Synchronized Time Series Models for National Accounts Data

Lucas Harlaar[1,2,3*] Jan van den Brakel[1,2] Siem Jan Koopman[3,4]

[1]Department of Quantitative Economics, Maastricht University, 6211 LM Maastricht, Netherlands
[2]Department of Research & Development, Statistics Netherlands, 6412 EX Heerlen, Netherlands
[3]Department of Econometrics, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, Netherlands
[4]Tinbergen Institute, 1018 WB Amsterdam, Netherlands

Abstract

We develop a multivariate time series model for the complete national accounts dataset which typically consists of at least nine variables. The dynamic features of the variables are characterized by a set of unobserved components that evolve stochastically over time and may represent trend, cyclical, and seasonal effects. The components can be uniquely associated with a single variable or can be common to multiple variables. The model treats expenditure and production variables jointly, subject to accounting-identity constraints, ensuring that their gross domestic product aggregates are consistent in both mean and variance by construction. The model constraints can be treated naturally within the Kalman filter. Parameter estimation is based on exact maximum likelihood which is feasible despite the high-dimensional parameter vector. This model-based framework allows for synchronized nowcasting, forecasting, seasonal adjustment, trend and cycle extraction for national accounts variables, including gross domestic product. We show that the statistical precision of signal extraction and forecasting improves when our framework is used instead of univariate and unconstrained model alternatives. When the model includes a common cycle component for all national accounts variables, it can be interpreted as a macroeconomic cycle indicator at current prices. We empirically illustrate our model-based methodology for data from national accounts of Germany, Italy, and the Netherlands, and show its significance for both official statistics and macroeconomic policy-making.

Keywords: Unobserved components, constrained state space model, linear restrictions, Kalman filter, national accounts, macroeconomic cycles, official statistics

*Corresponding author: Lucas Harlaar, Department of Quantitative Economics, Maastricht University, Tongersestraat 53, 6211 LM Maastricht, Netherlands. Email: lucas.harlaar@maastrichtuniversity.nl

1 Introduction

Gross domestic product (GDP) is constructed from a set of coherently and consistently aggregated macroeconomic indicators of a country, which is known as the *national accounts.* National statistical institutes (NSIs) are subject to international accounting standards when they produce national accounts; see United Nations et al. (2009) and Eurostat (2013b, 2014). We consider two sides of national accounts: production and expenditure; both sides add up to GDP[1]. On the production side, GDP is the sum of the added value of all productive sectors in the economy, to which taxes are added and subsidies on products are subtracted. On the expenditure side, GDP is the sum of all payments on domestic goods and services. More specifically, it is the sum of final consumption, gross capital formation, and net exports. In the past, small statistical discrepancies could exist between the aggregates of the two different approaches or with GDP itself. This is a natural consequence, given that national accounts are compiled from many different data sources. Today, these definitions are strictly adhered to when the national accounts are evaluated at current prices (no inflation correction). The reason being that NSIs apply data adjustment procedures on source data such that they satisfy the theoretical national accounting identities; see, for example, van der Ploeg (1985) who proposed a combination of methods for this purpose. However, statistical discrepancies can still persist at constant prices (chain-linked volumes) or when seasonally and calendar adjusted measures are calculated; see the discussion in Bikker et al. (2019).

The national accounts sheet is compiled by accumulating information from a variety of sources on the demand and supply sides of the economy, both primary and secondary sources, including sample estimates obtained from surveys among firms and households (primary) and statistical measures derived from registers and administrations (secondary). Since conducting sample surveys with large sample sizes is involved, the resulting statistical information only becomes available at low publication frequencies, such as per year or per quarter. In addition, statistical production processes are complex and time-consuming, often resulting in substantial publication lags. Aiming to lower the implications of these publication lags for policymakers, NSIs are making a trade-off between accuracy and timeliness of initial (flash) estimates, despite the inconvenience that such estimates are subject to revisions. This trade-off has led to many research contributions in the literature on developing model-based approaches to nowcasting and short-term forecasting of

[1]We do not consider the income approach to national accounts in our study.

macroeconomic time series variables. These contributions have ranged from the use of dynamic factor models, see e.g. Giannone et al. (2008); Modugno (2013); Proietti et al. (2021) and the recent discussion in Cascaldi-Garcia et al. (2024), use of mixed-data sampling (MIDAS) regression methods, see e.g. Andreou et al. (2013); Guérin and Marcellino (2013); Smith (2016), the use of mixed-frequency vector autoregressions (VARs), see e.g. Giannone et al. (2009); Kuzin et al. (2011); Hecq et al. (2022); Cimadomo et al. (2022)), the use of dynamic structural general equilibrium (DSGE) models, see e.g. Del Negro and Schorfheide (2013); Giannone et al. (2016), and the use of unobserved components time series models, see e.g. Harvey (2006); Scott and Varian (2014); Kohns and Bhattacharjee (2023); Harlaar et al. (2025)). Monitoring the current state of the economy using an econometric or statistical model has emerged and developed extensively in the past twenty years; see e.g. Stundziene et al. (2024). The model-based approaches, including those listed above, have in common that they do not account for the consistency of establishing GDP using either the production route or the expenditure route. Hence, model-based approaches are not tailored to the needs of statistical production. No attempt is made to provide *consistent* and *synchronized* predictions of the current state of all macroeconomic time series variables present in national accounts simultaneously. However, it is acknowledged that nowcasting and publication of (flash) estimates in advance are challenges that need to be faced by official statistics; see Bell et al. (2024). We aim to bridge these challenges from both macroeconomic and official statistics perspectives.

To tackle these challenges, we adopt the class of unobserved component (UC) time series models (also called *structural time series models* in the literature) as advocated by Harvey (1989). This class of models provides basic descriptions of the dynamic features of individual time series variables, and it lends itself naturally to multivariate extensions. Furthermore, once our solutions are provided in detail for UC models as presented below, they also facilitate the understanding of how the solutions can be implemented in other model-based approaches such as those associated with dynamic factor, MIDAS, VAR, DSGE and other models, because they are closely aligned to UC models in their own ways. The multivariate UC model also provides an effective decomposition of dynamic features into components that can easily be interpreted. In particular, economic time series variables are assumed to be generated by long- and short-term dynamic components which can be associated with e.g. trend, cycle and seasonal effects; see Harvey (1989). The use of UC models in the multivariate modeling of economic time series has been fully explored in Harvey and

Koopman (1997). Azevedo et al. (2006) have considered a multivariate UC model to track the business cycle from a panel of economic indicators. Similarly, Aruoba et al. (2009) have extracted a single factor from a larger panel of economic indicators which they argue can serve as proxy of real activity of the economy. In more general dynamic factor models for economic variables (Azevedo et al., 2006; Giannone et al., 2008; Bańbura et al., 2013; Proietti et al., 2021), multiple factors are selected to obtain a good description of all dynamic features. When the number of factors increases, already starting from three and higher, it becomes more difficult to provide (economic) interpretations to factors; in the overview of Cascaldi-Garcia et al. (2024) it has been advocated to refrain from giving interpretations. However, Bikker et al. (2019) argue that it can be key to attach interpretations to factors to avoid discrepancies between aggregate series and the sum of their breakdown. In particular, adjusting for seasonal and calendar effects according to Eurostat standards (Eurostat, 2015) causes discrepancies between GDP and national accounts, which is where the use of a multivariate UC model can be beneficial for correcting for these discrepancies. For example, as an alternative to conventional seasonal-adjustment methods such as X12ARIMA, our proposed UC model can also produce seasonally adjusted figures that obey the aforementioned national accounts constraints. Consequently, UC and related models are increasingly adopted by (applied) researchers and practitioners at NSIs, central banks and economic policy units.

A particular challenge in official statistics is to address statistical discrepancies in national accounts, whether they are caused by measurement errors in source data, or by seasonal adjustment methods. Another source of statistical inconsistency is concerned with model-based predictions. More specifically, the aggregate of the nowcasts and out-of-sample forecasts of the individual variables in the national accounts are not necessarily consistent between GDP approaches. The aim of this study is to show that individual national account forecasts can be synchronized directly within a model-based approach such that their aggregates can naturally serve as a common consistent GDP forecast. This differs from the classical forecast reconciliation literature which requires a two-step procedure (see e.g. Athanasopoulos et al. (2009); Hyndman et al. (2011); Wickramasuriya et al. (2019)) and is more in line with a branch of literature that in Athanasopoulos et al. (2024, Section 3.9) is referred to as “in-built coherence”. The key to our solution is a particular implementation of the Kalman filter and related state space methods (Durbin and Koopman, 2012). We show that forecasts obtained from a multivariate UC model can be synchronized by introducing linear restrictions at the level of the unobserved components themselves. The work on enforcing linear restrictions in state space methods dates back to Doran (1992) who has advocated to treat linear restrictions in the same way as observations. More developments of this perspective on imposing linear restrictions are reported in Pandher (2002, 2007) and Pizzinga (2012). Our solution is closely related to enforcing a zero-sum constraint on seasonal effects (Harvey and Koopman, 1993; Proietti, 2000). The same zero sum device has recently been adopted for synchronizing level interventions (to enforce complete substitution within the panel of economic variables) in a multivariate UC model (Smith et al., 2024; Hungnes et al., 2024). We generalize the zero sum device further by having the entire state vector with all unobserved components, denoted by $\boldsymbol{\alpha}_t$, being subject to the linear restriction $\boldsymbol{W}\boldsymbol{\alpha}_t = \mathbf{0}$, where $\boldsymbol{W}$ is a pre-designed matrix and $\mathbf{0}$ is a vector of zeros. In this way, we can synchronize the individual components on both sides of national accounts (production and expenditure) and synchronize different linear combinations of the components, which are needed for different tasks such as trend extraction, seasonal adjustment, prediction, and nowcasting.

Combining different breakdowns of GDP in a single multivariate UC time series model has the advantage that both cross-sectional and temporal information are used simultaneously to obtain more precise estimates of the dynamic features (trend, seasonal, cycle). Incorporating these known restrictive structures for synchronization will lead to increased estimation accuracy as it enforces more interdependencies between variables and their components in the model. The main contribution of the paper is the development of a comprehensive tool that extracts dynamic components accurately from national accounts data, and provides synchronized predictions and projections for purposes such as de-trending, seasonal adjustment, nowcasting and forecasting. This makes it relevant to both economists at policy institutes and practitioners in official statistics. We provide empirical evidence of our model-based approach for the national accounts of three Euro area countries: Germany, Italy and the Netherlands.

The remainder of the paper is structured as follows. Section 2 introduces the multivariate UC model for national accounts and discusses how state space methods are modified to accomplish synchronization. Section 3 describes the data for the national accounts of three countries. Section 4 presents the empirical results and derives a macroeconomic cycle indicator at current prices. Section 5 shows the ability of our model to obtain consistent and synchronized predictions of aggregate GDP and constituent national accounts simultaneously. Section 6 concludes.

2 Multivariate unobserved components for national accounts data

2.1 Systems of National Accounts

The United Nations statistics division[2] defines global guidelines on national accounting called the System of National Accounts (SNA, 2008) on which Eurostat based its most recent European system of national and regional accounts rules, called ESA 2010 (Eurostat, 2013a). Eurostat considers the following breakdowns of GDP by expenditure and production approach given in Table 1.

Table 1: National accounts of expenditure and production approaches for Eurostat.

Label	Expenditure approach		Label	Production approach	
C	Consumption	$+$	A	Agriculture, forestry, fishing	$+$
GCF	Gross Capital Formation	$+$	$B-E$	Industry (except construction)	$+$
EX	Exports	$-$	F	Construction	$+$
IM	Imports	$=$	$G-I$	Wholesale and retail trade, transport, accommodation and food service activities	$+$
			J	Information and communication	$+$
			K	Financial and insurance activities	$+$
			L	Real estate activities	$+$
			$M-N$	Professional, scientific and technical activities; administrative and support service activities	$+$
			$O-Q$	Public administration, defense, education, human health and social work activities	$+$
			$R-U$	Arts, entertainment and recreation; other service activities; activities of household and extra-territorial organizations and bodies	$+$
			TLS	Product Taxes less Subsidies	$=$
GDP	GDP at current prices		GDP	GDP at current prices	

There are also alternative but related SNAs in use. For example, the one used by Statistics Netherlands as published on Statline, the official publication website of Statistics Netherlands. It contains the same national accounts categorization according to the Eurostat binding standards in Table 1. However, NSIs are free to publish their own alternative categories. This alternative depiction of the expenditure and production variables from Statline is presented in Table 2. In this paper, we consider these two alternative breakdowns of GDP to illustrate the flexibility and generality of our method.

[2] in collaboration with the World Bank, IMF, OECD, and European Commission, see United Nations et al. (2009)

Table 2: National accounts of expenditure and production approaches for Statistics Netherlands.

