UNIVERSITA DEGLI STUDI DI MILANO BICOCCA

Department of Economics, Management and Statistics
PhD in Economics

Ramiz Salman
ID: 834578

Identification of Common Economic Cycles Using Optimal Multivariate Filters

A thesis presented for the degree of Doctor of Philosophy

PhD Supervisor: Prof. Matteo Pelagatti

PhD Coordinator: Prof. Matteo Manera

33rd Cycle
May 2021, Milan

Summary

This thesis includes two essays that are focused on developing multivariate filter approaches to be used for extracting common cyclical components where the common components can be used as an estimator of a business cycle.

The first chapter aims to develop an optimal multivariate filter in order to extract common cyclical components of macroeconomic indicators. The filter allows macroeconomic series to be modeled as a phase shifted version of a coinciding business cycle (BC) while keeping other time series components such as the stochastic trend and idiosyncratic shocks intact (i.e. they are individually specified for each series). Earlier studies of Rünstler (2004), Valle e Azevedo et al. (2006) have applied phase shift in the form of a delay parameter when specifying lead-lag cycles. However, the lead-lag relationship is defined by rotating the baseline cycle which leads to loss of information. This deficiency is especially important if one considers working in continuous time. Therefore, this paper improves on the former technique by allowing a more flexible phase shift mechanism on the original BC. This in turn should lead to more realistic estimates and filters considering that the underlying data is generated through a continuous time framework. The study starts by presenting a structure for bi-variate time series system and then extends to model to incorporate a structure for three time series and beyond. Kalman filter and smoothing recursions are applied to compute the smoothed cycle estimates and to construct the likelihood function. Using simulated data, we test both model specifications by carrying out a grid search of the initial delay parameter to see the likelihood behavior as the parameter moves into fractional neighborhoods. Afterwards, applying the methodology to a set of EU countries and macroeconomic indicators; the study aims to shed light to the presence of cyclical heterogeneity at country level economic activity for major EU member states. A second empirical study provides analysis on how the model can be implemented for assigning a lead/lag ordering to three main economic indicators of a single country.

The second chapter implements a multivariate non-parametric filtering approach; the Vertical Multivariate Singular Spectrum Analysis (V-MSSA) of Hassani and Mahmoudvand (2013) and Golyandina et al. (2013) to be applied for identifying a common economic cycle indicator. The methodology is a data-driven and a rigorous procedure that can decompose a time series into many sub components. By exploiting this ability of the SSA, the paper aims to first extract cyclical components based on frequency characteristics and then follow by choosing only common cyclical component pairs with-in the business cycle frequency spectrum. These components will then be aggregated for constructing an EU region wide Business cycle indicator.

The chapter outlines each steps of the algorithm that will eventually identify the SSA filter to act as a band-pass filter. The study then proceeds with simulation based data where the common cycle can be controlled and extracted a priori as a benchmark to the SSA-based filter estimates. The study follows with an empirical analysis similar to the framework set in Valle e Azevedo et al. (2006) with the aim to identify a Euro region business cycle indicator. The SSA based filter estimate is compared with Euro region economic activity indicators; the EuroCoin and the quarterly GDP growth rate of the EU area. Our results presents evidence of a successful alternative for tracing the cyclical position of the EU economy from a much smaller data set. Moreover, the constructed indicator also could serve as an unobserved proxy for a monthly growth cycle. A further analysis is also conducted to reveal whether the SSA based approach can be considered as an alternative to parametric filtering methods. In this framework, the study presents results from the empirical analysis of common cycle extraction using two multivariate UCM alternatives. Comparing results with the SSA based filter, we are able to reinforce our initial findings and use the parametric based results as a robustness check.

Acknowledgements

I started my Ph.D. at the University of Milan-Bicocca 4 years ago. I am very excited to have finalized my thesis project in which several people have supported me during the time.

Firstly, I am grateful to my supervisor, Professor Matteo Pelagatti for his research support, encouragement and expertise on both theoretical and technical aspects of this study. This work would not have been possible without him. I am grateful that he inspired me and that he introduced me to the area of multivariate filters and their application to macroeconometrics.

I am thankful to the PhD program committee for a very valuable academic experience and for addressing many of my Ph.D related issues, research travel, amongst others.

I would like to thank Professor Alain Hecq of Maastricht University, my host supervisor at the Department of Quantitative Economics, for allowing me to use the University facilities during my Ph.D research visit and also providing me an office where I continued to work and improve on my thesis.

I express also my sincere gratitude to my family, my mother Ayşe for everything she has given me in this life, my brother Harun and father Omar for their patience and support.

