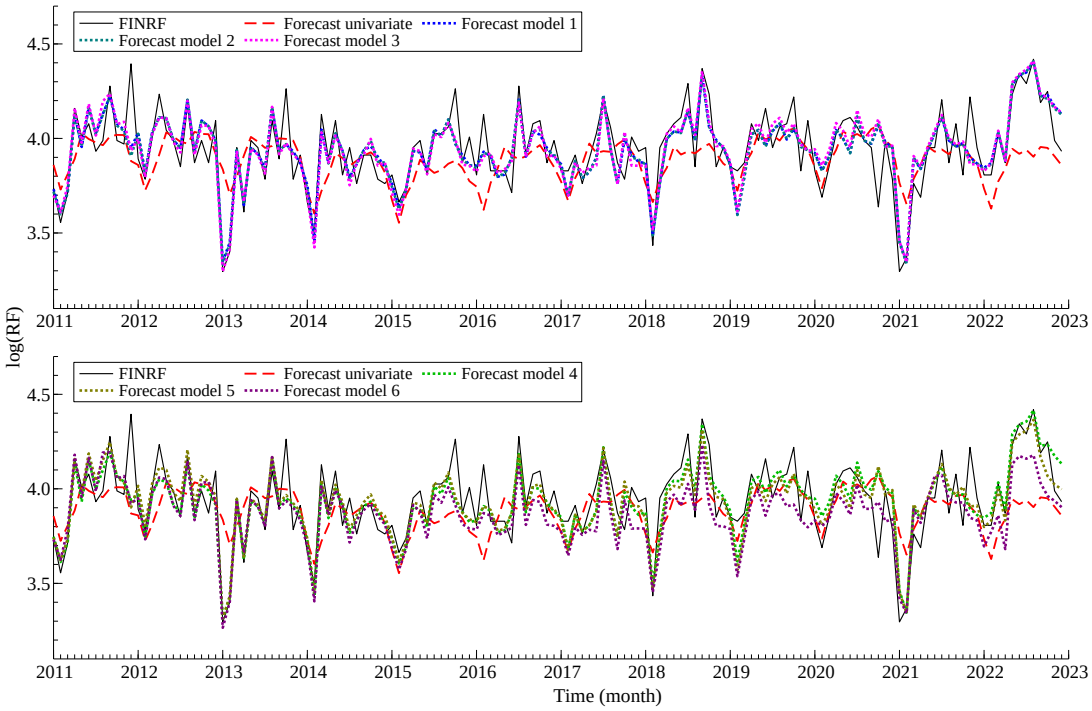

Figure B15: 1- to 12-step ahead forecasts of the FINRF series of the different models in scenario 4, over the forecasting years 2011-2022.

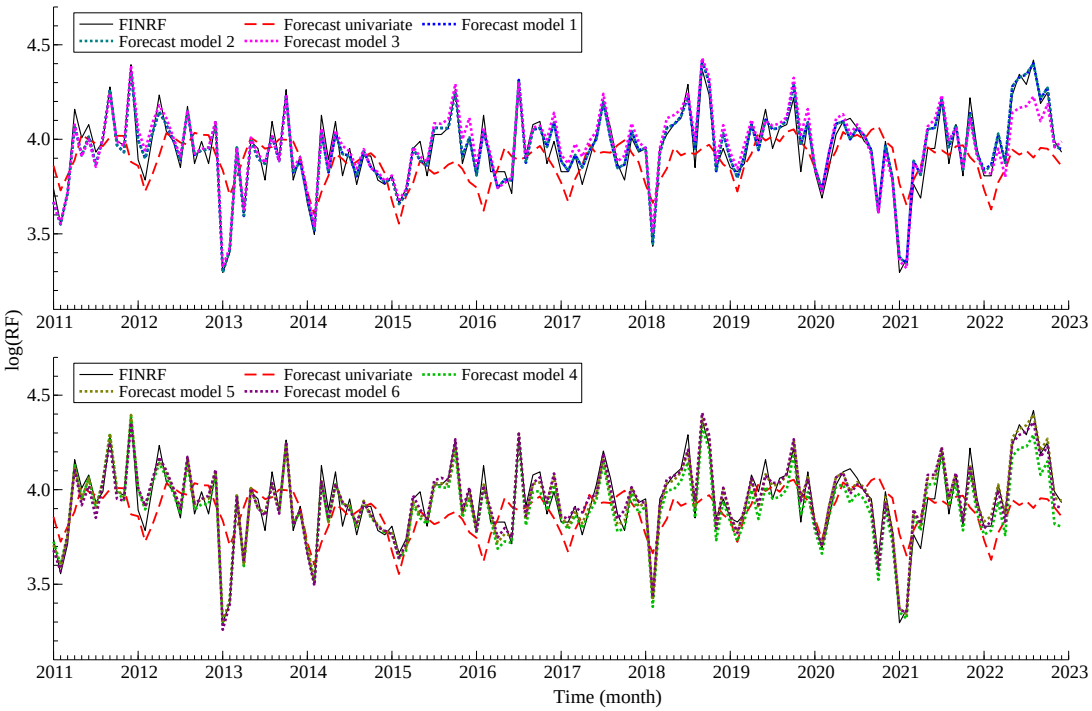

Figure B16: 1- to 12-step ahead forecasts of the FINRF series of the different models in scenario 5, over the forecasting years 2011-2022.

C Supplementary material Section 4

C.1 State space form

The Labour Force model can be written in linear state space form from (A.1)-(A.2). The measurement equation can be written as:

$$\begin{pmatrix} y_t \\ x_{t,1} \\ x_{t,2} \\ x_{t,3} \end{pmatrix} = \left[\begin{array}{c} \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \end{pmatrix} \otimes \boldsymbol{I}_4 \end{array} \begin{array}{c} \begin{pmatrix} 0 & 0 & 0 \\ k_{t,1} & 0 & 0 \\ 0 & k_{t,2} & 0 \\ 0 & 0 & k_{t,3} \end{pmatrix} \end{array} \right] \boldsymbol{\alpha}_t + \begin{pmatrix} \varepsilon_{t,y} \\ \varepsilon_{t,1} \\ \varepsilon_{t,2} \\ \varepsilon_{t,3} \end{pmatrix}, \tag{C.1}$$

where the state vector $\boldsymbol{\alpha}_t$ is specified as follows:

$$\boldsymbol{\alpha}_t = \begin{pmatrix} \boldsymbol{\mu}_t' & \boldsymbol{\nu}_t' & \boldsymbol{\gamma}_t' & \boldsymbol{u}_t' \end{pmatrix}', \quad \boldsymbol{\mu}_t = \begin{pmatrix} \mu_{t,y} & \mu_{t,x_1} & \mu_{t,x_2} & \mu_{t,x_3} \end{pmatrix}', \quad \boldsymbol{\nu}_t = \begin{pmatrix} \nu_{t,y} & \nu_{t,x_1} & \nu_{t,x_2} & \nu_{t,x_3} \end{pmatrix}',$$
$$\boldsymbol{\gamma}_t = \begin{pmatrix} \gamma_{t,y} & \gamma_{t,x_1} & \gamma_{t,x_2} & \gamma_{t,x_3} & \cdots & \gamma_{t-2,y} & \gamma_{t-2,x_1} & \gamma_{t-2,x_2} & \gamma_{t-2,x_3} \end{pmatrix}', \quad \boldsymbol{u}_t = \begin{pmatrix} u_{t,1} & u_{t,2} & u_{t,3} \end{pmatrix}',$$

which contains the trend, dummy seasonal and sampling error terms. As discussed the idiosyncratic noise of the target and LFS auxiliary series are Gaussian i.i.d. sequences with the following diagonal irregular covariance matrix:

$$\boldsymbol{\Sigma}_\varepsilon = \begin{bmatrix} \sigma^2_{\varepsilon_y} & 0 & 0 & 0 \\ 0 & \sigma^2_{\varepsilon_x} & 0 & 0 \\ 0 & 0 & \sigma^2_{\varepsilon_x} & 0 \\ 0 & 0 & 0 & \sigma^2_{\varepsilon_x} \end{bmatrix}. \tag{C.2}$$