Label	Expenditure approach		Label	Production approach	
C	Consumption	$+$	AIG	Goods from Agriculture and Industry	$+$
I	Investment	$+$	CS	Commercial Services	$+$
IC	Inventory Change	$+$	NCS	Non-Commercial Services	$+$
EX	Exports	$-$	TLS	Product Taxes less Subsidies	$=$
IM	Imports	$=$			
GDP	GDP at current prices		GDP	GDP at current prices	

2.2 Model specification

Denote $n_E \times 1$ vector of time series variables of the expenditure approach by $\boldsymbol{y}_{E,t}$ and $n_P \times 1$ vector of time series variables of the production approach by $\boldsymbol{y}_{P,t}$, where the time index t refers to a quarterly frequency, with $t = 1, \ldots, T$, and define $N = n_E + n_P$ as the total number of time series variables in national accounts and with $N \times 1$ data vector $\boldsymbol{y}_t = (\boldsymbol{y}'_{E,t}, \boldsymbol{y}'_{P,t})'$. The unobserved components time series (UC) model for $y_{i,t}$, the ith variable of the observation vector $\boldsymbol{y}_t$ consists of trend $\mu_{i,t}$, seasonal $\gamma_{i,t}$ and cycle $\psi_{i,t}$ components, and can be written as

$$y_{i,t} = \mu_{i,t} + \gamma_{i,t} + \psi_{i,t} + \varepsilon_{i,t}, \qquad \varepsilon_{i,t} \sim \text{NID}\left(0, \sigma^2_{\varepsilon_i}\right), \quad i = 1, \ldots, n_E, n_E + 1 \ldots, N, \tag{1}$$

where irregular sequences $\{\varepsilon_{i,1}, \ldots, \varepsilon_{i,T}\}$ are normal, independent and identically distributed (NID), are not correlated between indices i, and have mean zero and variance $\sigma^2_{\varepsilon_i} > 0$, for $i = 1, \ldots, N$. The latent trend component $\mu_{i,t}$ in the model is able to capture non-stationary trending features in the time series by specifying it as the stochastic process

$$\mu_{i,t+1} = \mu_{i,t} + \nu_{i,t}, \qquad \nu_{i,t+1} = \nu_{i,t} + q_t\zeta_{i,t}, \qquad \zeta_{i,t} \sim \text{NID}\left(0, \sigma^2_{\zeta}\right), \tag{2}$$

as in Harvey (1989), where the increase $\nu_{i,t}$ in trend $\mu_{i,t+1}$ is specified as a random walk process with the NID trend innovation sequences $\{\zeta_{i,1}, \ldots, \zeta_{i,T}\}$ not being correlated between the indices i and between all irregular sequences. This trend specification can be referred to as the *smooth local linear trend* model or the *integrated random walk*. Trend innovations are scaled by a time-varying scalar q_t, which is a fixed function of time and accounts for sudden and sharp trend changes; for example, those due to unprecedented events such as the COVID-19 outbreak; see Section 2.7 for further discussion and specification of the trend. The quarterly observations for the variables in national accounts are not necessarily seasonally adjusted and therefore we include a seasonal component that is sufficiently flexible to capture the seasonal effects throughout the time series. In the same spirit as for the trend component, Harvey (1989) proposes a particular stochastically evolving seasonal process using a set of trigonometric equations. In particular, we model the seasonal component as

$$\gamma_{i,t} = \sum_{j=1}^{(s/2)} \gamma_{j,i,t}, \tag{3}$$

with

$$\begin{aligned} \gamma_{j,i,t+1} &= \cos(\lambda_j)\gamma_{j,i,t} + \sin(\lambda_j)\gamma^*_{j,i,t} + \omega_{j,i,t}, \qquad \lambda_j = 2\pi\, j/s, \\ \gamma^*_{j,i,t+1} &= -\sin(\lambda_j)\gamma_{j,i,t} + \cos(\lambda_j)\gamma^*_{j,i,t} + \omega^*_{j,i,t}, \qquad j = 1, \ldots, [s/2], \end{aligned} \tag{4}$$

where λ_j represents the seasonal frequency of $\gamma_{j,i,t}$, innovations $\omega_{j,i,t}$ and $\omega^*_{j,i,t}$ are mutually and serially independent NID $\left(0, \sigma^2_\omega\right)$ disturbances with the same time-invariant variance σ^2_ω, and $[s/2]$ equals $s/2$ for s even, and $(s-1)/2$ for s odd. For example, for a quarterly time series, we have $s = 4$, $\lambda_1 = 0.5\pi$ and $\lambda_2 = \pi$. Since $\sin(\pi) = 0$, the variable $\gamma^*_{2,i,t}$ can be removed from the model. Proietti (2000) discusses alternative specifications for the seasonal component $\gamma_{i,t}$.

Three cycle model specifications are considered. First, the UC model with *similar cycles*, introduced by Harvey and Koopman (1997), refers to $\psi_{i,t} = c_{i,t}$ with:

$$\begin{aligned} c_{i,t+1} &= \rho_c \cos(\lambda_c) c_{i,t} + \rho_c \sin(\lambda_c) c^*_{i,t} + \tilde{\omega}_{i,t}, \\ c^*_{i,t+1} &= -\rho_c \sin(\lambda_c) c_{i,t} + \rho_c \cos(\lambda_c) c^*_{i,t} + \tilde{\omega}^*_{i,t}, \end{aligned} \tag{5}$$

where $0 < \rho_c < 1$ is the damping factor, λ_c is the cyclical frequency, innovations $\tilde{\omega}_{i,t}$ and $\tilde{\omega}^*_{i,t}$ are mutually and serially independent NID $\left(0, \sigma^2_{\tilde{\omega}}\right)$ disturbances with time-invariant variance $\tilde{\omega}^2 > 0$, and the pairs $(\tilde{\omega}_{i,t}, \tilde{\omega}^*_{i,t})$ and $(\tilde{\omega}_{j,t}, \tilde{\omega}^*_{j,t})$ are uncorrelated for $i \neq j = 1, \ldots, N$. The periodicity of the cycle is given by $2\pi/\lambda_c$. Hence, the cycles are all independent of each other, but they share the same coefficients ρ_c, λ_c and $\sigma^2_{\tilde{\omega}}$, and, therefore, the same dynamic properties. The second cycle model specification has a *common cycle* shared by all national accounts. We have $\psi_{i,t} = \vartheta_i c_t$, where $\vartheta_1 = 1$ and $\vartheta_2, \ldots, \vartheta_N$ are loadings that we can treat as unknown model coefficients, implying that $c_t \equiv c_{1,t}$. The loadings allow each series to exhibit the common cycle with a distinct amplitude. The common cycle c_t has the same dynamic stochastic updating structure as in (5). When the cycle period and damping factor estimates fall within admissible regions, c_t can be interpreted as the business cycle: when $6 \leq 2\pi/\hat{\lambda}_c \leq 48$ for $s = 4$ (quarterly observations) and $0 < \hat{\rho}_c < 1$. Finally, our third specification of the UC model is without a cyclical component and is obtained simply by having $\psi_{i,t} = 0$. This model only contains trend, seasonal and irregular components, and can be referred to as the *basic structural model* (Harvey, 1989).

2.3 State space formulation

Each UC model specification can be represented as a linear Gaussian multivariate state space model for the $N \times 1$ observation vector $\boldsymbol{y}_t = (\boldsymbol{y}'_{E,t}, \boldsymbol{y}'_{P,t})'$ given by the measurement and transition equations

$$\boldsymbol{y}_t = \boldsymbol{Z}\,\boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\alpha}_{t+1} = \boldsymbol{T}\,\boldsymbol{\alpha}_t + \boldsymbol{R}\,\boldsymbol{\eta}_t, \tag{6}$$

respectively, with $m \times 1$ state vector $\boldsymbol{\alpha}_t$ containing the trend, seasonal, and cycle elements present in the model (1), $N \times 1$ vector $\boldsymbol{\varepsilon}_t$ containing irregular elements $\varepsilon_{i,t}$ for $i = 1, \ldots, N$, $g \times 1$ innovation vector $\boldsymbol{\eta}_t$ containing the transition innovations of equations (2), (4) and (5). In particular, we have $\boldsymbol{\varepsilon}_t \sim \text{NID}(\boldsymbol{0}_N, \boldsymbol{H})$ and $\boldsymbol{\eta}_t \sim \text{NID}(\boldsymbol{0}_g, \boldsymbol{Q})$ where $\boldsymbol{0}_n$ is an $n \times 1$ vector of zeros, and $\boldsymbol{H}$ and $\boldsymbol{Q}$ are variance matrices of dimensions $N \times N$ and $g \times g$, respectively. The system matrices $\boldsymbol{Z}$, $\boldsymbol{T}$ and $\boldsymbol{R}$ can be regarded as selection matrices and consist only of unity and zero values, except that the matrix $\boldsymbol{T}$ may also contain cos() and sin() functions to treat seasonal and cyclical transitions.

As an illustration, take model (1) for quarterly national accounts time series, $s = 4$, with a common cycle. Then $y_{i,t}$ is decomposed by trend $\mu_{i,t}$, seasonal $\gamma_{1,i,t} + \gamma_{2,i,t}$ and common cycle $\psi_{i,t} = \vartheta_i c_t$ with the state space representation (6) given by

$$\boldsymbol{\alpha}_t = \left(\boldsymbol{\mu}'_t, \boldsymbol{\nu}'_t, \boldsymbol{\gamma}'_{1,t}, \boldsymbol{\gamma}^{*\prime}_{1,t}, \boldsymbol{\gamma}'_{2,t}, c_t, c^*_t\right)', \qquad \boldsymbol{\eta}_t = \left(\boldsymbol{\zeta}'_t, \boldsymbol{\omega}'_{1,t}, \boldsymbol{\omega}^{*\prime}_{1,t}, \boldsymbol{\omega}'_{2,t}, \tilde{\omega}_t, \tilde{\omega}^*_t\right)', \tag{7}$$

where $\boldsymbol{x}_t = (x_{1,t}, \ldots, x_{N,t})'$, for $x = \mu, \nu, \gamma_1, \gamma^*_1, \gamma_2, \zeta, \omega_1, \omega^*_1, \omega_2$, with system matrices

$$\boldsymbol{Z} = [\boldsymbol{Z}_\mu, \boldsymbol{Z}_\gamma, \boldsymbol{\vartheta}, \boldsymbol{0}_N], \quad \boldsymbol{T} = \text{diag}\left[\boldsymbol{T}_\mu, \boldsymbol{T}_\gamma, \rho_c \boldsymbol{C}(\lambda_c)\right], \quad \boldsymbol{R} = \text{diag}[(\boldsymbol{0}_4, \boldsymbol{I}_4)' \otimes \boldsymbol{I}_N, \boldsymbol{I}_2], \tag{8}$$

and

$$\boldsymbol{Z}_\mu = (1, 0) \otimes \boldsymbol{I}_N, \qquad \boldsymbol{Z}_\gamma = (1, 0, 1) \otimes \boldsymbol{I}_N, \qquad \boldsymbol{\vartheta} = (\vartheta_1, \ldots, \vartheta_N)',$$
$$\boldsymbol{T}_\mu = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \otimes \boldsymbol{I}_N, \qquad \boldsymbol{T}_\gamma = \begin{bmatrix} \boldsymbol{C}(\lambda_1) & \boldsymbol{0}_2 \\ \boldsymbol{0}'_2 & -1 \end{bmatrix} \otimes \boldsymbol{I}_N, \qquad \boldsymbol{C}(\lambda) = \begin{bmatrix} \cos(\lambda) & \sin(\lambda) \\ -\sin(\lambda) & \cos(\lambda) \end{bmatrix}.$$

More details of this state space formulation for UC models are given in Appendix A.