Contents

List of Figures

List of Tables

Chapter 1

Extracting Phase Shifted Common Cycles using Optimal Multivariate Filters

1.1 Introduction

The main aim of filters used in Business Cycle (BC) analysis is to extract cyclical components inherit in time series data. Popular techniques in extracting BC from economic data could be referred to Hodrick and Prescott (1997), Baxter and King (1999), Gomez (2001) and Christiano and Fitzgerald (2003). However, despite gaining vast amount of popularity, there is to be found certain reservations regarding the usage of these methodologies. Therefore, papers advocating some caution and to not hastily apply filtering methodologies without thoroughly understanding certain drawbacks is evident from the literature. The HP filter is a low pass filter but shares exhibits certain drawbacks. Its gain function is monotonically decreasing and hence is unable to filter out the high frequency components of a time series. This leads to a non smoothed filtered cycle if the time series has a non-zero spectrum at high frequencies. Further, Cogley and Nason (1995) illustrate that the HP filtered cycle can exhibit strong correlation between lags even though the DGP indicates otherwise and therefore generate a spurious relation.[1] With the aim of achieving a smoother cycle estimate, Baxter and King (1999) (BK) improves some of the drawbacks of the HP filter. This is done by introducing a symmetric finite (K) order moving average when constructing the gain function. However the side back to this is the loss of data and a continuing albeit weaker presence of spurious regression.[2] In this framework, Schenk-Hoppé (2001) and Ivanov (2015) show findings of spurious cycles generated by the application of HP and BP filters. Similar papers that examine the disadvantages of these filters are Mohr (2005), Bates and Ndiaye (2014), Grant and Chan (2017). The ideal band-pass filter in Christiano and Fitzgerald (2003), (CF) is able to generate an ideal band-pass filter which is more successful in subduing the irregular frequencies while also exploiting the full sample. In terms of application to economic data, papers of Orphanides and Norden (2002) and Kamber et al. (2018) compare output gap estimates from these filters to find that none are reliable. Nilsson and Gyomai (2011) carry out a turning point analysis to find that HP filtered cycles achieve a better performance to CF filtered cycles. Van Aarle et al. (2008) and De Haan et al. (2008), Larsson and Tamas (2012) find little differences between the filtering techniques as they provide high correlation between filtered cycles of HP, BK and CF.

1.1.1 Unobserved Component Model in Business Cycle Literature

The Unobserved Component Model (UCM) methodology in this context, provide a recipe to encounter possible spurious relations that maybe be identified in the previously explained filters. Initially applied by Harvey and Jaeger (1993) to quarterly GDP series, the methodology has gained popularity through the years. Moreover, as pointed by Harvey and Trimbur (2003), Valle e Azevedo et al. (2006) the extracted cycle from their UCM is found to be smooth and clearly defined while showing characteristics of an ideal band-pass filter. Pedregal et al. (2001), Mitra and Sinclair

[1] See Nelson and Kang (1981) for the initial critique of the filter.

[2] See Woitek et al. (1998) for a technical discussion.

(2012) state the superiority of simultaneously extracting trend and cycles in a time series with respect to HP and bandpass filtering. Doménech and Gomez (2006), Basistha and Nelson (2007), Planas et al. (2008) and Harvey (2011) convey their preference for a UCM rather than the common filters for analyzing output gap cycles. Hwa Oh and Zivot (2006), Delle Monache and Harvey (2011) highlight the flexibility of applying the UCM while Proietti (2006) states that UCM should be prioritized when cycles are to be extracted from ARIMA processes. The UCM has also been an important tool in the construction for the proposed filter of Harvey and Trimbur (2003). By repetitively substituting the error component in the constructed stochastic cycle with a stochastic cycle of similar features, the authors are able to achieve a smoothed cycle with features similar to one extracted from a band-pass filter. (Valle e Azevedo et al., 2006), Harvey et al. (2007), Gonzalez et al. (2015) apply the filtering technique to achieve smoothed cycles in their analysis. Perron and Wada (2009) compares cycle estimates from commonly used techniques (i.e. UCM, HP, BK, amongst others). Their findings show that UCM gives the closest result to the NBER cycle chronology while HP and BK faring less similar and tending to underestimate episodes of hard recessions and strong expansion periods. Finally, Başţürk et al. (2014) and Perron and Wada (2016) compare HP filtered cycles with UCM filtered ones when breaks are allowed in the trend component of a time series such as slope and level shifts. They find that UCM filtered cycles are able to better account for dynamics inherent in the low frequency component.