The state vector's dynamic transitioning can be described as follows:

$$\boldsymbol{\alpha}_{t+1} = \begin{bmatrix} \begin{pmatrix} \boldsymbol{T}_\mu & \boldsymbol{0}_{8\times 12} \\ \boldsymbol{0}_{12\times 8} & \boldsymbol{T}_\gamma \end{pmatrix} & \boldsymbol{0}_{20\times 3} \\ \boldsymbol{0}_{3\times 20} & \begin{pmatrix} \phi & 0 & 0 \\ 0 & \phi & 0 \\ 0 & 0 & \phi \end{pmatrix} \end{bmatrix} \boldsymbol{\alpha}_t + \boldsymbol{\eta}_t, \qquad \boldsymbol{\eta}_t \sim \mathbf{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_\eta), \tag{C.3}$$

where ϕ is the monthly autoregressive coefficient of the sampling error, retrieved from van den Brakel and Michiels (2021) and treated as known. We opt for a local linear trend model and a stochastic seasonal dummy model, again see Harvey (1989) and Durbin and Koopman (2012), by specifying:

$$\boldsymbol{T}_\mu = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \otimes \boldsymbol{I}_4, \quad \boldsymbol{T}_\gamma = \begin{bmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \otimes \boldsymbol{I}_4, \quad \boldsymbol{\eta}_t = \begin{pmatrix} \boldsymbol{\xi}_t' & \boldsymbol{\zeta}_t' & \boldsymbol{\omega}_t' & \boldsymbol{0}_{1\times 8} & \boldsymbol{e}_t' \end{pmatrix}',$$
$$\boldsymbol{\xi}_t = \begin{pmatrix} \xi_{t,y} \\ \xi_{t,x_1} \\ \xi_{t,x_2} \\ \xi_{t,x_3} \end{pmatrix}, \quad \boldsymbol{\zeta}_t = \begin{pmatrix} \zeta_{t,y} \\ \zeta_{t,x_1} \\ \zeta_{t,x_2} \\ \zeta_{t,x_3} \end{pmatrix}, \quad \boldsymbol{\omega}_t = \begin{pmatrix} \omega_{t,y} \\ \omega_{t,x_1} \\ \omega_{t,x_2} \\ \omega_{t,x_3} \end{pmatrix}, \quad \boldsymbol{e}_t = \begin{pmatrix} e_{t,1} \\ e_{t,2} \\ e_{t,3} \end{pmatrix}.$$

The covariance matrix of the disturbance terms is a blockdiagonal matrix of:

$$\boldsymbol{\Sigma}_\eta = \text{diag}(\boldsymbol{\Sigma}_\xi, \boldsymbol{\Sigma}_\zeta, \boldsymbol{\Sigma}_\omega, \boldsymbol{0}_{8\times 8}, \boldsymbol{\Sigma}_e), \tag{C.4}$$

where,

$$\boldsymbol{\Sigma}_e = \begin{bmatrix} \sigma_e^2 & 0 & 0 \\ 0 & \sigma_e^2 & 0 \\ 0 & 0 & \sigma_e^2 \end{bmatrix}, \tag{C.5}$$

and it is expected that $\sigma_e^2 = 1$ due to the scaling of the sampling error.

C.2 Tables

Table C1: Parameter estimates with constant (left) and time-varying (right) variances for trend disturbance terms.

Model	1	2	3	4	1	2	3	4
$\hat{\sigma}^2_{\xi_1}$	0.1235	1.133e-08	0.0897	0.0914	0.0580	0.0026	0.0346	0.0346
$\hat{\sigma}^2_{\xi_2}$	-	-	0.0475	0.0464	-	-	0.0308	0.0324
$\hat{\theta}_{\mu_1,2}$	0.3278	1926.0761	0.4350	0.4371	0.3830	3.6643	0.3410	0.2819
$\hat{\theta}_{\mu_1,3}$	0.3776	1784.4226	0.4788	0.4673	0.4160	3.4528	0.4089	0.3625
$\hat{\theta}_{\mu_1,4}$	0.4553	1675.4192	0.3776	0.4069	0.4684	3.3569	0.5497	0.5748
$\hat{\theta}_{\mu_2,3}$	-	-	0.7975	0.8017	-	-	0.7533	0.7718
$\hat{\theta}_{\mu_2,4}$	-	-	-0.0140	-0.0359	-	-	-0.2062	-0.1616
$\hat{\sigma}^2_{\zeta_1}$	0.0072	0.0340	0.0179	0.0176	0.0051	0.0296	0.0097	0.0100
$\hat{\sigma}^2_{\zeta_2}$	-	-	0.0007	0.0007	-	-	0.0009	0.0009
$\hat{\theta}_{\nu_1,2}$	1.1395	0.5876	0.6442	0.6461	1.1465	0.6122	0.7551	0.7431
$\hat{\theta}_{\nu_1,3}$	1.1814	0.6770	0.7726	0.7715	1.1822	0.6772	0.8442	0.8273
$\hat{\theta}_{\nu_1,4}$	1.2859	0.7931	0.9999	0.9925	1.2823	0.7735	0.9944	0.9726
$\hat{\theta}_{\nu_2,3}$	-	-	0.9542	0.9548	-	-	0.9741	0.9756
$\hat{\theta}_{\nu_2,4}$	-	-	1.0569	1.0719	-	-	1.0901	1.1104
$\hat{\sigma}^2_{\omega_1}$	0.0029	0	0.0039	0.0041	0.0043	0.0036	0.0053	0.0049
$\hat{\sigma}^2_{\omega_2}$	-	0	-	3.393e-11	-	3.920e-11	-	4.530e-12
$\hat{\sigma}^2_{\omega_3}$	-	0	-	0.0002	-	0.0008	-	0.0007
$\hat{\sigma}^2_{\omega_4}$	-	0	-	4.369e-13	-	9.617e-13	-	7.222e-12
$\hat{\theta}_{\gamma_1,2}$	0.0086	-	0.0909	-	-0.2181	-	-0.1520	-
$\hat{\theta}_{\gamma_1,3}$	-0.4189	-	-0.2201	-	-0.6051	-	-0.4324	-
$\hat{\theta}_{\gamma_1,4}$	0.0714	-	0.0483	-	-0.0316	-	-0.0606	-
$\hat{\sigma}^2_e$	1.1865	0.9902	0.4702	0.4652	1.1565	1.0536	0.5551	0.5659
$\hat{\sigma}^2_{\varepsilon_y}$	2.933e-18	0.0535	4.697e-08	1.015e-19	8.417e-13	0.0172	0.0022	0.0038
$\hat{\sigma}^2_{\varepsilon_x}$	1.243e-07	0.0049	0.0151	0.0155	6.369e-10	1.004e-13	0.0125	0.0117

Table C2: Time-varying Covid correction factors.

Component	Factor	Quarter					
		2019(2)	2019(3)	2019(4)	2020(1)	2020(2)	2020(3)
Level	$\sigma^2_{\xi,t}$	1	10	20	20	10	1
Slope	$\sigma^2_{\zeta,t}$	1	2	4	4	2	1

Table C3: MSE Forecasting performance based on the 5-step ahead forecasts per forecasting year and in total, with constant and time-varying* variances for trend disturbance terms.