2.4 Imposing linear restrictions

Recall that the aggregate variable is GDP. Let Imports observations be multiplied by -1 for ease of notation. By definition, the following relation between the national accounts in $\boldsymbol{y}_{E,t}$, $\boldsymbol{y}_{P,t}$, and their aggregate must hold $GDP_t \equiv \sum_{i=1}^{n_E} y_{E,i,t} \equiv \sum_{i=1}^{n_P} y_{P,i,t}$. In case of model (1), this translates to the following restriction on the unobserved components

$$\sum_{i=1}^{n_E}(\mu_{i,t} + \gamma_{i,t} + \psi_{i,t} + \varepsilon_{i,t}) \; - \sum_{i=n_E+1}^{N} (\mu_{i,t} + \gamma_{i,t} + \psi_{i,t} + \varepsilon_{i,t}) = 0. \tag{9}$$

A clear benefit of UC models is that the components are directly interpretable. This naturally allows for the restriction to be adjusted to accommodate various needs. For example, Bikker et al. (2019) highlight the importance of synchronized seasonally-adjusted (SA) national accounts and GDP forecasting for NSIs. Synchronization of each dynamic feature in the multiple time series of the national accounts leads to more compelling and interpretable results. We can simply update the restriction of (9) to accommodate the SA predictions by imposing the zero restriction separately for each component. Since our aim is to predict the signal subject to the zero-sum restriction, we consider the restrictions for the three components, trend, seasonal, and cycle,

$$\sum_{i=1}^{n_E}\mu_{i,t} \; - \sum_{i=n_E+1}^{N} \mu_{i,t} = 0, \qquad \sum_{i=1}^{n_E}\gamma_{i,t} \; - \sum_{i=n_E+1}^{N} \gamma_{i,t} = 0, \qquad \sum_{i=1}^{n_E}\psi_{i,t} \; - \sum_{i=n_E+1}^{N} \psi_{i,t} = 0. \tag{10}$$

State space model formulations can be altered to accommodate linear restrictions on the state vector. For example, a temporal zero-sum restriction is often required for the seasonal component in which the yearly sum of seasonal effects must be zero to ensure the identification of the trend component. Imposing cross-sectional zero-sum restrictions as in (10) require a projection of the covariance matrix of the state disturbances onto the zero-sum subspace, see e.g. Harvey and Koopman (1993), Proietti (2000), Hungnes et al. (2024) and Smith et al. (2024). In what follows, we generalize this constrained state space model approach to comply with the linear national accounts restrictions in (10). The zero-sum restrictions on the unobserved state vector can be written in the form of $\mathbf{1}_m'\boldsymbol{\alpha}_t = 0$, where $\mathbf{1}_m$ is the $m \times 1$ vector of unity values. This is a specific case of imposing the general set of linear restrictions

$$\boldsymbol{W}\boldsymbol{\alpha}_t = \mathbf{0}_r, \tag{11}$$

where the $r \times m$ matrix $\boldsymbol{W}$ is known and r is the number of linear restrictions. For our state space model (6)–(8), we can impose five zero-sum restrictions by having $\boldsymbol{W} = [\boldsymbol{I}_5 \otimes \boldsymbol{w}, \boldsymbol{0}_5, \boldsymbol{0}_5]$ with the $1 \times N$ vector $\boldsymbol{w} = (\boldsymbol{1}'_{n_E}, -\boldsymbol{1}'_{n_P})$. In this way, the matrix $\boldsymbol{W}$ projects all idiosyncratic trend and seasonal components into a zero-sum subspace so that total expenditure is balanced against total production. We will show that a model-based enforcement of this restriction can be implemented for each t through the variance matrix $\boldsymbol{Q}$. In addition, it requires restrictions on the initial state vector α_1 and, in case of a common cycle, on the cycle loading vector $\boldsymbol{\vartheta}$.

2.5 Enforcing the restrictions on each component

To impose restrictions (11), the sub-matrices in $\boldsymbol{Q}$ need to be of reduced rank, since the state space model loses degrees of freedom when imposing linear UC restrictions. More specifically, let $\boldsymbol{A}$ denote a matrix whose columns form a basis for the null space of $\boldsymbol{w}$, i.e. $\text{Col}(\boldsymbol{A}) = \text{Nul}(\boldsymbol{w})$. Since $\boldsymbol{w}$ is a $1 \times N$ non-zero row vector defining one linear constraint on any vector $\boldsymbol{x}$ such that $\boldsymbol{wx} = 0$, it leaves $N-1$ degrees of freedom. Therefore, its null space can be spanned by $N-1$ linearly independent column vectors $\boldsymbol{x}$. Hence, $\boldsymbol{A}$ has dimension $N \times (N-1)$. The covariance sub-matrices of the state disturbances are then defined as follows.

Trend One trend disturbance specification is considered; $\boldsymbol{Q}_\zeta = \boldsymbol{A}\boldsymbol{\Sigma}_\zeta\boldsymbol{A}'$, where $\boldsymbol{\Sigma}$ reflects a $(N-1) \times (N-1)$ *scalar* variance matrix. This variance matrix structure consists of the identity matrix scaled by a non-negative value, i.e. $\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{I}_{(N-1)}$ where $\sigma^2 \geq 0$ is a scalar variance. Pre- and postmultiplying $\boldsymbol{\Sigma}_\zeta$ with $\boldsymbol{A}$ and $\boldsymbol{A}'$ respectively ensures that $\boldsymbol{w}\boldsymbol{Q}_\zeta\boldsymbol{w}' = 0$. This specification indicates that the long-term trends within the time series are modeled by a smooth trend model with correlated disturbances, both between and within the GDP approaches, which share a scalar variance. We notice that the correlations are merely a consequence of the zero-sum constraint.

Seasonal Furthermore, each series exhibits correlated short-term recurring seasonal effects with time-invariant stochastic disturbance variance that is shared across series and harmonics. Therefore, $\boldsymbol{Q}_\omega = \boldsymbol{I}_3 \otimes (\boldsymbol{A}\boldsymbol{\Sigma}_\omega\boldsymbol{A}')$. Again, correlations are solely induced by the restrictions. For this reason, we impose the restriction separately per harmonic, to avoid correlations between their elements. However, this is merely a modeling choice; imposing a single restriction on all seasonal state vector elements can be equally valid.

Cycle The common cycle model contains one long-term mean-reverting component with time-invariant disturbance variance $\sigma^2_{\tilde{\omega}}$. In state space form, this implies that $\boldsymbol{Q}_{\tilde{\omega}} = \sigma^2_{\tilde{\omega}}\boldsymbol{I}_2$. The common cycle is normalized based on the first time series in the observation vector, which is Consumption, i.e. the first element in $N \times 1$ loading vector $\boldsymbol{\vartheta}$ equals one. The zero-sum cycle restriction in this model specification is imposed on $\boldsymbol{\vartheta}$. The last element is set such that the restriction $\boldsymbol{w}\boldsymbol{\vartheta} = 0$ is enforced. All other loadings are freely estimated by maximum likelihood as part of parameter vector $\boldsymbol{\theta}$. In the similar cycle model, $\boldsymbol{Q}_{\tilde{\omega}} = \boldsymbol{I}_2 \otimes (\boldsymbol{A}\boldsymbol{\Sigma}_{\tilde{\omega}}\boldsymbol{A}')$, creating N 'similar' cycles since they share the damping factor, period and variance parameters. They are subject to the same constraint treatment as the slope and seasonal disturbances, where again both harmonics are restricted separately.

2.6 Enforcing the restrictions on the initial state vector

Finally, the initial state vector is assumed to be $\boldsymbol{\alpha}_1 \sim N(\boldsymbol{a}_1, \boldsymbol{P}_1)$. The Kalman filter is initialized with mean $\boldsymbol{a}_1$, subject to restriction $\boldsymbol{W}\boldsymbol{a}_1 = \boldsymbol{0}$, and with variance $\boldsymbol{P}_1$ given by

$$\boldsymbol{P}_1 = \mathrm{diag}\left[\kappa\boldsymbol{Q}_\zeta,\ \kappa\boldsymbol{Q}_\zeta,\ \kappa\boldsymbol{Q}_\omega,\ \frac{1}{1-\rho_c^2}\boldsymbol{Q}_{\tilde{\omega}}\right], \tag{12}$$

where κ is the diffuse prior $\kappa \to \infty$. This initial state variance matrix implies a diffuse initialization for the level, slope, and seasonal components which are still subject to the deterministic zero-sum restrictions. To facilitate implementing the constraints and for all practical purposes as set out by Harvey (1989), we simply set a numerical value for κ, that is, $\kappa = 10^6$. The cycle components are stationary and are therefore initialized with finite variance matrices which are also subject to the restrictions. In particular, this specification implies $\boldsymbol{W}\boldsymbol{a}_1 = \boldsymbol{0}$ and $\boldsymbol{W}\boldsymbol{P}_1\boldsymbol{W}' = \mathbf{O}$. The transition matrix (14) preserves the null space of $\boldsymbol{W}$ in all UC models that are considered in this paper. Hence, the mean and variance restrictions hold for all state vectors at $t = 2, \ldots, T$.

2.7 COVID treatment

In recent years, many suggestions have been proposed in the literature on how to deal with extreme fluctuations during the COVID-19 pandemic from March 2020 to the end of 2022; see, e.g. Lenza and Primiceri (2022), Schorfheide and Song (2021), Fernandes et al. (2023), Carriero et al. (2024), van den Brakel et al. (2022) and Bobeica and Hartwig (2023). For example, in Lenza and Primiceri (2022), the error term of a vector autoregressive model is scaled upward during the COVID-19 period to account for the large macroeconomic shocks. For the UC models in this study, we suggest scaling the trend innovation term of the slope $\zeta_{i,t}$ during periods affected by COVID-19. We have $\mathbb{V}\text{ar}(q_t\zeta_{i,t}) = q_t^2\sigma_\zeta^2$, for $i = 1,\ldots,N$, where σ_ζ^2 is the slope disturbance variance and q_t^2 is a strictly positive time-varying scale. This specification is common to all innovation sequences $\zeta_{1,t},\ldots,\zeta_{N,t}$. As a result, we replace $\boldsymbol{Q}_\zeta$ by $\boldsymbol{Q}_{\zeta_t} = q_t^2\boldsymbol{Q}_\zeta$ and implicitly have $\boldsymbol{Q}_t$ rather than $\boldsymbol{Q}$, in Section 2.5. The factor q_t^2 equals 1 for $t = 1,\ldots,t^*-3$, where t^* denotes the start of the COVID-19 outbreak in the first quarter of 2020. It is assumed that the COVID-19 outbreak lasted three quarters and we have $q_{t^*-2}^2 = q_{t^*-1}^2 = q_{t^*}^2 = \bar{q}$ where $\bar{q} \geq 1$. In the last stretch of the COVID-19 pandemic, we have $q_{t^*+h}^2 = 1 + (\bar{q}-1)\rho^h$, for $h = 1,\ldots,T-t^*$ and where $0 \leq \rho \leq 1$, to reflect that the impact of COVID-19 faded over some period of time. In this specification, the time-varying scale factor q_t^2 decays at a rate ρ after the outbreak. It follows from the trend specification in equations (1) and (2) that this re-scaling of the slope disturbance variance has a delayed effect of two time periods on $y_{i,t}$. Therefore, and contrary to Lenza and Primiceri (2022), the scaling is started before time t^*, the start of the outbreak. Both $\bar{q}$ and ρ can be treated as additional unknown parameters that need to be estimated or can be determined a priori based on external information. This trend specification embeds a time-invariant variance for slope disturbances by having $\bar{q} = 1$ and $\rho = 0$. The modeling of long-term trends within the time series by a smooth trend model with time-varying and uncorrelated slope disturbances seems reasonable for national account time series. In this way, the dynamic progression of the trend is only exhibited through stochastic changes in the slope component. It enforces the characteristic smooth evolution of trend changes present in macroeconomic time series, while being flexible enough to capture the disruptions at the start of 2020. Local level and local linear trend specifications, where the level component also has a stochastic disturbance term, have been considered as well. However, these more flexible specifications are not considered further, as they tend to take up cyclical dynamics as well.

2.8 Parameter estimation

Parameter estimation and signal extraction are carried out by software developed in Ox 9 (Doornik, 2021) and with use of subroutines in the SsfPack library for linear state space modeling (Koopman et al., 2008). Visualizations are created with the ggplot2 library (Wickham, 2016) in R (R Core Team, 2024). We maximize the log-likelihood function to obtain an estimate of the parameter vector $\boldsymbol{\theta}$ using the BFGS numerical optimizer. The parameter vector includes all variance, COVID-19 and, when applicable, additional cycle coefficients. In our empirical study, we show that it is computationally feasible to obtain the exact maximum likelihood estimates in this way despite the high dimensional state and parameter vectors. It is well documented that standard numerical optimization routines may fail to locate the global maximum of the log-likelihood function. This becomes more of an issue when the complexity of the model increases, the size of $\boldsymbol{\theta}$ increases, or constraints are added. In order to control this numerical imperfection to some extent, we suggest to fit the model without constraints first. These resulting estimates are then used as starting values to make the search for the parameters of the more comprehensive constrained model more efficient.

Other estimation strategies that could help the parameter search for multivariate UC models including trend and cyclical components can be found in e.g. van den Brakel and Krieg (2016), Rünstler and Vlekke (2018) and de Winter et al. (2022). Rünstler and Vlekke (2018) suggest to initially fit univariate models to each series in order to obtain preliminary estimates of key parameters. De Winter et al. (2022) argue it is not feasible to estimate all parameters directly. They develop a four-step strategy based on pre-set signal-to-noise ratios for variance parameters and iteratively fixing and re-estimating parts of the parameter vector. This shows similarities with the suggestion of van den Brakel and Krieg (2016) to run the numerical optimization procedure for a subset of parameters while the others are kept fixed and iterate this approach until optimal estimates are obtained.