1.1.2 UCM and Phase Shift in Business Cycle Analysis

Our paper centers on UCM methodology and its extended version of accommodating a phase shift mechanism. Henceforth, it will be beneficial to also include literature with similar focus. Rünstler (2004), was first to introduce the phase shift structure such that stochastic cycles can be modeled as continuously shifted versions of a base cycle. Rotating the base cycle by a defined angle, his model generates cycles which can be considered a shifted version of the common cycle. Examples of its application on BC's are Valle e Azevedo et al. (2006), Koopman and Azevedo (2008), Creal et al. (2010), Moës (2012), Chen and Mills (2012). The application of the phase shift mechanism has also been extended to financial cycles (FC).[3] Chen and Mills (2012), Rünstler and Vlekke, 2015, Galati et al. (2016), Koopman et al. (2016) incorporate the UCM with a phase shift mechanism to analyze lead-lag dynamics at both BC and FC frequencies.

[3]According to Claessens et al. (2012) and Drehmann et al. (2012) FC's are more persistent and have greater amplitudes wrt BC's and influence steepness of BC movements during economic downturns and peaks.

1.2 Methodology

1.2.1 Phase Shift with Angular Shift

The phase shift mechanism works such that the shifted cyclical components that are extracted are rotated versions of a baseline cycle which is itself extracted from an indicator such as Industrial production (IP). The angle that the common cycle rotates is a product of two variables; the frequency (λ) and the phase shift (r). In this way, one can compare the relationship between cyclical components by the amount of phase shift with respect to a baseline cycle. The downside of the angular transformation to achieve a phase shift movement is that, the original cycle is no longer retained. For an econometrician point of view, this means the information content embedded in the data is not being used in an optimally efficient way. This is due to loss of information. More specifically, the method is unable to exploit extra information content through the lead-lag dynamics contained in the data. To make matters clearer, let us consider two economic indicators y_t and z_t driven by cyclical and an irregular components. Let us also assume that the cycle component in z_t is generated by the shifted version of the cycle component in y_t

$$\begin{aligned} y_t &= \psi_t + \varepsilon_t \\ z_t &= \phi_t + u_t \end{aligned} \tag{1.1}$$

where $\varepsilon_t, u_t \sim \text{WN}$, and the relation between the cyclical components is $\phi_t = \psi_{t+r}$ with $r \in \mathbb{R}$. According to this definition, ψ_t is the common cycle, ϕ_t is its shifted version by time lag of r. The stochastic cycle in a UCM framework can be shown as

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda & \sin\lambda \\ -\sin\lambda & \cos\lambda \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix} \tag{1.2}$$

where $\rho \in (0, 1]$ is a decay parameter, $\lambda \in (0, \phi]$ is an angle denoted in radians and represents the frequency of the cycles (ψ_t, ψ_t^*) and the error terms κ_t, κ_t^* are uncorrelated white noise processes with equal variance of σ_κ^2.[4] The compact form of (1.2) can be written as

$$\boldsymbol{\psi}_t = \rho \boldsymbol{T}(\lambda) + \boldsymbol{\psi}_{t-1} + \boldsymbol{\kappa}_t \tag{1.3}$$

where $\boldsymbol{T}(\lambda)$ is a rotation matrix with the property $\boldsymbol{T}^r = \boldsymbol{T}(\lambda)$, the error vector in compact form follows $\boldsymbol{\kappa}_t = (\kappa_t, \kappa_t^*)' \sim NID(\boldsymbol{0}, \sigma_{\kappa,jj}^2 \boldsymbol{I}^2)$ with $\sigma_{\kappa,jj}^2 = \sigma_\kappa^2 \; \forall \, j \in \mathbb{N}$. The idea is that $\boldsymbol{T}(\lambda)$ allows $\boldsymbol{\psi}_{t-1}$ to rotate by an angle (λ) on a circle that lies in the 2-dimensional plane with radius equal to amplitude (see Figure ref here!). In equation (1.3), $\boldsymbol{\psi}_{t-1}$ will be moving along a smaller circumference than the amplitude of the common cycle since the first term includes the decay parameter so that we have $\rho \boldsymbol{T}(\lambda) \boldsymbol{\psi}_{t-1}$. Finally, the error terms are added to reach the total sum which gives the final coordinates (ψ_t, ψ_t^*).

[4] The stochastic cycles (ψ_t, ψ_t^*) are the projections of the stochastic cycle in compact form $(\boldsymbol{\psi}_t)$ on the $(x, y) - axis$.