Scenario	Model	2011	2012	2013	2014	2015	2016	2017	2018	2019	(2019)	2020	(2020)	2021	(2021)	Total	(Total)
1	BSTSM	3.231	1.695	0.308	0.326	0.983	0.276	0.062	0.086	0.101	0.101	6.074	6.074	17.488	26.429	2.785	3.597
2	1	2.563	0.709	0.092	0.809	0.547	0.385	0.082	0.063	0.133	0.126	2.384	6.086	2.599	15.002	0.942	2.406
	2	2.079	0.421	0.103	0.633	0.437	0.361	0.070	0.037	0.084	0.082	2.310	6.439	2.793	17.144	0.848	2.528
	3	0.988	0.626	0.067	0.774	0.578	0.478	0.126	0.113	0.217	0.209	2.241	6.581	2.243	6.879	0.768	1.584
	4	1.506	0.710	0.291	1.207	0.116	0.331	0.377	0.052	0.059	0.058	1.666	2.354	2.011	9.732	0.757	1.521
3	1	2.584	0.701	0.110	0.701	0.573	0.414	0.098	0.119	0.097	0.090	2.155	5.722	2.946	15.716	0.954	2.439
	2	2.291	0.392	0.103	0.535	0.456	0.391	0.090	0.086	0.060	0.058	2.126	5.969	3.022	17.922	0.868	2.572
	3	0.992	1.471	0.202	1.214	0.414	0.136	0.697	0.238	0.038	0.037	1.492	2.356	2.237	2.851	0.830	0.964
	4	1.654	1.452	0.361	1.363	0.103	0.086	0.537	0.150	0.040	0.039	1.479	2.270	2.300	10.415	0.866	1.675
4	1	2.456	0.554	0.016	0.396	0.462	0.437	0.129	0.189	0.177	0.154	2.236	5.709	2.667	13.062	0.884	2.142
	2	1.744	0.377	0.132	0.704	0.367	0.373	0.042	0.062	0.078	0.092	1.938	5.071	3.006	17.919	0.802	2.444
	3	1.060	0.630	0.219	0.873	0.207	0.277	0.471	0.233	0.165	0.157	1.527	2.854	2.867	12.879	0.775	1.806
	4	1.513	0.560	0.323	1.010	0.119	0.142	0.366	0.430	0.101	0.089	1.485	3.377	2.953	14.023	0.818	1.995

*Brackets refer to results including correction.

Table C4: MAE Forecasting performance based on the 5-step ahead forecasts per forecasting year and in total, with constant and time-varying* variances for trend disturbance terms.

Scenario	Model	2011	2012	2013	2014	2015	2016	2017	2018	2019	(2019)	2020	(2020)	2021	(2021)	Total	(Total)
1	BSTSM	1.432	1.173	0.462	0.480	0.951	0.368	0.226	0.275	0.279	0.279	2.266	2.266	3.882	4.442	1.072	1.123
2	1	1.560	0.701	0.267	0.846	0.710	0.531	0.222	0.197	0.352	0.335	1.459	2.202	1.268	3.725	0.738	1.027
	2	1.335	0.568	0.311	0.737	0.646	0.517	0.206	0.147	0.284	0.280	1.416	2.296	1.302	3.986	0.679	1.002
	3	0.944	0.632	0.188	0.813	0.704	0.609	0.281	0.283	0.441	0.425	1.420	2.257	1.185	1.884	0.682	0.820
	4	1.187	0.732	0.422	0.953	0.303	0.524	0.536	0.197	0.229	0.221	1.159	1.370	1.284	2.847	0.684	0.845
3	1	1.557	0.701	0.284	0.795	0.731	0.575	0.259	0.291	0.307	0.290	1.401	2.145	1.339	3.801	0.749	1.039
	2	1.365	0.569	0.300	0.687	0.663	0.558	0.259	0.239	0.241	0.236	1.375	2.224	1.349	4.052	0.692	1.014
	3	0.772	0.996	0.407	1.065	0.473	0.324	0.754	0.431	0.181	0.170	1.154	1.397	1.333	1.387	0.717	0.743
	4	1.199	1.182	0.596	1.027	0.287	0.281	0.652	0.371	0.184	0.174	1.114	1.336	1.377	2.932	0.752	0.913
4	1	1.528	0.565	0.110	0.523	0.645	0.517	0.340	0.398	0.380	0.358	1.326	2.149	1.336	3.591	0.697	0.975
	2	1.201	0.567	0.338	0.813	0.595	0.515	0.184	0.211	0.274	0.291	1.296	1.954	1.309	4.082	0.664	0.977
	3	0.968	0.708	0.305	0.861	0.378	0.389	0.646	0.426	0.401	0.390	1.158	1.480	1.293	3.417	0.685	0.906
	4	1.167	0.685	0.402	0.926	0.314	0.301	0.512	0.655	0.304	0.287	1.139	1.619	1.335	3.559	0.704	0.948

*Brackets refer to results including correction.

C.3 Figures

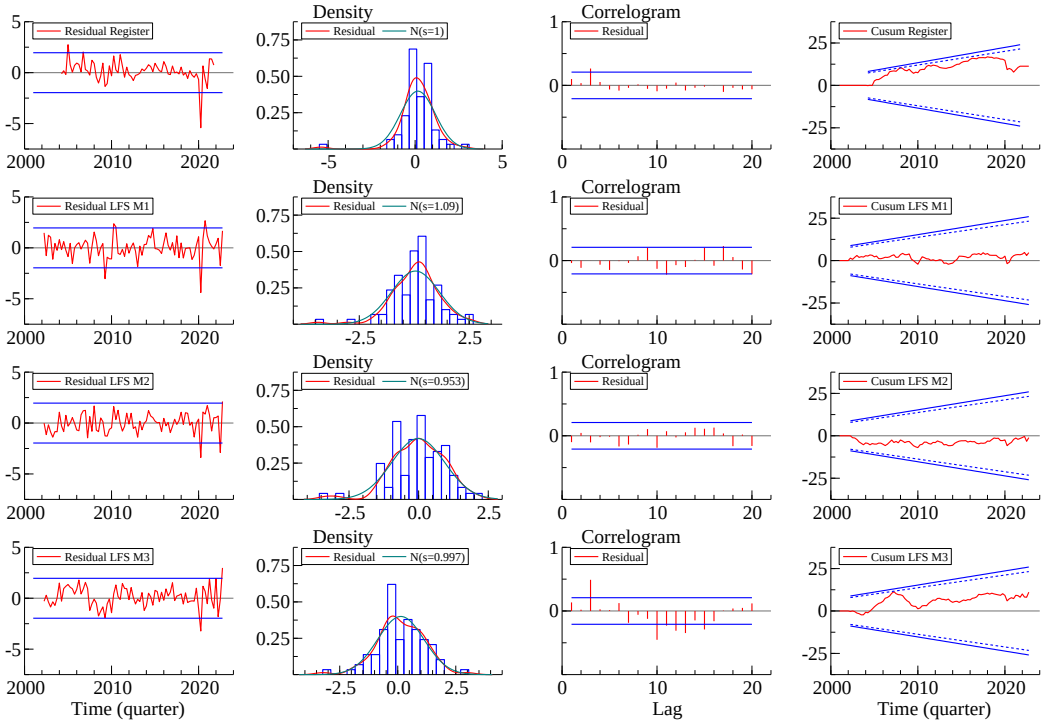

Figure C1: Innovations diagnostics of Model 1, obtained with Kalman smoother.

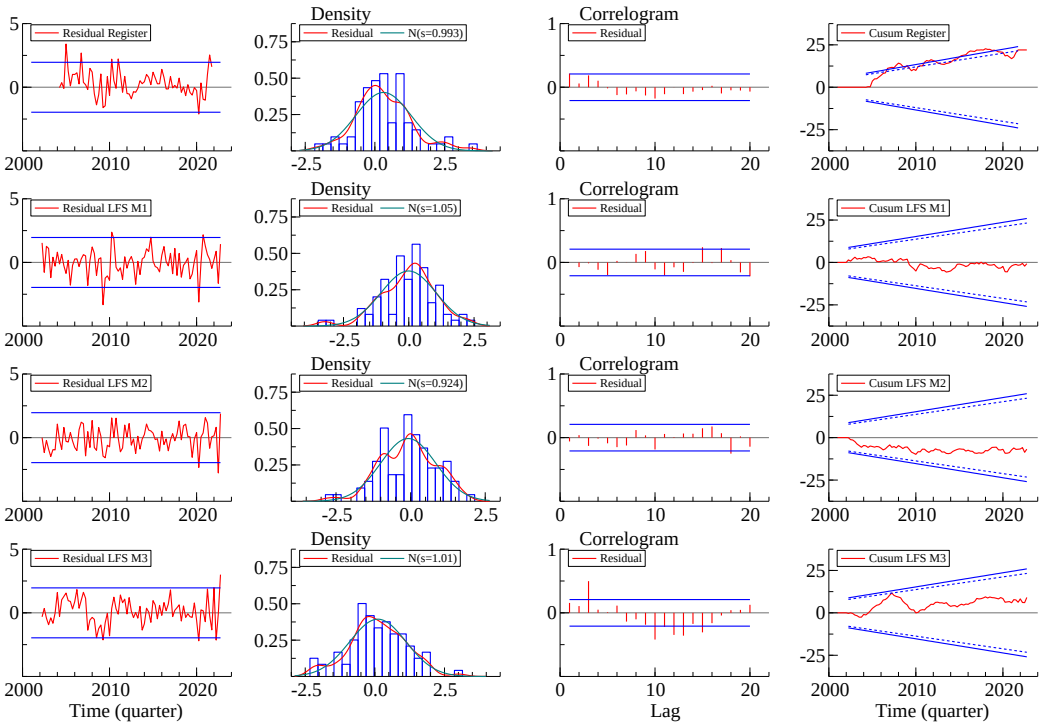

Figure C2: Innovations diagnostics of Model 1 including time-varying variances for the trend disturbance terms at the start of the Covid outbreak.