3 Data description

The proposed constrained UC model is applied to national accounts data of Germany, Italy and the Netherlands. Due to the presence of linear (additive) restrictions on unobserved components in the model, we keep the data in levels despite it being common to take logs in macroeconomic data for empirical studies. Furthermore, we exclude Imports from the expenditure approach to GDP; for further details, see Tables 1 and 2. In our subsequent empirical application, when imposing the UC restrictions, we simply treat Imports as part of the production approach instead. This is mathematically equivalent to multiplying all Import observations by -1, which is done in Sections 2.4 and 2.3 for notational purposes only. It preserves the original interpretation.

3.1 Germany & Italy

To conduct our analysis, national accounts data of the expenditure and production approaches to GDP are retrieved from the OECD data explorer for Germany and Italy. As a result, $N = 15$ since $\boldsymbol{y}_{E,t}$ and $\boldsymbol{y}_{P,t}$ are 4×1 and 11×1 respectively, following from Table 1. Germany's observation window is 1991Q1-2023Q4, while Italy's sample is starting four years later, hence that sample ranges from 1995Q1-2023Q4. All their national accounts are observed and published on a quarterly frequency. Therefore, $T = 132$ for Germany and $T = 116$ for Italy. No outliers were removed for either country. That results in 1,980 and 1,740 data points respectively, on which we fit our proposed UC models. All observations are re-scaled to be expressed in 10 billion euros, to improve the stability of the numerical optimizer that was used to maximize the log-likelihood function. The more common log-transformation is not applicable as was mentioned above. To account for the excessive COVID-19 shocks, we model the trend according to the discussion in Sections 2.2 and 2.7, where both q_t^2 and ρ are estimated by maximum likelihood.

3.2 The Netherlands

To show the flexibility and generalizability of our method, the expenditure and production breakdowns of GDP in the Netherlands, as shown in Table 2, are retrieved from the official database of Statistics Netherlands. This implies that (1) now reflects a 9-dimensional model, as $\boldsymbol{y}_{E,t}$ and $\boldsymbol{y}_{P,t}$ are 5×1 and 4×1, respectively, using the alternative definitions of Statline. The dimensions in (2)-(10) also change accordingly. Similar to Italy, observations of the Dutch national accounts start with the first quarter of 1995. Consequently, its sample ranges from 1995Q1-2023Q4 implying that $T = 116$ as well. This provides a total of 1,044 data points, from which two outliers were removed after consultation with the national accounts team at Statistics Netherlands. The COVID-19 treatment and re-scaling of the observations are handled in a similar manner as for the German and Italian national accounts.

4 Empirical application

In this empirical study, we consider three constrained multivariate UC models for each of the national accounts data sets of Germany, Italy, and the Netherlands: model with similar cycles, model with a common cycle, and model without a cycle. We present the parameter estimates of the three models and discuss goodness-of-fit indicators. We further verify the adequacy of the models by presenting residual diagnostic test statistics and discuss the macroeconomic interpretation of the results.

4.1 Estimated parameters

The maximum likelihood estimates of the parameters in the three different multivariate UC models are presented in Table 3. In particular, Table 3 reports the estimates of the common scalar variances σ_ζ^2 and σ_ω^2, corresponding to the stochastic slope and seasonal state disturbances, respectively, together with the estimate of σ_c^2 which is the common scalar variance of the cycle states rather than the variance of its corresponding state disturbances. Since the cycle is assumed to be a stationary autoregressive component, we have $\sigma_c^2 = \sigma_{\tilde{\omega}}^2/(1-\rho_c^2)$. As we estimate σ_c^2, the variance of the cycle disturbance follows naturally as $\sigma_{\tilde{\omega}}^2 = (1-\rho_c^2)\sigma_c^2$, which implies that as ρ_c tends to unity, $\sigma_{\tilde{\omega}}^2$ tends to zero, and the cycle process $c_{i,t}$ reverts to an invariant sinusoidal pattern but is still stochastic, since the initial cycle state is random and well-defined, that is, $c_{i,1} \sim \mathrm{NID}(0, \sigma_c^2)$. This framework leads to a more stable numerical optimization process. Further, we place logarithmic transformations on most parameters to impose appropriate bounds on the parameter space. Table 3 reports the point estimates (after taking anti-logs) and 95% confidence intervals, which are asymmetric due to the anti-log transformations.

Moreover, Table 3 shows the estimated time-dependent variance factor of the slope disturbance variance around the start of the COVID-19 outbreak and corresponding rate of decay. At last, the cycle period (in quarters) and damping factor estimates are presented. When models contain no cyclical component, we observe an increased variance of the trend which indicates that part of the long-term oscillation in the series is now captured by the trend. However, the majority is (incorrectly) allocated to noise in the observations rather than signal, an indication that the models with a cycle result in a better model fit compared to the bsm. The corresponding estimated irregular variances are presented in Table 4. Furthermore, Figure 1 depicts the signal-to-noise ratios of the cycle and irregular for each series. For the similar cycle model it is defined as $\mathrm{SNR}_i = \sigma_{\tilde{\omega}}^2/\sigma_{\varepsilon_i}^2$, while for the common cycle model it is $\mathrm{SNR}_i = (\vartheta_i^2\sigma_{\tilde{\omega}}^2)/\sigma_{\varepsilon_i}^2$. As shown in Table 4, when irregular variances tend to zero, the corresponding signal-to-noise ratio goes to infinity and is therefore left out of the graph. Since the similar cycle model for Italy fails to capture cyclical patterns the

Table 3: Maximum likelihood estimates of model parameters across countries and constrained multivariate UC model types, with 95% confidence intervals in brackets.

	Germany			Italy			Netherlands		
	no cycle	similar	common	no cycle	similar	common	no cycle	similar	common
Variances state									
$\hat{\sigma}^2_\zeta$ $(\times 10^{-4})$	17.93 [14.44, 22.26]	5.00 [3.54, 7.08]	7.21 [5.97, 8.72]	2.24 [1.56, 3.21]	2.24 [1.59, 3.16]	1.39 [1.07, 1.81]	4.21 [3.04, 5.83]	0.33 [0.18, 0.61]	1.24 [0.85, 1.79]
$\hat{\sigma}^2_\omega$ $(\times 10^{-5})$	2.86 [1.52, 5.38]	1.50 [0.80, 2.84]	2.06 [1.21, 3.52]	1.19 [0.72, 1.98]	1.19 [0.74, 1.92]	1.46 [1.00, 2.13]	3.98 [2.91, 5.45]	4.24 [3.02, 5.96]	4.99 [3.30, 7.54]
$\hat{\sigma}^2_c$ $(\times 10^{-2})$		0.69 [0.41, 1.15]	8.86 [6.55, 11.99]		*	17.56 [11.34, 27.21]		3.17 [1.91, 5.26]	0.38 [0.22, 0.63]
Trend flexibility									
$\hat{\bar{q}}$	1329 [819, 2158]	4327 [2504, 7475]	710 [373, 1353]	4155 [2526, 6835]	4155 [2506, 6891]	1652 [947, 2882]	836 [481, 1456]	11066 [5125, 23893]	1605 [843, 3058]
$\hat{\rho}$	0.69 [0.64, 0.74]	0.70 [0.65, 0.75]	0.86 [0.75, 0.93]	0.66 [0.60, 0.71]	0.66 [0.60, 0.71]	0.80 [0.73, 0.86]	0.87 [0.79, 0.92]	0.87 [0.80, 0.92]	0.89 [0.80, 0.95]
Cycle properties									
$2\pi/\hat{\lambda}_c$		15.27 [12.44, 18.75]	19.30 [13.71, 27.18]		*	22.04 [16.56, 29.33]		25.40 [23.50, 27.44]	18.85 [11.15, 31.86]
$\hat{\rho}_c$		0.92 [0.88, 0.95]	0.84 [0.81, 0.87]		*	0.85 [0.81, 0.88]		0.98 [0.96, 0.98]	0.87 [0.82, 0.91]

NOTE: Logarithmic transformations of the parameters to place appropriate bounds on the parameter space lead to asymmetric confidence intervals. * For Italy, the similar cycle model has cycle parameters that tend to boundary conditions during estimation, essentially leading to the same model specification as the model without cycle.

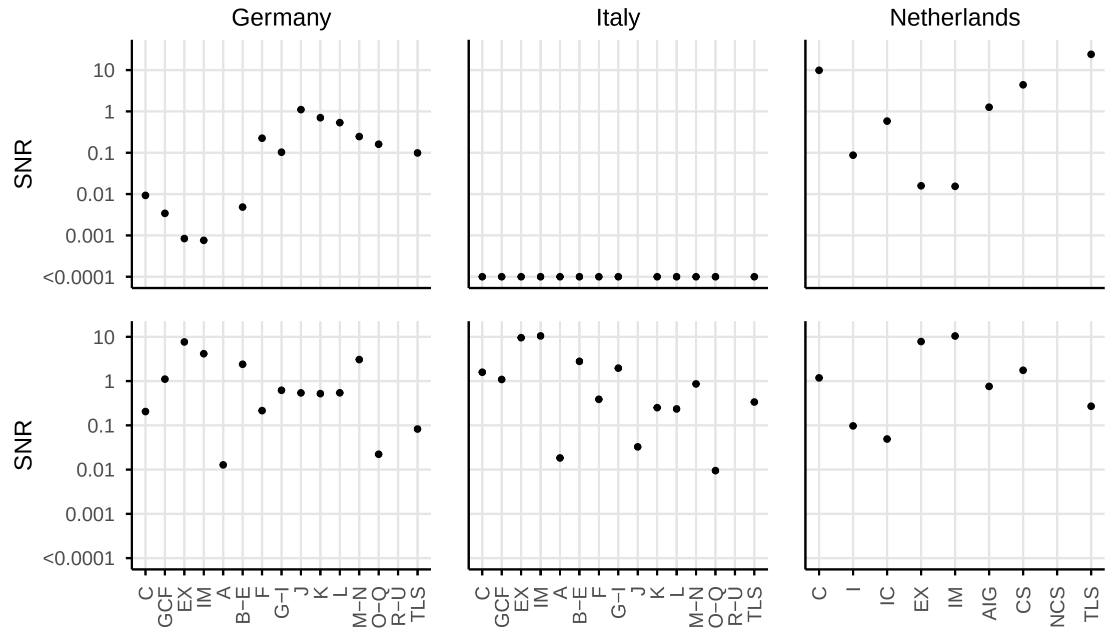

Figure 1: Signal-to-noise ratios (SNRs) of cycle and irregular components across series, countries and (cycle) model types. Results for the similar cycle (top row) and common cycle (bottom row) models are presented for countries Germany (left column), Italy (middle column), and the Netherlands (right column). Note that the y-axes are depicted on the $\log_{10}$ scale for visual purposes.

resulting signal-to-noise ratios become zero. These results support a multivariate UC model with trend, seasonal, cycle and noise components from a statistical perspective. Whether statistical evidence for a shared economic cycle can also be supported from a macroeconomic perspective is discussed below.

4.2 Goodness-of-fit measures

To verify whether the inclusion of similar cycles or a common cycle in the model leads to an improvement of the in-sample fit of the data, compared to a model without any cycle, we consider the maximized log-likelihood value, the corresponding Bayesian Information Criteria (BIC), the mean squared error (MSE) of the prediction residuals and the coefficient of determination R_d^2 where the total sum of squares is based on the data in differences. These goodness-of-fit statistics are presented in Table 5 for each country, and for the three model specifications.

The constrained multivariate UC model with common cycle fits the data the best for all countries in terms of the lowest values for BIC and MSE, and of the highest values for likelihood and R_d^2. These results provide some evidence that the national accounts data exhibit a shared underlying

Table 4: Maximum likelihood estimates of irregular variances across countries and constrained multivariate UC model types, with 95% confidence intervals in brackets.

	Germany			Italy			Netherlands		
	no cycle	similar	common	no cycle	similar	common	no cycle	similar	common
Variances irregular									
$\hat{\sigma}^2_{\varepsilon_1}$ $(\times 10^{-2})$	9.15	11.17	12.90	6.93	6.93	3.08	0.083	0.016	0.078
	[6.19, 13.53]	[7.76, 16.06]	[9.28, 17.92]	[4.81, 9.98]	[4.87, 9.85]	[2.22, 4.29]	[0.049, 0.139]	[0.0017, 0.14]	[0.039, 0.16]
$\hat{\sigma}^2_{\varepsilon_2}$ $(\times 10^{-1})$	2.70	3.03	1.40	1.17	1.17	0.51	0.20	0.18	0.18
	[1.68, 4.36]	[2.23, 4.14]	[0.99, 2.00]	[0.82, 1.67]	[0.87, 1.57]	[0.37, 0.70]	[0.14, 0.27]	[0.13, 0.25]	[0.11, 0.28]
$\hat{\sigma}^2_{\varepsilon_3}$ $(\times 10^{-3})$	1083	1236	103	239	239	17.69	4.04	2.65	3.74
	[757, 1550]	[937, 1630]	[69, 153]	[172, 332]	[172, 332]	[10.20, 30.69]	[2.79, 5.85]	[1.52, 4.63]	[2.02, 6.92]
$\hat{\sigma}^2_{\varepsilon_4}$ $(\times 10^{-1})$	11.69	13.65	1.55	3.76	3.76	0.21	1.53	0.97	0.12
	[8.10, 16.88]	[10.41, 17.91]	[1.06, 2.26]	[2.78, 5.10]	[2.81, 5.05]	[0.12, 0.39]	[1.13, 2.08]	[0.71, 1.34]	[0.052, 0.27]
$\hat{\sigma}^2_{\varepsilon_5}$ $(\times 10^{-4})$	*	*	6.08	7.64	7.64	7.53	1539	1003	82
			[4.02, 9.20]	[4.76, 12.26]	[4.63, 12.60]	[4.57, 12.40]	[1145, 2067]	[729, 1380]	[47, 143]
$\hat{\sigma}^2_{\varepsilon_6}$ $(\times 10^{-3})$	188	214	51.49	33.57	33.57	8.55	2.38	1.22	2.05
	[125, 281]	[155, 295]	[32.83, 80.74]	[23.14, 48.71]	[23.86, 47.23]	[5.43, 13.48]	[1.60, 3.54]	[0.61, 2.46]	[1.26, 3.33]
$\hat{\sigma}^2_{\varepsilon_7}$ $(\times 10^{-3})$	4.30	4.61	4.56	1.34	1.34	1.14	1.64	0.35	1.27
	[2.03, 9.09]	[3.08, 6.91]	[3.21, 6.48]	[0.84, 2.13]	[0.86, 2.09]	[0.71, 1.83]	[1.06, 2.54]	[0.047, 2.60]	[0.61, 2.63]
$\hat{\sigma}^2_{\varepsilon_8}$ $(\times 10^{-3})$	9.54	10.06	9.32	6.84	6.84	2.74	*	*	*
	[5.22, 17.45]	[6.87, 14.74]	[4.99, 17.39]	[3.33, 14.02]	[3.32, 14.08]	[1.47, 5.13]			
$\hat{\sigma}^2_{\varepsilon_9}$ $(\times 10^{-4})$	10.83	9.38	11.63	*	*	0.73	6.12	0.64	6.13
	[4.34, 27.03]	[5.13, 17.15]	[7.51, 18.01]			[0.38, 1.42]	[3.74, 10.02]	[0.023, 18.13]	[1.83, 20.60]
$\hat{\sigma}^2_{\varepsilon_{10}}$ $(\times 10^{-3})$	1.70	1.47	3.39	2.95	2.95	2.82			
	[0.83, 3.47]	[0.81, 2.69]	[1.97, 5.82]	[1.84, 4.72]	[1.90, 4.58]	[1.92, 4.13]			
$\hat{\sigma}^2_{\varepsilon_{11}}$ $(\times 10^{-4})$	20.63	19.39	23.56	3.73	3.73	3.79			
	[10.20, 41.74]	[12.35, 30.43]	[14.32, 38.75]	[2.19, 6.36]	[2.22, 6.26]	[2.22, 6.47]			
$\hat{\sigma}^2_{\varepsilon_{12}}$ $(\times 10^{-3})$	4.18	4.20	2.39	1.48	1.48	1.21			
	[2.33, 7.49]	[2.71, 6.52]	[1.58, 3.62]	[0.90, 2.45]	[0.92, 2.38]	[0.73, 1.98]			
$\hat{\sigma}^2_{\varepsilon_{13}}$ $(\times 10^{-3})$	4.74	6.45	5.93	7.25	7.25	6.86			
	[2.69, 8.32]	[3.82, 10.89]	[3.99, 8.80]	[5.17, 10.16]	[5.19, 10.12]	[4.56, 10.33]			
$\hat{\sigma}^2_{\varepsilon_{14}}$	*	*	*	*	*	*			
$\hat{\sigma}^2_{\varepsilon_{15}}$ $(\times 10^{-2})$	0.98	1.05	1.07	1.50	1.50	0.96			
	[0.49, 1.95]	[0.73, 1.50]	[0.68, 1.69]	[0.99, 2.27]	[1.00, 2.24]	[0.66, 1.40]			

NOTE: Logarithmic transformations of the parameters to place appropriate bounds on the parameter space lead to asymmetric confidence intervals.
* For some of the series, the irregular component drops out as the estimated variance parameter tends to the zero lower-bound.

Table 5: In-sample fit measures for different model specifications (similar, common, and no cycle) estimated for Germany, Italy, and the Netherlands.

							In-sample fit measures			
Country	Model	N	T	p	m	d	log-likelihood	BIC	MSE	R_d^2
Germany	no cycle	15	132	19	75	75	-301.709	0.537	0.619	0.140
	similar	15	132	22	105	75	-268.112	0.510	0.631	0.161
	common	15	132	36	77	75	71.884	0.201	0.365	0.300
Italy	no cycle	15	116	19	75	75	569.197	-0.397	0.196	0.422
	similar	15	116	22	105	75	569.196	-0.390	0.196	0.422
	common	15	116	36	77	75	960.879	-0.807	0.118	0.594
Netherlands	no cycle	9	116	13	45	45	349.950	-0.406	0.124	0.308
	similar	9	116	16	63	45	398.342	-0.485	0.104	0.361
	common	9	116	24	47	45	603.610	-0.842	0.081	0.453

NOTE: The table reports the log-likelihood, Bayesian Information Criterion (BIC), mean squared error (MSE), and R^2 for time series data (R_d^2). Furthermore, N = no. of time series, T = no. of observations, $p = \dim(\boldsymbol{\theta})$, $m = \dim(\boldsymbol{\alpha}_t)$ and d = no. of diffuse states.

cyclical pattern that drives their movement over time. This is confirmed by the results for all fit measures presented in Table 5. A similar argument can be made that the models with similar cycles and without cycle fail to capture crucial dynamics in the data. The maximum likelihood estimates reported in Table 3 show that some of the cycle parameters tend to go to boundary conditions during the numerical optimization process. For example, the similar cycle model essentially fails to extract any cyclical dynamics from most variables of Italy.

4.3 Residual diagnostic checking

To assess whether the model assumptions, including constraints, are empirically valid, we compute diagnostic test statistics for normality, heteroscedasticity, and autocorrelation of standardized residuals (Durbin and Koopman, 2012, Section 2.12). The statistics with corresponding critical values (at $\alpha = 0.01$) are summarized in Figure 2. The critical value of the autocorrelation statistics (Ljung-Box) depends on the number of sample autocorrelations considered (in our study, 12) and on the number of model parameters in each model (see Table 5). Therefore, the critical values are slightly different among the three models.

Table 6 collects all violations of the null hypotheses, which is shown in relation to the total number of series, for each country and model type. The normality assumption of the error terms is clearly violated for the majority of series, regardless of the model type, for the national accounts

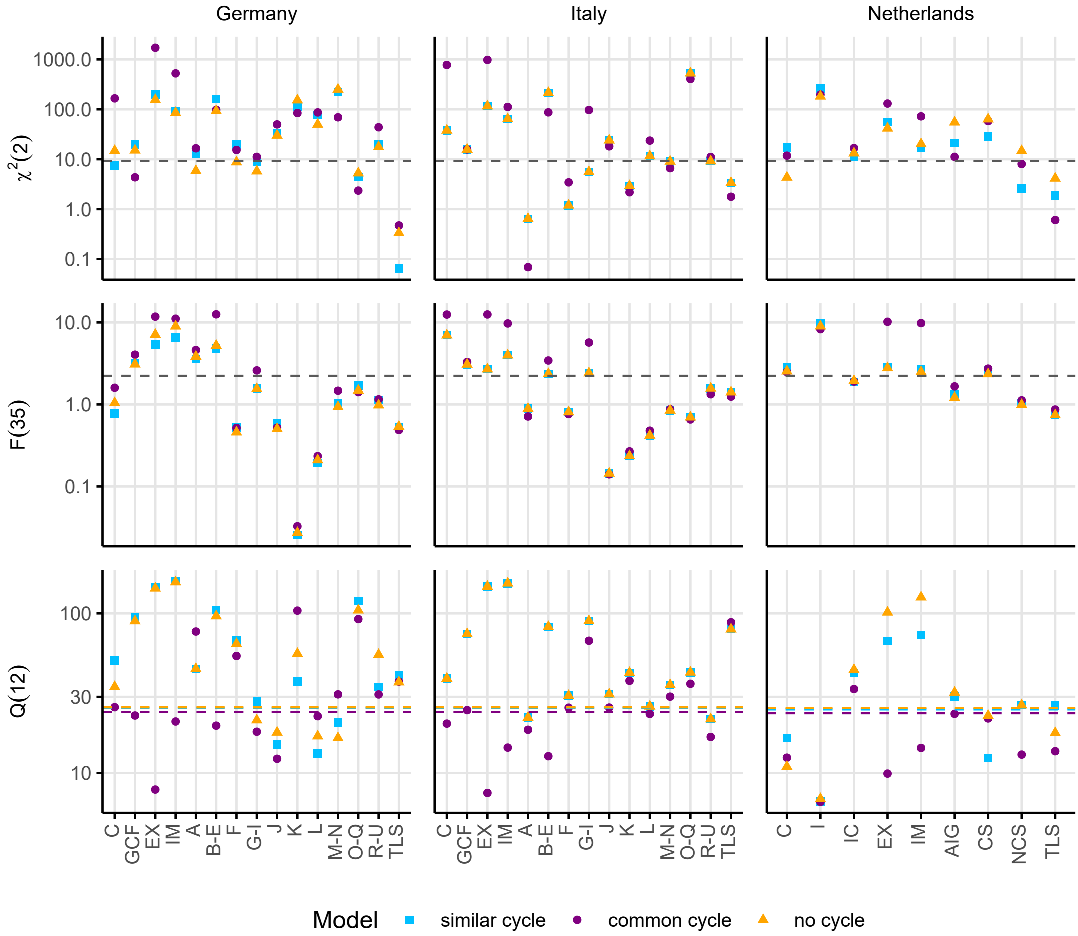

Figure 2: Residual diagnostics tests for normality (Bowman-Shenton, χ^2 with 2 degrees of freedom (df)), heteroscedasticity (F-test with 35 df) and serial correlation (Ljung-Box, 12 lags, χ^2 with df $= 12-(p/N)-1$, where $p = \dim(\boldsymbol{\theta})$ and $N = \dim(\boldsymbol{y}_t)$) of the standardized prediction errors of the individual variables in the three constrained multivariate UC models considered (similar, common, no cycle) for countries Germany (left column), Italy (middle column), and the Netherlands (right column). The y-axes show the values of the respective test statistics on the $\log_{10}$ scale. The dashed line depicts the critical value of the diagnostic test statistic, at $\alpha = 0.01$. The critical value for Ljung-Box changes slightly as it depends on number of model parameters per individual variable in the national accounts. The x-axes depict their indicators found in Table 1 and Table 2.

data of Germany and the Netherlands. For Italy, there are slightly less series for which the null hypothesis of normally distributed residuals can be rejected. In case it cannot be assumed that the prediction residuals are normally distributed, the Kalman filter and smoother remain to deliver minimum mean squared error estimates of the state vector (Harvey, 1989, Chapter 3). Consequently, this result does not raise fundamental concerns about the validity of inferences drawn from the empirical analysis. The heteroscedasticity test is the ratio of two sums of squared stan-

Table 6: Number of rejected diagnostic test statistics as depicted from Figure 2.

Country	Test	similar cycle	common cycle	no cycle	N
Germany	Normality	11 (73.3%)	12 (80.0%)	10 (66.7%)	15
	Heteroscedasticity	5 (33.3%)	6 (40.0%)	5 (33.3%)	15
	Ljung-Box	12 (80.0%)	8 (53.3%)	11 (73.3%)	15
Italy	Normality	8 (53.3%)	10 (66.7%)	8 (53.3%)	15
	Heteroscedasticity	6 (40.0%)	6 (40.0%)	6 (40.0%)	15
	Ljung-Box	13 (86.7%)	8 (53.3%)	13 (86.7%)	15
Netherlands	Normality	7 (77.8%)	7 (77.8%)	7 (77.8%)	9
	Heteroscedasticity	5 (55.6%)	5 (55.6%)	5 (55.6%)	9
	Ljung-Box	6 (66.7%)	1 (11.1%)	5 (55.6%)	9

dardized residuals, for two mutually exclusive subsets of 35 entries, taken from the beginning and the end of the sample period. Each subset comprises roughly a third of the entire sample. The null hypothesis of homoskedastic standardized residuals is rejected for roughly half of all series across all model types for the Netherlands. The results are generally better for the other two countries. It appears that the inclusion of the volatile period caused by the COVID-19 outbreak explains the homoscedasticity violations. As discussed in Section 2.7, the trend is temporarily made more flexible with the treatment decaying after a certain amount of time. However, the impact of the COVID-19 outbreak on national accounts varies significantly. Even with a special treatment of COVID-19, large residuals at the end of the sample period remain present for some of the series, violating the assumption of time-invariant variance of errors. The test to detect the presence of autocorrelation in the residuals indicates quite some violations of the null hypothesis of no autocorrelation. However, the common cycle model clearly outperforms the other two model specifications.