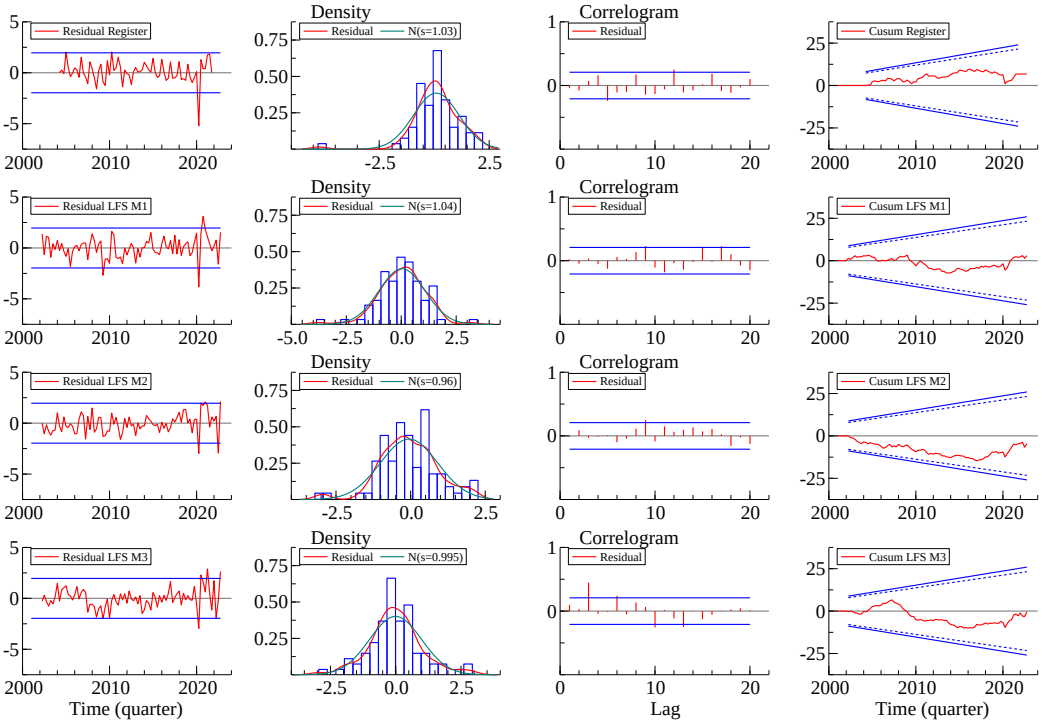

Figure C3: Innovations diagnostics of Model 2, obtained with Kalman smoother.

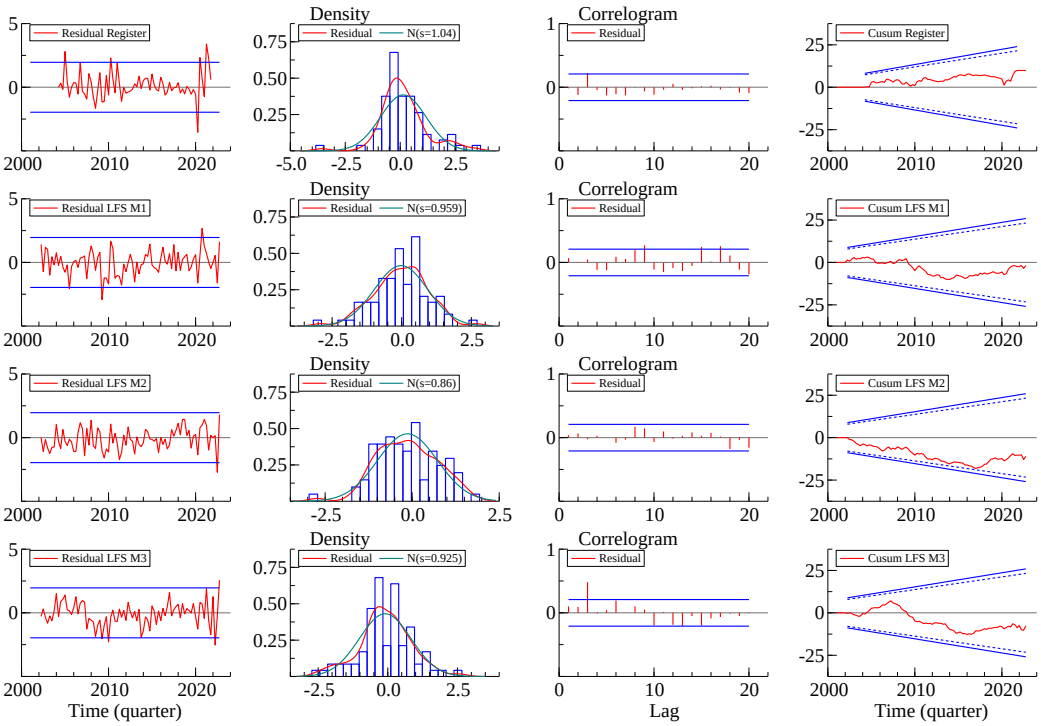

Figure C4: Innovations diagnostics of Model 2 including time-varying variances for the trend disturbance terms at the start of the Covid outbreak.

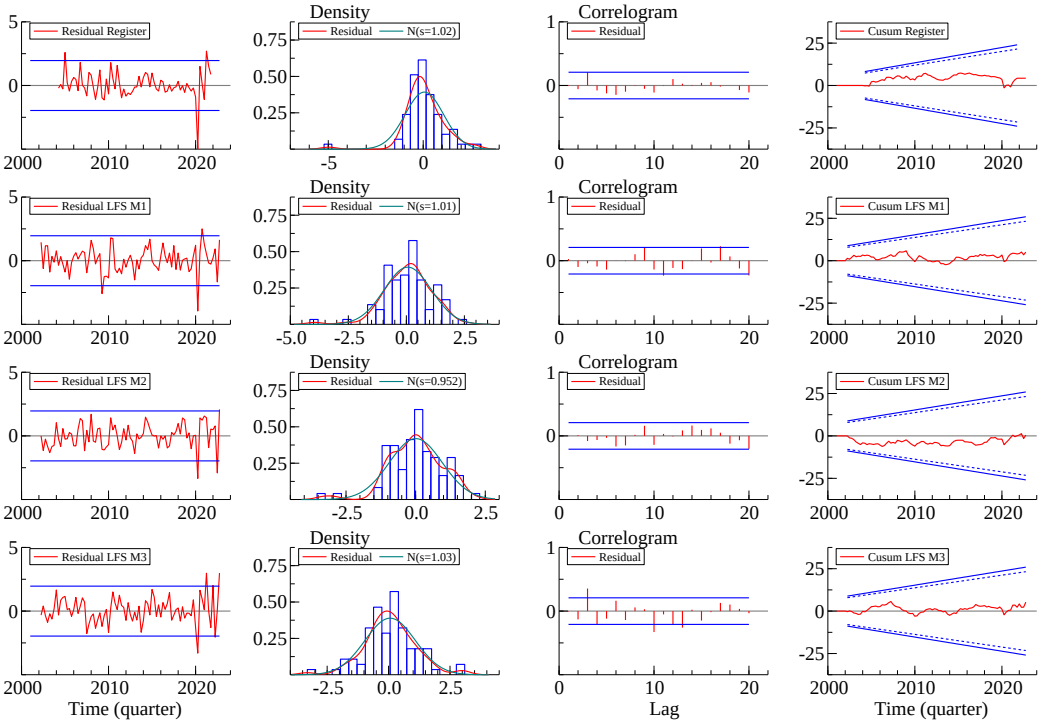

Figure C5: Innovations diagnostics of Model 4, obtained with Kalman smoother.