4.4 Signal extraction of common cycle component

The Kalman filter and smoother provide estimates of the common cycle from our preferred multivariate UC model, and are presented in Figure 3, for the three countries considered in this study. We notice that these estimated cycles have taken into account the national accounts constraints which includes the common cycle loading constraint: $\sum_{i=1}^{n_E} \hat{\vartheta}_i = \sum_{i=n_E+1}^{N} \hat{\vartheta}_i$; it ensures that the aggregated loadings are identical between the expenditure and production approaches. For comparisons, the smoothed cycle of a univariate UC model (smooth trend plus seasonal plus cycle plus irregular) for the GDP aggregate of the three countries is also displayed in each graph. We emphasize that the common cycle is normalized by the first series in $\boldsymbol{y}_t$, that is, Consumption, since we pre-set its loading to unity, $\vartheta_1 = 1$. The aggregated loadings, corresponding to the series of expenditure or production approaches, can be used to rescale the common cycle. In this way, we can have a balanced comparison of our common cycle estimate with the one extracted directly from the univariate model.

The graphs indicate that the cycles follow the fluctuations of their respective economies over the last decades. We notice that all extracted cycles are not based on the classical business cycle construction from logged GDP since the additive property of the NA constraints prohibited the log-transformation of the data. Hence, the observed cycle is extracted from the national accounts or GDP data at current prices, without taking logs or any other non-linear data transformations. The oscillation is therefore proportional to the level of the trend in the series. Since the trend clearly drifts upward over time for the majority of these macroeconomic series, the fluctuations of the cycle are somewhat larger at the end of the sample. Recall from Table 3 that the parameter estimates corresponding to the cycle component imply persistent cycle dynamics ($\hat{\rho}_c$ ranges from 0.84 to 0.87) and with typical business cycle period lengths ($2\pi/4\hat{\lambda}_c$ ranges from $18.85/4 = 4.7$ to $22.04/4 = 5.5$ years). The smoothed estimates of the cycle components presented in Figure 3 confirm the cyclical patterns that resemble the fluctuations in economic and business conditions. Hence, we may interpret the common cycle as an indicator of macroeconomic activity at current prices. Imposing the NA constraints in a non-linear multivariate setting on log-transformed national accounts data would be an interesting direction for further research. In this way, the extracted common cycle will have a more classical interpretation of the business cycle.

Figure 3 further demonstrates that the estimated cycle of the constrained multivariate national accounts model resembles the cycle extracted from GDP itself. The latter is the standard procedure for filtering/smoothing economic cycles in the literature. However, our results also suggest that a multivariate approach using the constituent components of GDP improves the precision of the estimated macroeconomic cycle indicator. The standard error decreases substantially for Germany and Italy, with notable improvements for the Netherlands as well. The uncertainty increases at the end of the sample. This is a consequence of temporarily increasing the variance of the slope disturbance to capture the volatile COVID-19 observations with an increased variance of the trend in the model. This amendment to the model also affects the estimation uncertainty of the cycle.

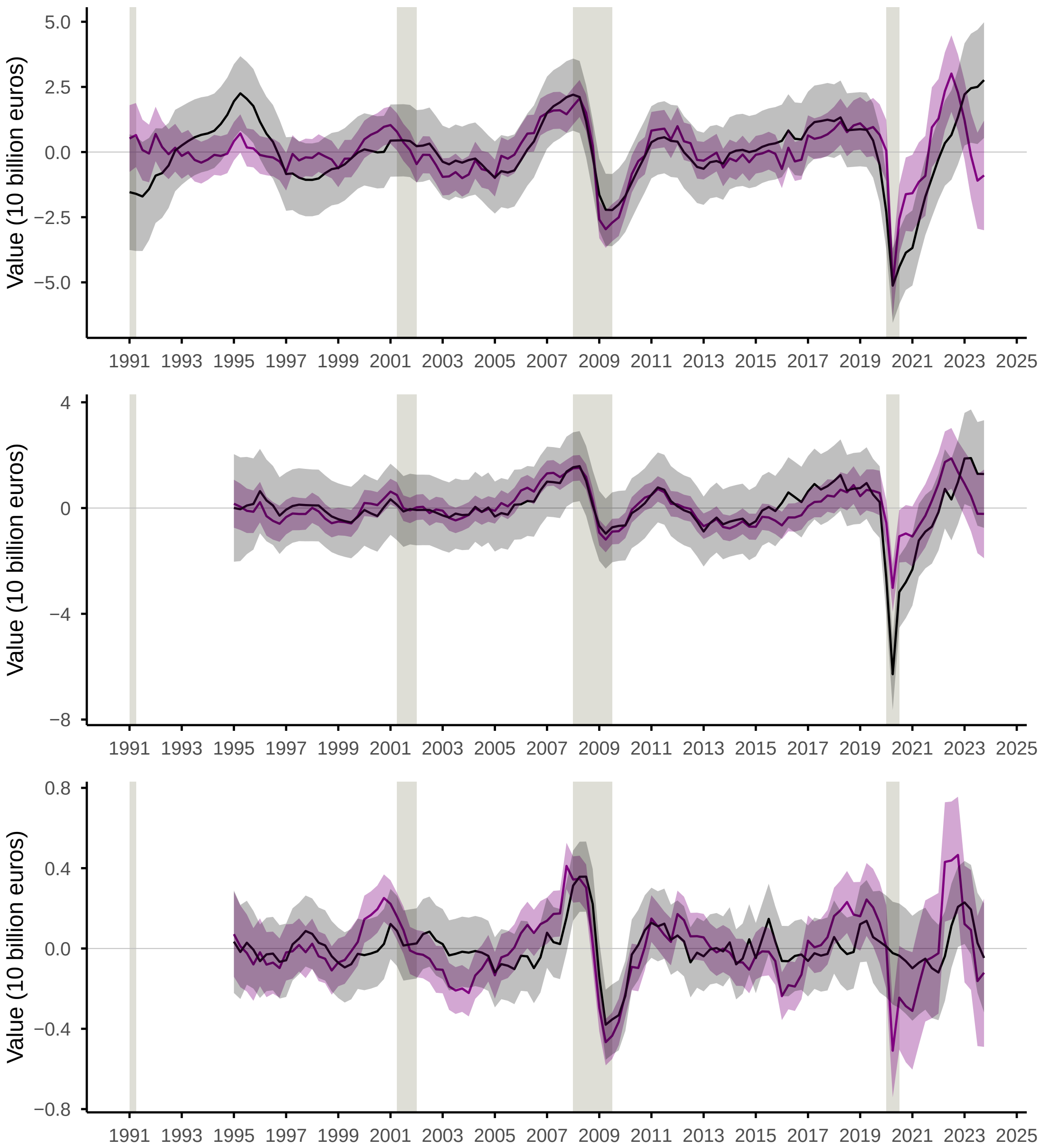

Figure 3: Smoothed estimates of the common cycle component of the constrained multivariate UC models (in purple), for countries Germany (top panel), Italy (middle panel) and the Netherlands (bottom panel). The aggregated common cycle loading estimates are used to scale the cycle to the level of GDP. The smoothed estimate of the cycle from a univariate UC model (smooth trend, seasonal, cycle plus irregular) applied to GDP is also displayed (in black) for each country. All estimates are accompanied by their model-implied 95% confidence intervals. Vertical shaded areas depict NBER US recession periods.

5 Synchronized predictions of GDP for official statistics

Next, we discuss model-based forecasting of the national accounts data set in a consistent and synchronized manner, using the restrictions implied by matrix $\boldsymbol{W}$ and vector $\boldsymbol{w}$ in Section 2.4. These restrictions are imposed on the covariance matrices of the disturbances in the transition equation and on the mean and covariance matrix of the initial state vector. It guarantees that the linear restrictions on the state vector carry over to predictions computed by the Kalman filter; see Durbin and Koopman (2012, Chapter 4) for a description of the Kalman filter. Consequently, the one-step ahead prediction of the observation vector $\boldsymbol{y}_t$ follows a scalar zero-sum restriction, that is $\boldsymbol{w}\,\hat{\boldsymbol{y}}_{t+1|t} = 0$, where $\hat{\boldsymbol{y}}_{t+1|t}$ is the one-step ahead prediction of $\boldsymbol{y}_t$. Forecasting with the state space model involves propagating the Kalman filter beyond the last time point T while treating future observations as missing. Multi-step forecasts are generated by iteratively applying this prediction step of the filter, hence the zero-sum restriction will be embedded in $\hat{\boldsymbol{y}}_{T+1|T}, \ldots, \hat{\boldsymbol{y}}_{T+h|T}$ for $h = 2, 3, \ldots$ as well.

To demonstrate the effectiveness of the constrained multivariate UC model in synchronizing forecasts, we have generated one-step ahead predictions for the national account data of Germany. The aggregated point forecasts of $\boldsymbol{y}_{E,t}$ and $\boldsymbol{y}_{P,t}$ are presented in Figure 4. Below these graphs, the corresponding discrepancy between $\sum_{i=1}^{n_E} \hat{y}_{E,i,t+1|t}$ and $\sum_{i=1}^{n_P} \hat{y}_{P,i,t+1|t}$, i.e. $\boldsymbol{w}\hat{\boldsymbol{y}}_{t+1|t}$, is plotted over time. The graphs in the upper half motivate why constraints should be introduced in the model-based forecasting exercise. The plots on the top left of the figure provide one-step ahead forecasts using two multivariate UC models with trend, seasonal, and common cycle components for the expenditure and production variables separately. The top right plot combines the variables into one model, leading to the unconstrained version of the common cycle model. These models are cast in the standard linear Gaussian state space form. The results clearly show that these models cannot create consistent initial estimates that are adequate for statistical production. The discrepancy reduces slightly by combining the variables in one model. When looking at the plots in the lower half of the figure, we clearly see the benefit of the constrained state space model. The discrepancy does not persist when the models are cast into constrained state space form and the forecasts are created as described above. The lower left graphs are the results of one-step ahead signal predictions (taking all state variables into consideration), which is compared to GDP at current prices. Finally, the lower right graphs show the flexibility of the model to assist statisticians