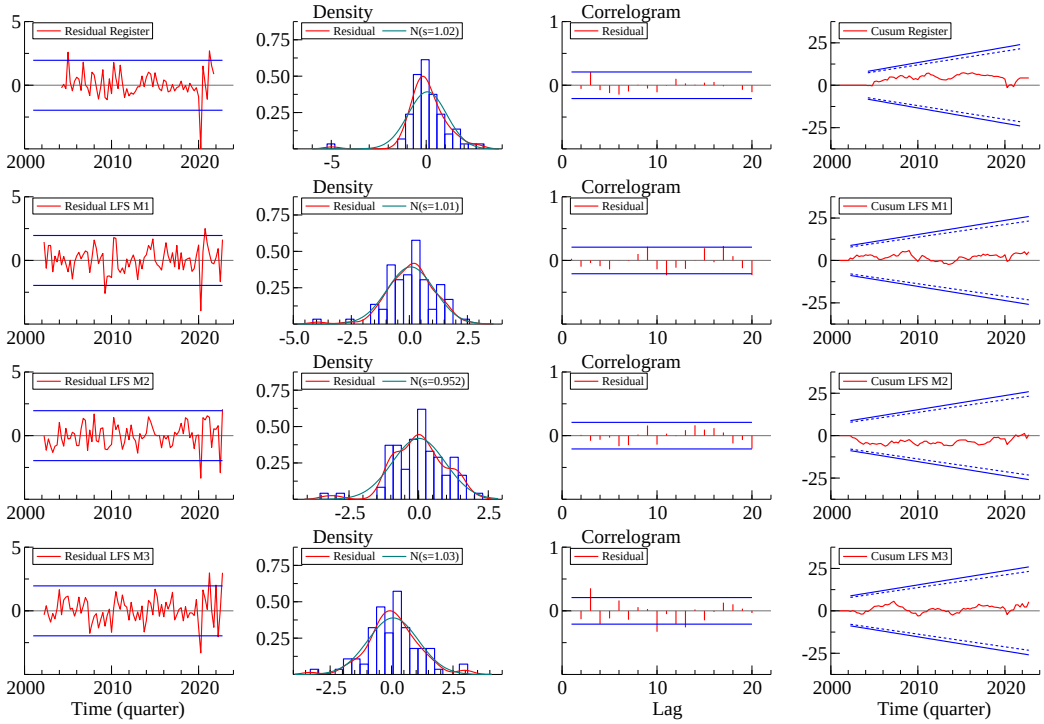

Figure C6: Innovations diagnostics of Model 4 including time-varying variances for the trend disturbance terms at the start of the Covid outbreak.

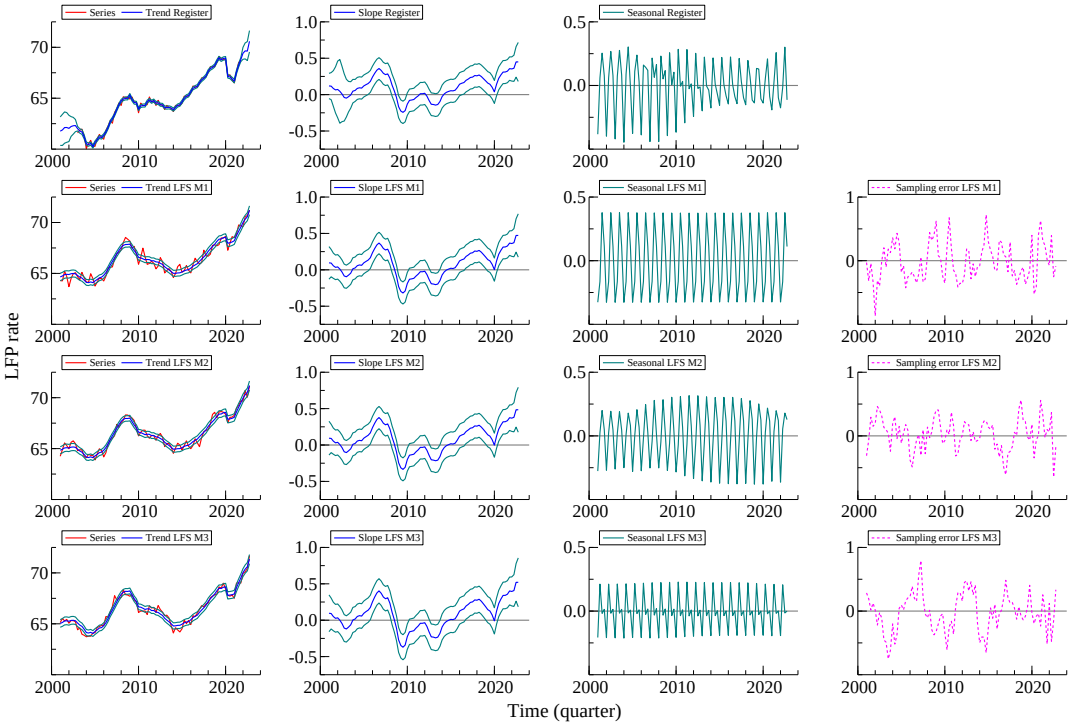

Figure C7: State vector estimates in Model 1, obtained with Kalman smoother.

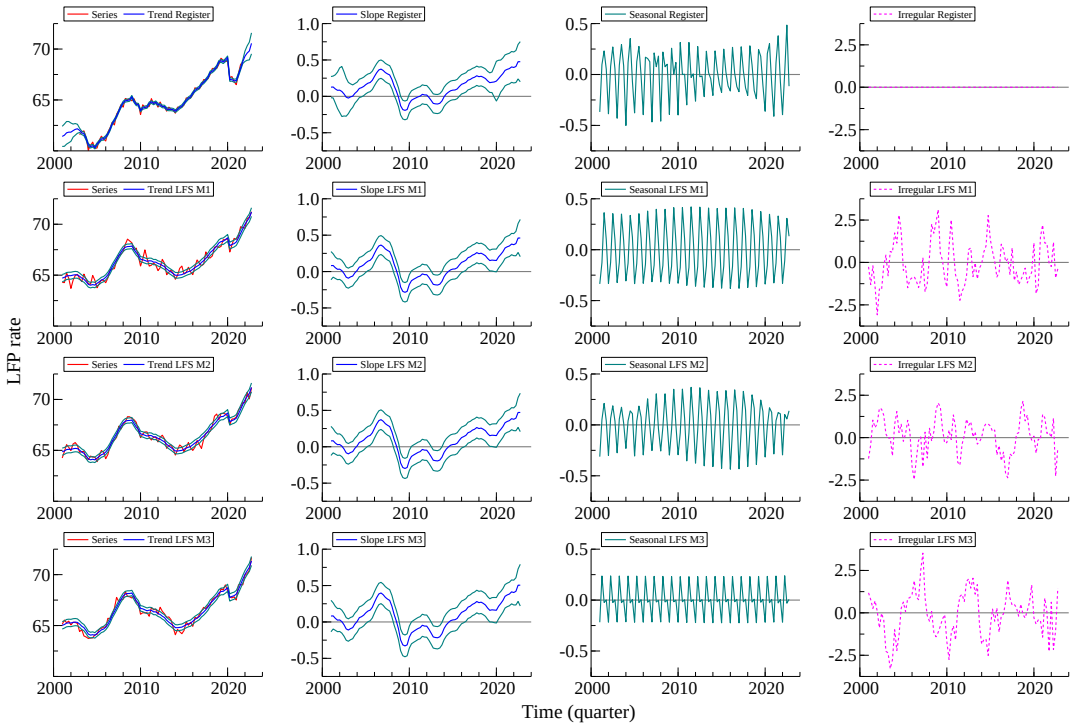

Figure C8: State vector estimates in Model 1 including time-varying variances for the trend disturbance terms at the start of the Covid outbreak.