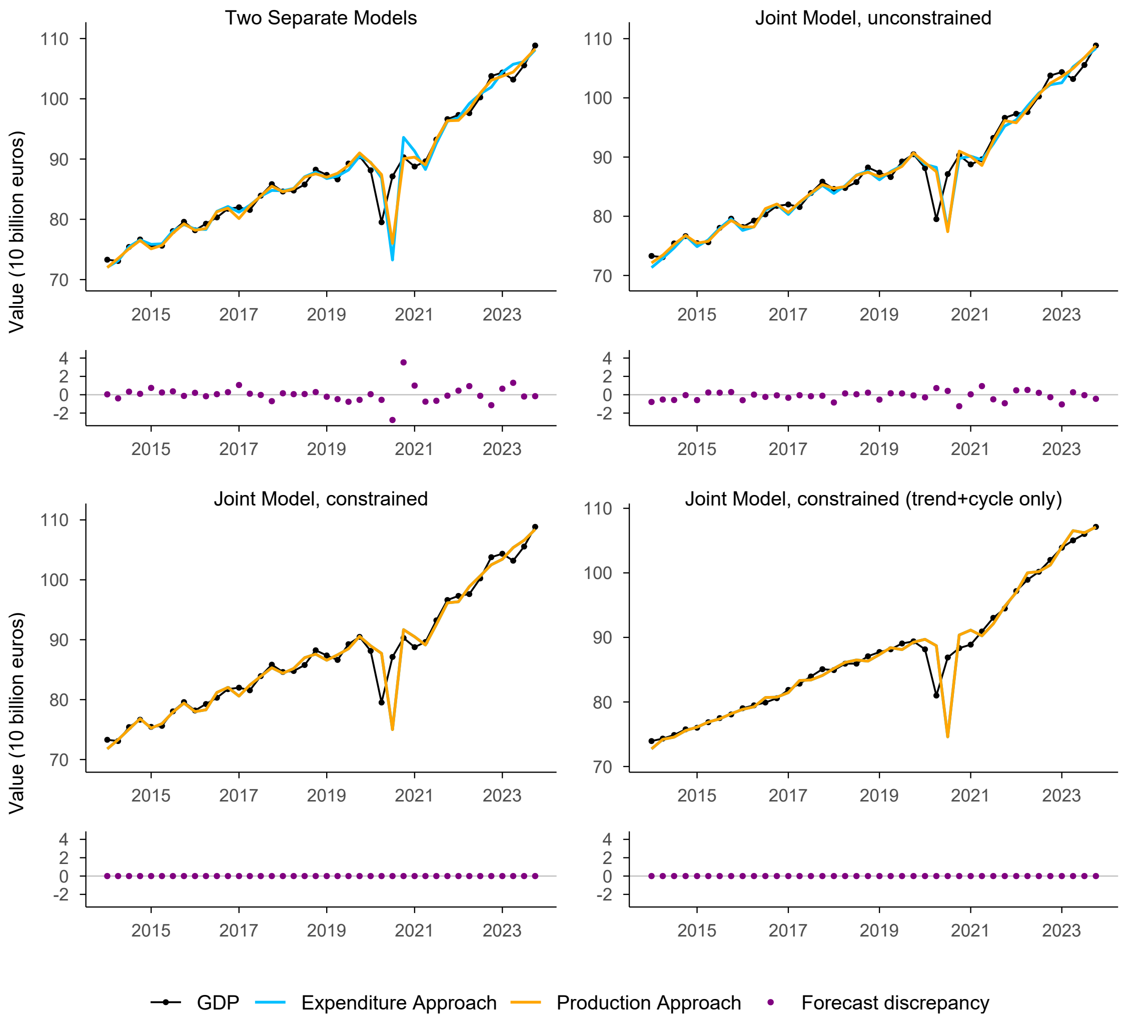

Figure 4: One-step ahead forecasts of GDP in Germany, by aggregating the point forecasts of the national accounts of both expenditure (in blue) and production approaches (in orange), against observed GDP (in black). In bottom-right panel, GDP data is seasonally adjusted (in black). The discrepancy between the aggregated point forecasts of both approaches is shown below each graph (in purple) for every time period t.

with model-based forecasting of seasonally adjusted variables. In this case, the aggregated trend and cycle state predictions are plotted alongside the seasonally adjusted GDP figures. Since the zero-sum restriction is applied separately per UC, as described in (10), the forecast discrepancy remains absent. We present the result for the common cycle model, but this choice is arbitrary. Similar illustrations can be based on models with similar cycles and no cycle. The common cycle model is simply selected because it performs best, as shown in Section 4.

Finally, we argue that our constrained multivariate UC model leads to more accurate signal extraction. We expect the predicted, filtered and smoothed estimates of the state vector to be more efficient because (1) the filter uses data from *all* national accounts variables rather than only the single variable GDP, (2) the restrictions act as additional information on the data generating processes that are incorporated in the system. Figure 5 illustrates this pattern by showing the standard errors of the predicted, filtered and smoothed signal estimates at the aggregated level of GDP using the full (un)constrained multivariate common cycle model ($N = 15$) and the univariate UC model. The results are presented for Germany. For a clear presentation, all models are without the COVID-19 modification and hence have time-invariant variances for the slope disturbance. Hence, the Kalman filter reaches steady state solutions for all models, see Figure 5 (upper panel). The standard error scales in the three panels of Figure 5 differ to facilitate better visual comparisons.

A clear pattern emerges when comparing the resulting standard errors of the GDP signal estimates (trend plus seasonal plus cycle) for the constrained, unconstrained and univariate models. The constrained multivariate UC model yields the lowest standard errors in all cases of one-step ahead prediction, filtering and smoothing. The two aggregated GDP standard errors of the unconstrained model show a reduction in the estimation error compared to the univariate model in the case of smoothing. It shows that the multivariate model benefits from the inclusion of cross-sectional information from the national accounts variables. In case of prediction and filtering, the steady state standard error corresponding to the aggregated expenditure variables in the multivariate model is larger than the univariate model equivalent. This result is of interest and is probably due to a mix of reasons. First, the NSI data for expenditure variables may have more noise, and therefore may contain less information, compared to the production variables. Second, the cross-section is much larger in the production breakdown of GDP ($n_P = 11$ for Germany) compared to the expenditure side ($n_E = 4$). Third, restrictions imposed on the covariance matrix of state disturbances for the full model may have more impact on the expenditure variables and less impact on the production variables. In case of smoothing, the noisy expenditure variables are subject to more averaging and this benefits the accuracy of their GDP estimate, much more so than for the less noisy single GDP composite variable used in the univariate model. Finally, the unconstrained model assumes that there are no contemporaneous correlations within trend and seasonal disturbances, while they exhibit one shared scalar variance each. This single variance parameter estimate is dominated by the trends or seasonal patterns of the 11 production series. In the constrained model, this single estimate is allocated over all corresponding (off-diagonal) elements in the co-

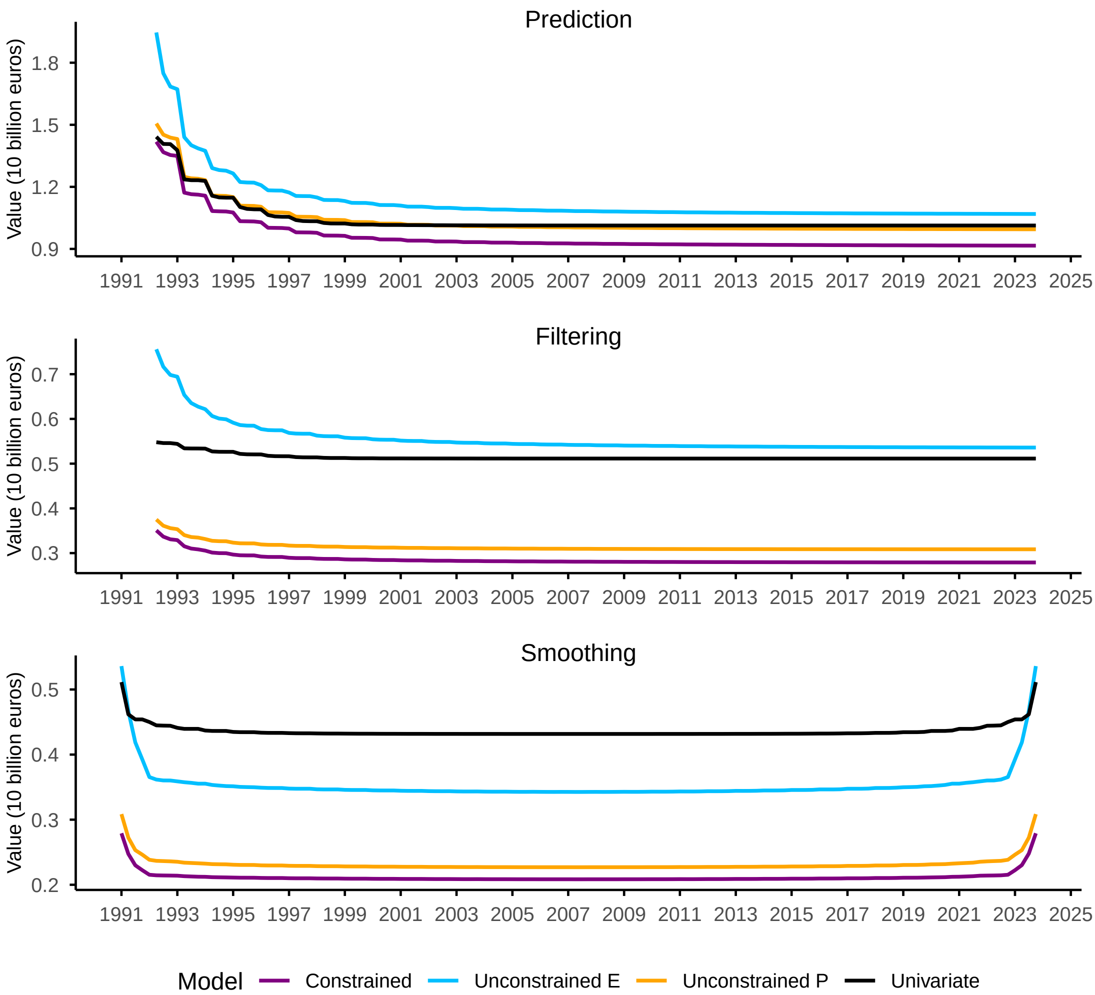

Figure 5: Standard errors of one-step prediction (upper panel), filtered (middle panel) and smoothed (lower panel) estimates of the GDP signal for Germany ($T = 132$), based on a univariate model (black, $N = 1$), aggregated production (orange, $N_P = 11$) and expenditure (blue, $N_E = 4$) variables from an unconstrained multivariate model ($N = 15$), as well as from a constrained model where they are equal by definition (purple, $N = 15$). All models are without the COVID-19 correction: slope disturbance variance is time-invariant.

variance matrix according to the restriction, lowering the disturbance variance on the diagonal, resulting in lower conditional variances of the state vector. It explains why the standard error from the production model is structurally lower than the one of the expenditure model. Overall, we have provided evidence that the constrained multivariate model can provide efficient, synchronized, and consistent signal estimates for extraction and forecasting.

6 Conclusion

This study investigates the use of the constrained multivariate unobserved components (UC) model for the complete set of national accounts data. The model is constrained by the introduction of linear restrictions that are imposed on the unobserved components. This is achieved by casting the model into a constrained state space form. The transition equation disturbances and the initial state vector are projected onto the zero-sum subspace through linear transformations of their respective covariance matrices. We illustrate that these restrictions are required for numerically consistent and synchronized forecasting of national accounts data of the expenditure and production sides of the economy, while both sides have GDP in common as their aggregated value. We show that the Kalman filter and smoother estimates of the unobserved components corresponding to the expenditure and production variables are numerically consistent and equal to GDP in an application to national accounts data of Germany, Italy and the Netherlands. Exact maximum likelihood estimation is adequate despite the additional model complexity caused by the constraints and the high-dimensional parameter and state vectors. We find clear evidence of macroeconomic cycles in the three EU countries. The proposed methodology provides a model-based alternative to the flash estimates of macroeconomic time series, currently used at NSIs. In addition, our method provides uncertainty measures for these initial estimates. It solves the problem of nowcasting and forecasting inconsistencies, both for traditional and seasonally adjusted variables, in the official statistics literature. Moreover, using the complete set of national accounts data provides more accurate filtering of macroeconomic cycles at central banks compared to the conventional univariate approach. Since the constrained multivariate UC model offers advancements for practitioners in all areas, we argue that the developed methodology is of significance from both macroeconomic policy and official statistics perspectives.

References

Andreou, E., Ghysels, E., and Kourtellos, A. (2013). Should macroeconomic forecasters use daily financial data and how? Journal of Business & Economic Statistics, 31(2):240–251.

Aruoba, S. B., Diebold, F. X., and Scotti, C. (2009). Real-time measurement of business conditions. Journal of Business & Economic Statistics, 27(4):417–427.

Athanasopoulos, G., Ahmed, R. A., and Hyndman, R. J. (2009). Hierarchical forecasts for australian domestic tourism. International Journal of Forecasting, 25(1):146–166.

Athanasopoulos, G., Hyndman, R. J., Kourentzes, N., and Panagiotelis, A. (2024). Forecast reconciliation: A review. International Journal of Forecasting, 40(2):430–456.

Azevedo, J., Koopman, S. J., and Rua, A. (2006). Tracking the business cycle of the euro area: a multivariate model-based band-pass filter. Journal of Business & Economic Statistics, 24(3):278–290.

Bańbura, M., Giannone, D., Modugno, M., and Reichlin, L. (2013). Now-casting and the real-time data flow. In Elliott, G., Granger, C., and Timmermann, A., editors, Handbook of economic forecasting, volume 2, pages 195–237. Elsevier.

Bell, W. R., Fonzo, T. D., Ladiray, D., McElroy, T. S., and Smith, P. A. (2024). Preface to time series special issue of the Journal of Official Statistics. Journal of Official Statistics, 40(4):499–507.

Bikker, R., van den Brakel, J., Krieg, S., Ouwehand, P., and van der Stegen, R. (2019). Consistent multivariate seasonal adjustment for Gross Domestic Product and its breakdown in expenditures. Journal of Official Statistics, 35(1):9–30.

Bobeica, E. and Hartwig, B. (2023). The Covid-19 shock and challenges for inflation modelling. International Journal of Forecasting, 39(1):519–539.

Carriero, A., Clark, T. E., Marcellino, M., and Mertens, E. (2024). Addressing Covid-19 outliers in BVARs with stochastic volatility. The Review of Economics and Statistics, 106(5):1403–1417.

Cascaldi-Garcia, D., Luciani, M., and Modugno, M. (2024). Lessons from nowcasting GDP across the world. In Clements, M. P. and Galvão, A. B., editors, Handbook of Research Methods and Applications in Macroeconomic Forecasting, pages 187–217. Edward Elgar Publishing.