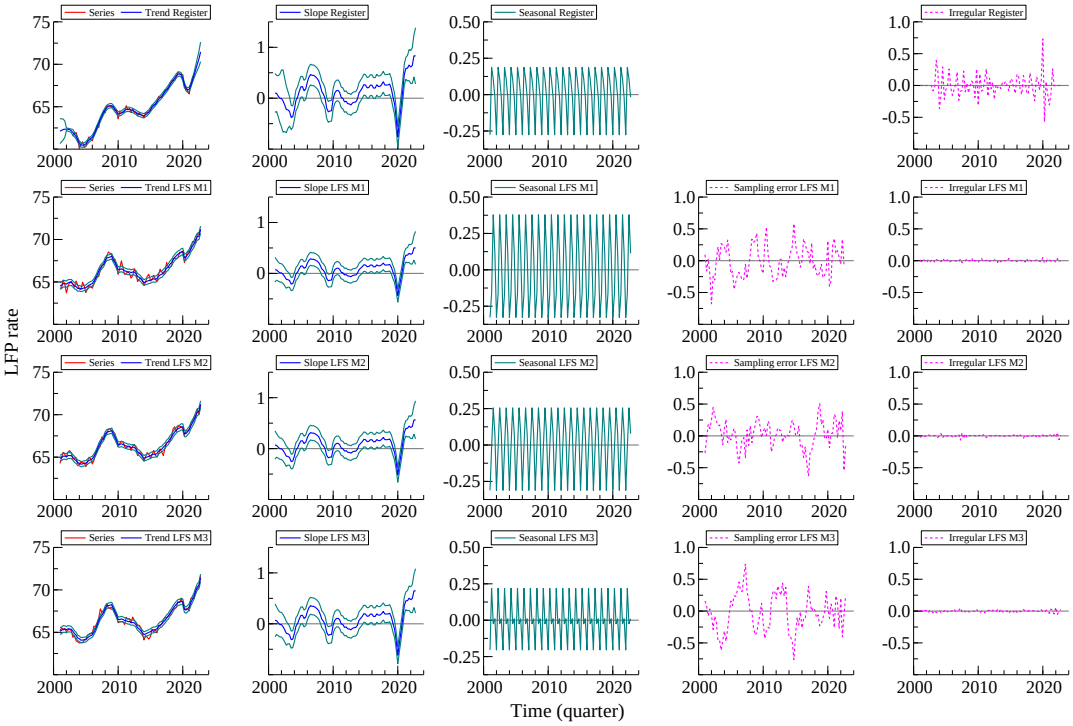

Figure C9: State vector estimates in Model 2, obtained with Kalman smoother.

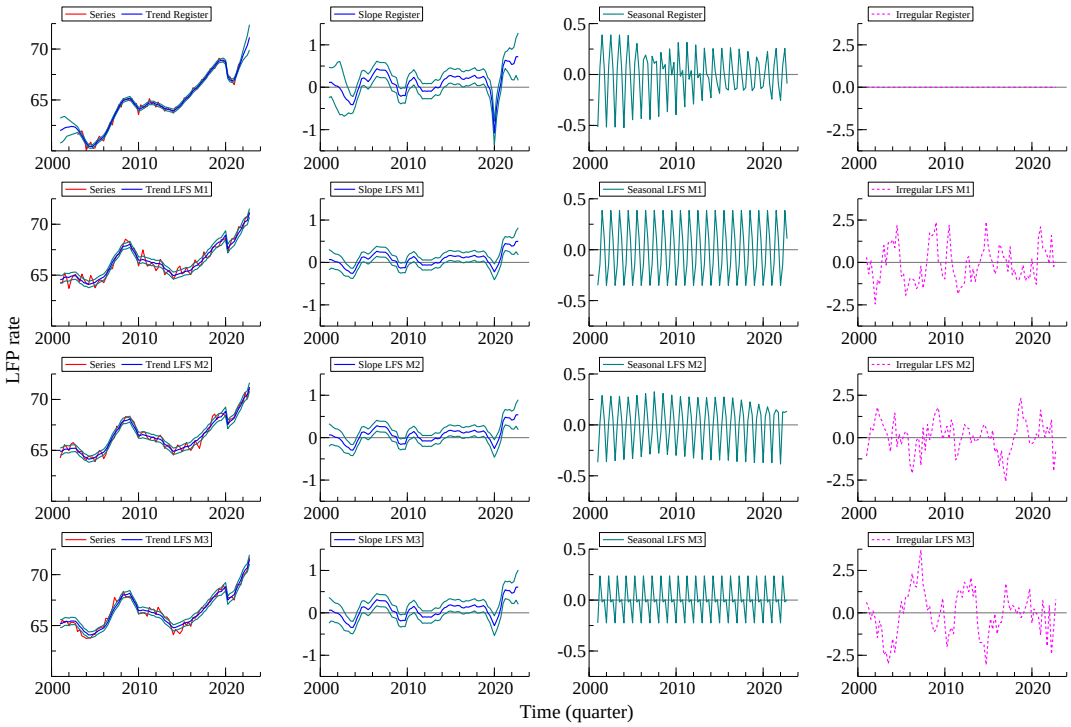

Figure C10: State vector estimates in Model 2 including time-varying variances for the trend disturbance terms at the start of the Covid outbreak.

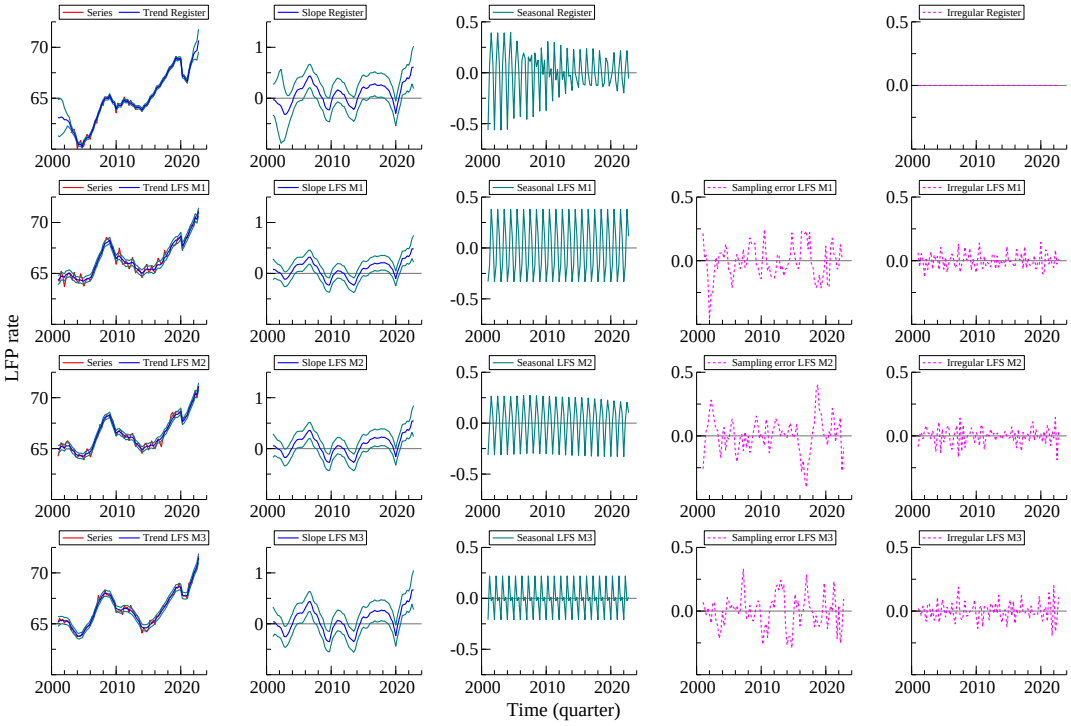

Figure C11: State vector estimates in Model 4, obtained with Kalman smoother.

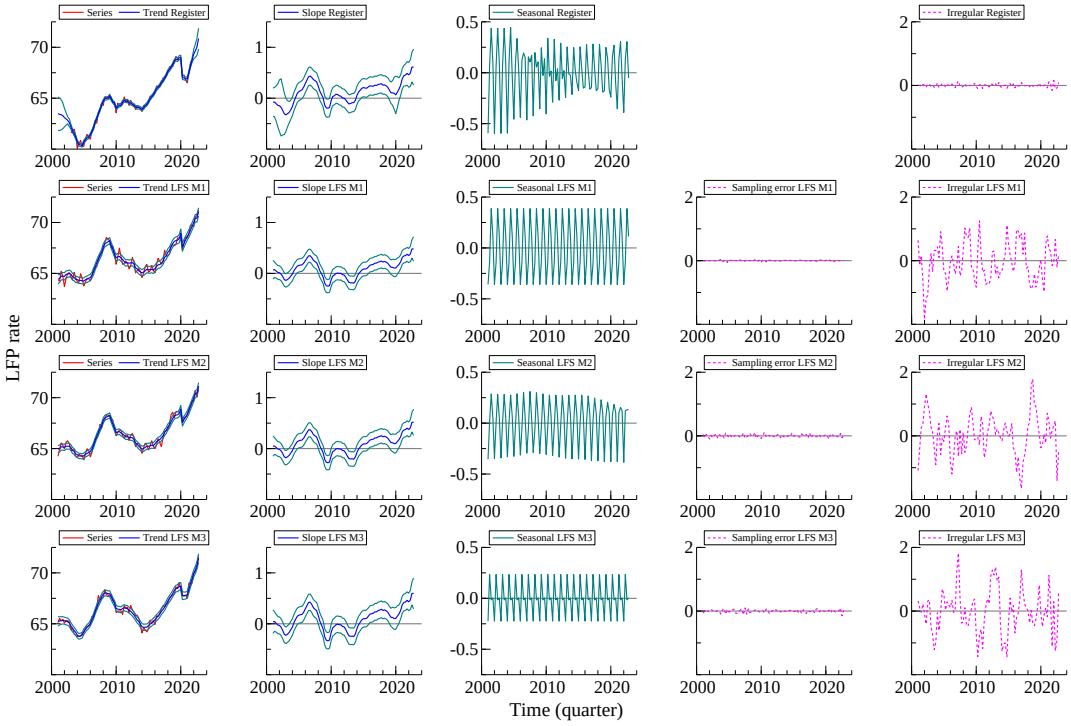

Figure C12: State vector estimates in Model 4 including time-varying variances for the trend disturbance terms at the start of the Covid outbreak.

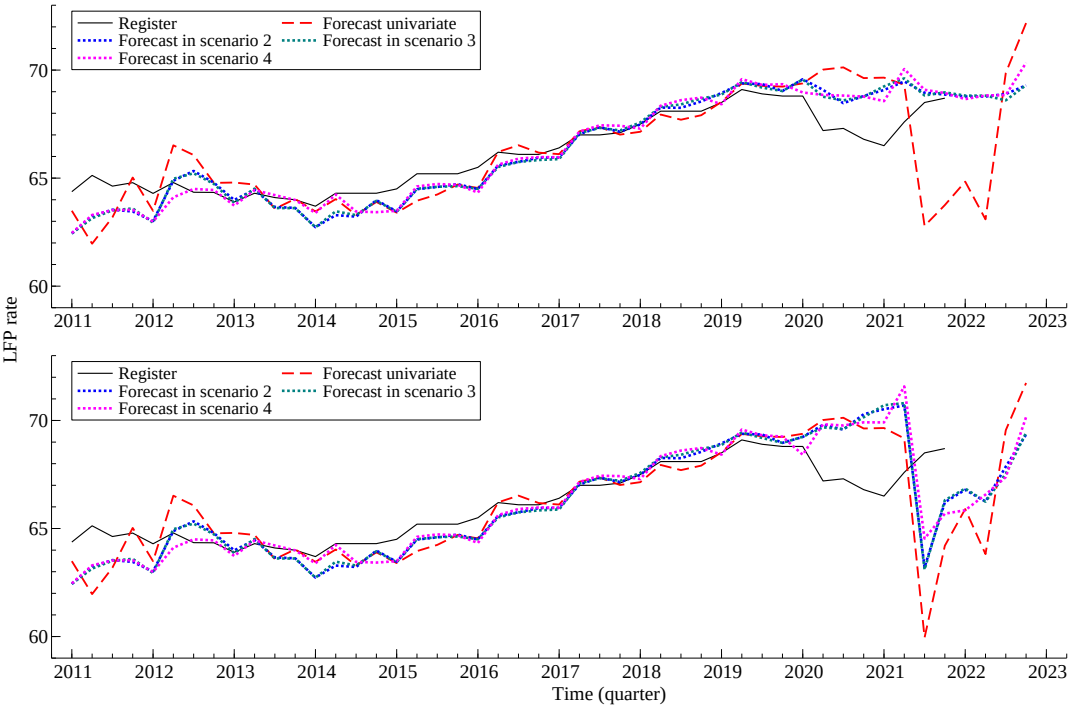

Figure C13: 5-step ahead forecasts of the register LPF rate series of Model 1, in the different forecast scenarios, constant (top) and time-varying (bottom) variances for trend disturbance terms.

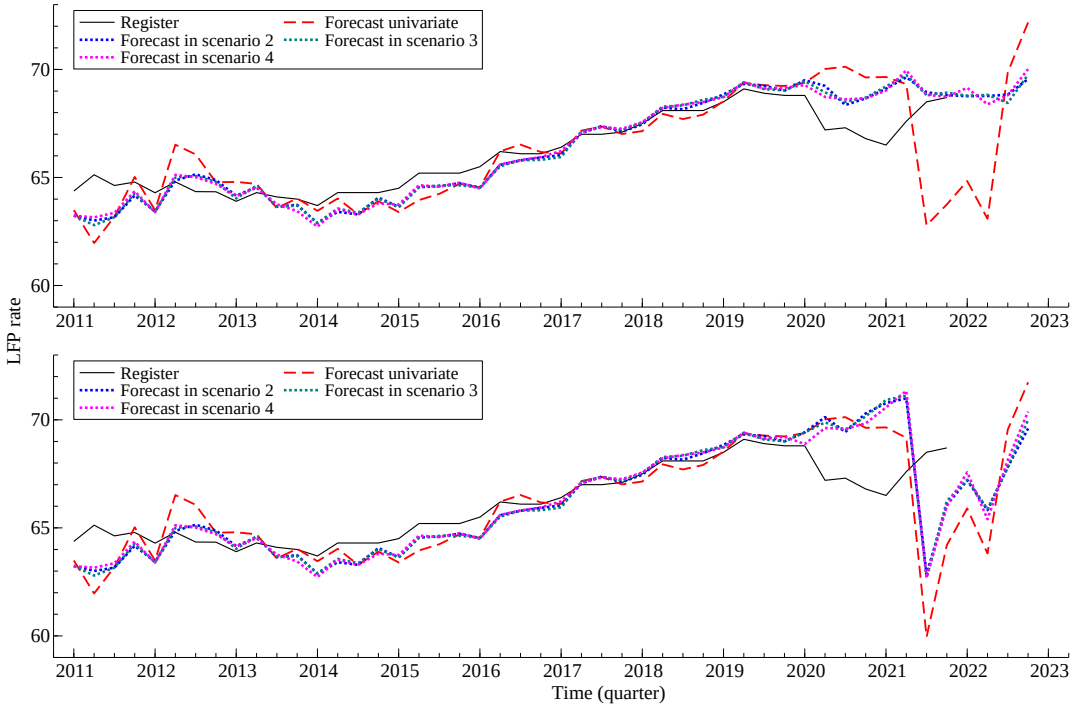

Figure C14: 5-step ahead forecasts of the register LPF rate series of Model 2, in the different forecast scenarios, constant (top) and time-varying (bottom) variances for trend disturbance terms.