Cimadomo, J., Giannone, D., Lenza, M., Monti, F., and Sokol, A. (2022). Nowcasting with large Bayesian vector autoregressions. Journal of Econometrics, 231(2):500–519. Special Issue: The Econometrics of Macroeconomic and Financial Data.

de Winter, J., Koopman, S. J., and Hindrayanto, I. (2022). Joint decomposition of business and financial cycles: Evidence from eight advanced economies. Oxford Bulletin of Economics and Statistics, 84(1):57–79.

Del Negro, M. and Schorfheide, F. (2013). Chapter 2 - DSGE model-based forecasting. In Elliott, G. and Timmermann, A., editors, Handbook of Economic Forecasting, volume 2 of Handbook of Economic Forecasting, pages 57–140. Elsevier.

Doornik, J. A. (2021). An Object-oriented Matrix Programming Language Ox 9. Timberlake.

Doran, H. E. (1992). Constraining Kalman filter and smoothing estimates to satisfy time-varying restrictions. The Review of Economics and Statistics, pages 568–572.

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

Eurostat (2013a). European system of accounts, esa 2010.

Eurostat (2013b). Handbook on quarterly National Accounts. Office for Official Publications of the European Communities.

Eurostat (2014). Eurostat Review on National Accounts and Macroeconomic Indicators. Publications Office of the European Union.

Eurostat (2015). ESS guidelines on seasonal adjustment.

Fernandes, M. C., Dutra, T. M., Dias, J. C., and Teixeira, J. C. (2023). Modelling output gaps in the euro area with structural breaks: The covid-19 recession. Economic Analysis and Policy, 78:1046–1058.

Giannone, D., Monti, F., and Reichlin, L. (2016). Exploiting the monthly data flow in structural forecasting. Journal of Monetary Economics, 84:201–215.

Giannone, D., Reichlin, L., and Simonelli, S. (2009). Nowcasting euro area economic activity in real time: the role of confidence indicators. National Institute Economic Review, 210:90–97.

Giannone, D., Reichlin, L., and Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. Journal of Monetary Economics, 55(4):665–676.

Guérin, P. and Marcellino, M. (2013). Markov-switching MIDAS models. Journal of Business & Economic Statistics, 31(1):45–56.

Harlaar, L., Commandeur, J., van den Brakel, J., Koopman, S. J., Bos, N., and Bijleveld, F. (2025). Statistical early warning models with applications. Journal of the Royal Statistical Society Series A: Statistics in Society, page qnaf193.

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

Harvey, A. C. (2006). Chapter 7 forecasting with unobserved components time series models. In Elliott, G., Granger, C., and Timmermann, A., editors, Handbook of economic forecasting, volume 1 of Handbook of Economic Forecasting, pages 327–412. Elsevier.

Harvey, A. C. and Koopman, S. J. (1993). Forecasting hourly electricity demand using time-varying splines. Journal of the American Statistical Association, 88(424):1228–1236.

Harvey, A. C. and Koopman, S. J. (1997). Multivariate structural time series models. In Systematic dynamics in economic and financial models, pages 269–298. John Wiley & Sons.

Hecq, A., Ternes, M., and Wilms, I. (2022). Hierarchical regularizers for mixed-frequency vector autoregressions. Journal of Computational and Graphical Statistics, 31(4):1076–1090.

Hungnes, H., Skjerpen, T., Hamre, J. I., Jansen, X. C., Pham, D. Q., and Sandvik, O. (2024). Structural break in the Norwegian labor force survey due to a redesign during a pandemic. Journal of Official Statistics, 40(1):122–160.

Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G., and Shang, H. L. (2011). Optimal combination forecasts for hierarchical time series. Computational Statistics Data Analysis, 55(9):2579–2589.

Kohns, D. and Bhattacharjee, A. (2023). Nowcasting growth using Google Trends data: A bayesian structural time series model. International Journal of Forecasting, 39(3):1384–1412.

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

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

Lenza, M. and Primiceri, G. E. (2022). How to estimate a vector autoregression after March 2020. Journal of Applied Econometrics, 37(4):688–699.

Modugno, M. (2013). Now-casting inflation using high frequency data. International Journal of Forecasting, 29(4):664–675.

Pandher, G. S. (2002). Forecasting multivariate time series with linear restrictions using constrained structural state-space models. Journal of Forecasting, 21(4):281–300.

Pandher, G. S. (2007). Modelling & controlling monetary and economic identities with constrained state space models. International Statistical Review, 75(2):150–169.

Pizzinga, A. (2012). Restricted Kalman filtering: Theory, methods, and application, volume 12. Springer Science & Business Media.

Proietti, T. (2000). Comparing seasonal components for structural time series models. International Journal of Forecasting, 16(2):247–260.

Proietti, T., Giovannelli, A., Ricchi, O., Citton, A., Tegami, C., and Tinti, C. (2021). Nowcasting GDP and its components in a data-rich environment: The merits of the indirect approach. International Journal of Forecasting, 37(4):1376–1398.

R Core Team (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rünstler, G. and Vlekke, M. (2018). Business, housing, and credit cycles. Journal of Applied Econometrics, 33(2):212–226.

Schorfheide, F. and Song, D. (2021). Real-time forecasting with a (standard) mixed-frequency VAR during a pandemic. Working Paper 29535, National Bureau of Economic Research.

Scott, S. L. and Varian, H. R. (2014). Predicting the present with Bayesian structural time series. International Journal of Mathematical Modelling and Numerical Optimisation, 5(1-2):4–23.

Smith, P. A. (2016). Google's MIDAS touch: Predicting uk unemployment with internet search data. Journal of Forecasting, 35(3):263–284.

Smith, P. A., van den Brakel, J., and Horsfield, G. (2024). Estimation of discontinuities from the introduction of tablet based data collection on the international passenger survey. Journal of Official Statistics, 40(4):748–782.

Stundziene, A., Pilinkiene, V., Bruneckiene, J., Grybauskas, A., Lukauskas, M., and Pekarskiene, I. (2024). Future directions in nowcasting economic activity: A systematic literature review. Journal of Economic Surveys, 38(4):1199–1233.

United Nations, Organisation for Economic Co-operation and Development, International Monetary Fund, European Commission, and World Bank (2009). System of National Accounts 2008. International Monetary Fund, USA.

van den Brakel, J. and Krieg, S. (2016). Small area estimation with state space common factor models for rotating panels. Journal of the Royal Statistical Society Series A: Statistics in Society, 179(4):763–791.

van den Brakel, J., Souren, M., and Krieg, S. (2022). Estimating monthly labour force figures during the Covid-19 pandamic in the Netherlands. Journal of the Royal Statistical Society Series A: Statistics in Society, 185(4):1560–1583.

van der Ploeg, F. (1985). Econometrics and inconsistencies in the National Accounts. Economic Modelling, 2(1):8–16.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wickramasuriya, S. L., Athanasopoulos, G., and Hyndman, R. J. (2019). Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. Journal of the American Statistical Association, 114(526):804–819.

A Appendix: state space formulations

The measurement equation of the constrained multivariate UC model with linear restrictions on the trend, seasonal and cycle components in state space form can be written as follows:

$$\begin{pmatrix} \boldsymbol{y}_{E,t} \\ \boldsymbol{y}_{P,t} \end{pmatrix} = \begin{bmatrix} \boldsymbol{Z}_\mu & \mathbf{O}_N & \boldsymbol{Z}_\gamma & \boldsymbol{Z}_\psi \end{bmatrix} \begin{pmatrix} \boldsymbol{\mu}_t \\ \boldsymbol{\nu}_t \\ \boldsymbol{\gamma}_t \\ \boldsymbol{\psi}_t \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_{E,t} \\ \boldsymbol{\varepsilon}_{P,t} \end{pmatrix}, \qquad \begin{pmatrix} \boldsymbol{\varepsilon}_{E,t} \\ \boldsymbol{\varepsilon}_{P,t} \end{pmatrix} \sim \boldsymbol{N}(\mathbf{0}, \boldsymbol{H}), \tag{13}$$

with $\boldsymbol{Z}_\mu = \boldsymbol{I}_N$, $\boldsymbol{Z}_\gamma = (1,0,1) \otimes \boldsymbol{I}_N$ and $\boldsymbol{Z}_\psi = (1,0) \otimes \boldsymbol{I}_N$, where $\boldsymbol{I}_N$ reflects an $N \times N$ identity matrix, and $\mathbf{O}_N$ refers to an $N \times N$ zero matrix. The observation vector $\boldsymbol{y}_t = (\boldsymbol{y}'_{E,t}, \boldsymbol{y}'_{P,t})'$ is $N \times 1$. The irregular term $\boldsymbol{\varepsilon}_t = (\boldsymbol{\varepsilon}'_{E,t}, \boldsymbol{\varepsilon}'_{P,t})'$ is a $N \times 1$ vector of mutually and serially independent disturbances (also of all state disturbance terms in the model), hence $\mathbb{E}(\boldsymbol{\varepsilon}_t) = \mathbf{0}$ is an $N \times 1$ vector of zeros and $\mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}'_t) = \boldsymbol{H}$ is an $N \times N$ diagonal matrix. The unobserved state vector contains $\boldsymbol{\mu}_t = (\boldsymbol{\mu}'_{E,t}, \boldsymbol{\mu}'_{P,t})'$ which is a $N \times 1$ vector of long-term stochastic trends, $\boldsymbol{\nu}_t = (\boldsymbol{\nu}'_{E,t}, \boldsymbol{\nu}'_{P,t})'$ which is a $N \times 1$ vector of stochastic trend slopes, $\boldsymbol{\gamma}_t = (\boldsymbol{\gamma}'_{E,1,t}, \boldsymbol{\gamma}'_{P,1,t}, \boldsymbol{\gamma}^{*\prime}_{E,1,t}, \boldsymbol{\gamma}^{*\prime}_{P,1,t}, \boldsymbol{\gamma}'_{E,2,t}, \boldsymbol{\gamma}'_{P,2,t})'$ which is a $3N \times 1$ vector of stochastic seasonal effects and $\boldsymbol{\psi}_t = (\boldsymbol{\psi}'_{E,t}, \boldsymbol{\psi}'_{P,t}, \boldsymbol{\psi}^{*\prime}_{E,t}, \boldsymbol{\psi}^{*\prime}_{P,t})'$ which is a $2N \times 1$ vector of stochastic cyclical patterns. This makes the total state vector $\boldsymbol{\alpha}_t = (\boldsymbol{\mu}'_t, \boldsymbol{\nu}'_t, \boldsymbol{\gamma}'_t, \boldsymbol{\psi}'_t)'$, which has $7N \times 1$ as dimensions. The transition equation of $\boldsymbol{\alpha}_t$ is as follows:

$$\begin{pmatrix} \boldsymbol{\mu}_{t+1} \\ \boldsymbol{\nu}_{t+1} \\ \boldsymbol{\gamma}_{t+1} \\ \boldsymbol{\psi}_{t+1} \end{pmatrix} = \begin{bmatrix} \boldsymbol{I}_N & \boldsymbol{I}_N & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \boldsymbol{I}_N & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \boldsymbol{C}_\gamma & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & \boldsymbol{C}_\psi \end{bmatrix} \begin{pmatrix} \boldsymbol{\mu}_t \\ \boldsymbol{\nu}_t \\ \boldsymbol{\gamma}_t \\ \boldsymbol{\psi}_t \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\zeta}_t \\ \boldsymbol{\omega}_t \\ \tilde{\boldsymbol{\omega}}_t \end{pmatrix}, \qquad \begin{pmatrix} \boldsymbol{\zeta}_t \\ \boldsymbol{\omega}_t \\ \tilde{\boldsymbol{\omega}}_t \end{pmatrix} \sim \boldsymbol{N}(\mathbf{0}, \boldsymbol{Q}_t) \tag{14}$$

where $\boldsymbol{C}_\gamma$ and $\boldsymbol{C}_\psi$ follow from (4) and (5) and all $\mathbf{O}$ are conformable zero matrices. Moreover,

$$\boldsymbol{Q}_t = \operatorname{diag}\left[\boldsymbol{Q}_{\zeta_t}, \boldsymbol{Q}_\omega, \boldsymbol{Q}_{\tilde{\omega}}\right] \tag{15}$$

where the slope disturbances of the trend have a time-varying covariance matrix due to COVID factor q_t^2, while the covariance matrices of the seasonal and cycle disturbances are time-invariant. Note that (15) indicates that the UC disturbance terms are independent from each other, which holds for all lags as well. The covariance matrices $\boldsymbol{Q}_{\zeta_t}$, $\boldsymbol{Q}_\omega$ and $\boldsymbol{Q}_{\tilde{\omega}}$ are affected by the linear restrictions on the states.