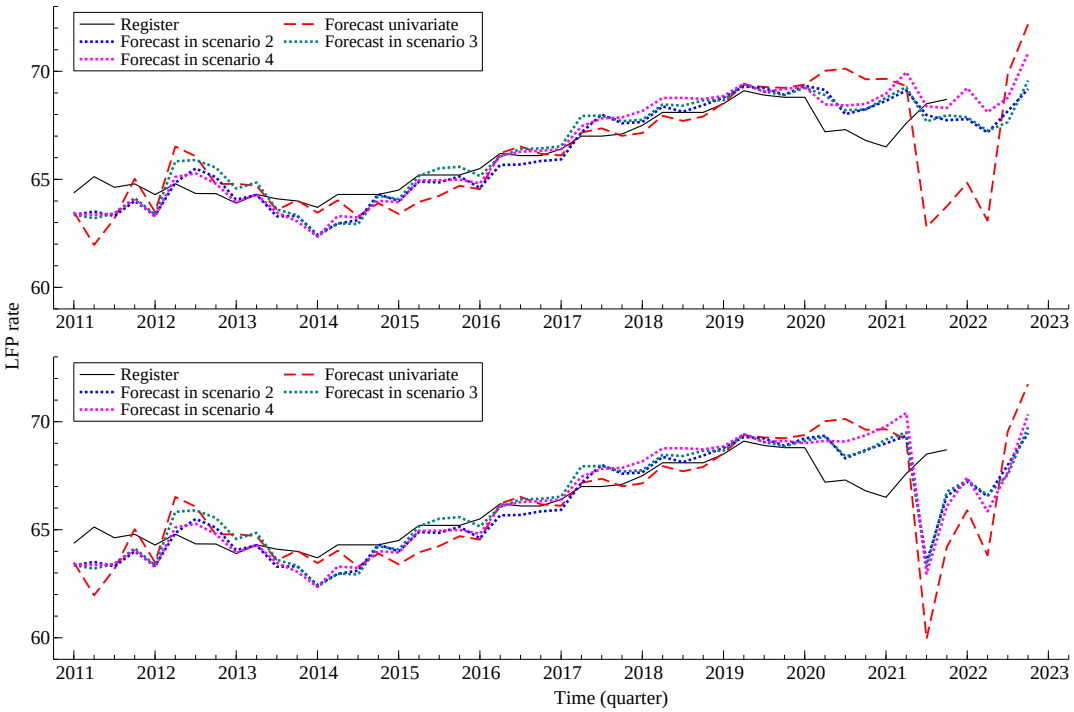

Figure C15: 5-step ahead forecasts of the register LPF rate series of Model 4, in the different forecast scenarios, constant (top) and time-varying (bottom) variances for trend disturbance terms.

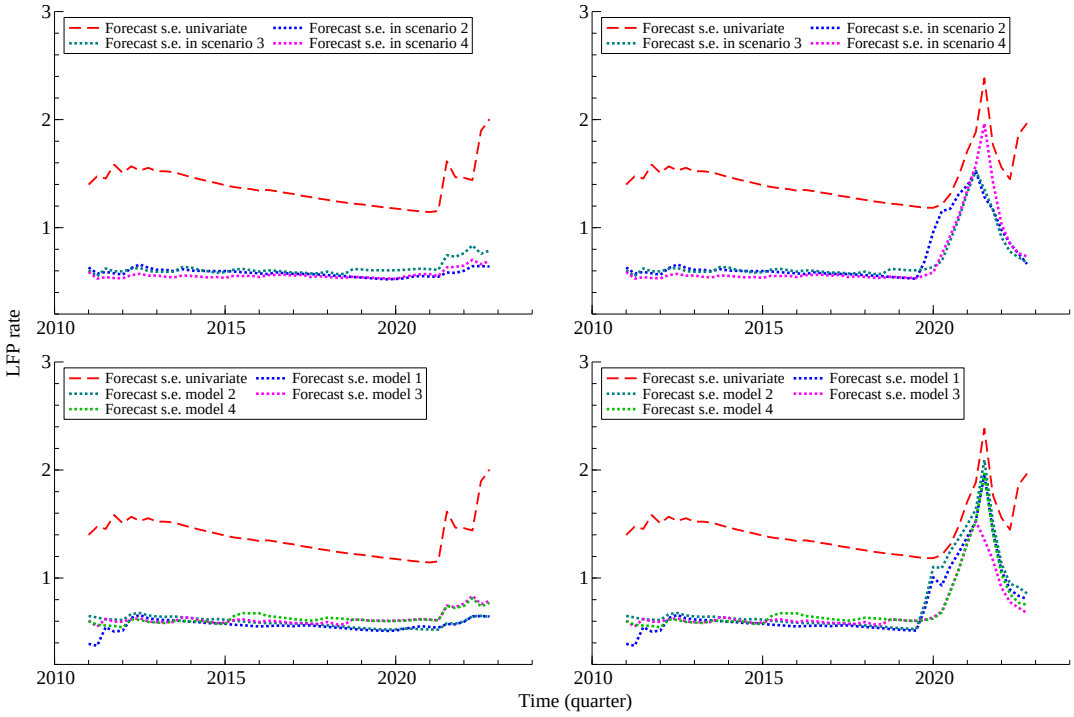

Figure C16: Forecast standard errors comparison from model 3 in the different scenarios and of the different models in scenario 3, with constant (left) and time-varying (right) variances for trend disturbance terms.

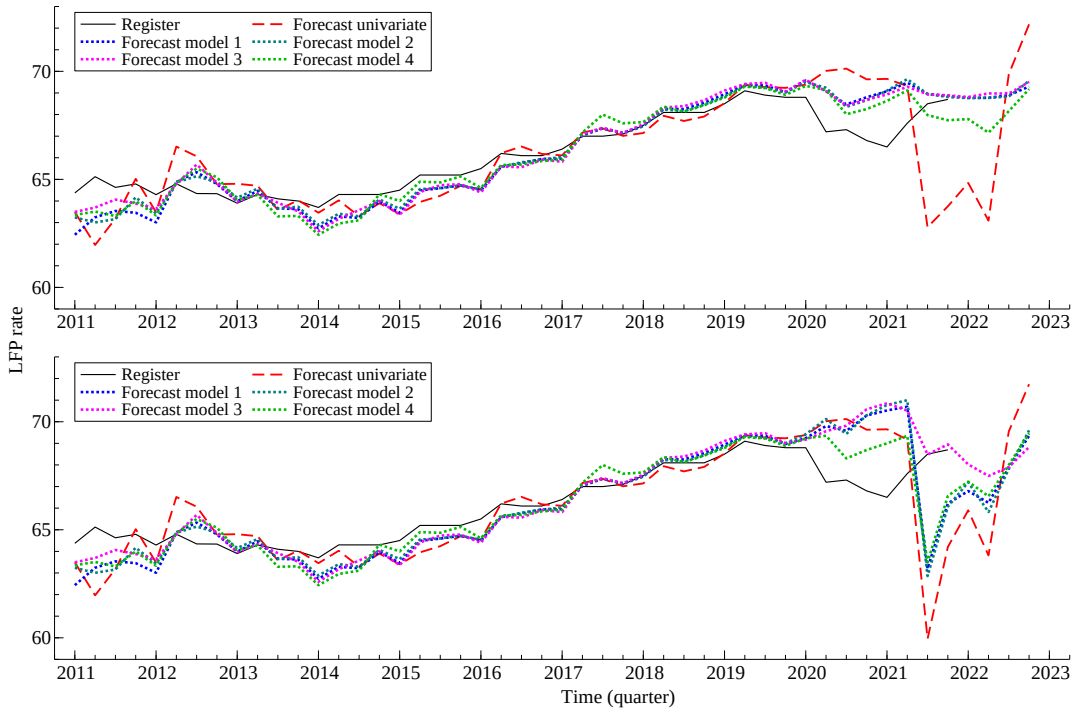

Figure C17: 5-step ahead forecasts of the register LPF rate series of the different models, in forecasting scenario 2, constant (top) and time-varying (bottom) variances for trend disturbances.

Figure C18: 5-step ahead forecasts of the register LPF rate series of the different models, in forecasting scenario 4, constant (top) and time-varying (bottom) variances for trend disturbances.