The phase shift mechanism in Rünstler (2004), uses the first row of $\boldsymbol{T}(\lambda)$ to end up with a shifted version of a cycle. Henceforth, equation (1.1) can be rewritten such that

$$\begin{aligned} y_t &= \psi_t + \varepsilon_t \\ z_t &= [\cos(r\lambda) \quad \sin(r\lambda)]\boldsymbol{\psi}_{t-1} + u_t \end{aligned} \tag{1.4}$$

The problem with equation (1.4), is that the shift in the cyclical component of z_t is not a sifted version of the original cycle ψ_t. The reason why Runstler's method is unable to achieve a shift of the exact cycle requires a short explanation. Let us first compare the optimal predictor of the stochastic cycle at time $t+h$ with its representation at $t+h$ using only its current value at time t. To reach the later representation, one needs to iterate h times equation (1.3) onto itself to reach the following

$$\psi_{t+h} = \rho^h \boldsymbol{T}^h \boldsymbol{\psi}_t + \sum_{j=0}^{h-1} \rho^j \boldsymbol{T}^j \kappa_{t+h-j}. \tag{1.5}$$

On the other hand, the prediction based on all past and current information available at t is

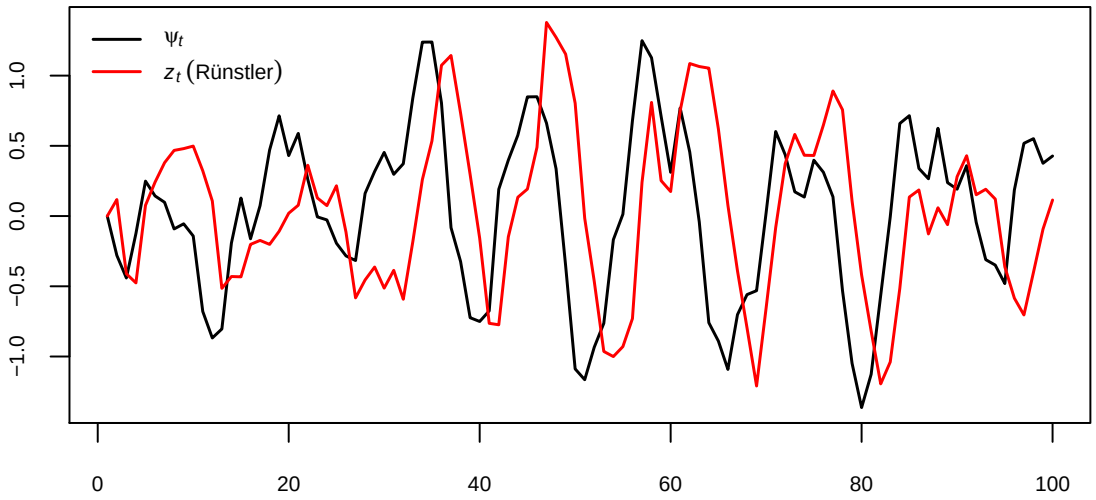
$$\begin{aligned} \widehat{\boldsymbol{\psi}}_{t+h|t} &= \mathbb{E}(\boldsymbol{\psi}_{t+h}|\boldsymbol{\psi}_t, \boldsymbol{\psi}_{t-1}, \cdots) = \mathbb{E}(\boldsymbol{\psi}_{t+h}|\boldsymbol{I}_t) \\ &= \rho^h [\cos(h\lambda) \quad \sin(h\lambda)]\boldsymbol{\psi}_t \end{aligned} \tag{1.6}$$

Neither the first element of ψ_{t+h} in equation (1.5) nor the conditional expectation of equation (1.6) are equivalent to the shifted cyclical component of z_t in equation (1.4). More specifically, Runstler's model fails to account for the information that is expressed in (1.5) and (1.6). The first component ρ^h in equation (1.5) requires the estimation of the decay parameter from the data set. This parameter is not included in the representation of z_t in equation (1.1). Additionally, the sequence of innovations in equation (1.5) are also disregarded in Runstler's equation for z_t.[5]

Figure 1.1 illustrates the drawbacks of Runstler method more clearly. For simplicity and to make our visual inspection more informative, we assume that z_t in equation (1.4) is only composed of the stochastic cycle shifted according to Runstler and excludes the error term u_t. It can be seen that the generated cycle does indeed exhibit a shift with respect to ψ_t. However, the exactness of the shift can be disputed by a quick visual inspection even though error components which might have further increased the difference have not been added to the series. The loss of information as was previously pointed in equations (1.5) and (1.6). Figure 1.1 also is a confirmation of the inability that Runstler methodology cannot account for the model parameter ρ and cycle innovation sequence $\sum_{j=0}^{h-1} \rho^j \boldsymbol{T}^j \boldsymbol{\kappa}_{t+h-j}$.

[5]Notice that z_t converges closer to an exact shift as ρ converges.

Figure 1.1: Simulation of Baseline cycle (ψ_t) and its Angular Shifted version (z_t): $r = 10, \rho = 0.9, \sigma_\kappa^2 = 0.05$



In this paper, we seek to improve upon this sub-optimal filtering procedure so that a higher degree of information can be extracted from the data and the filtered cycle becomes closer to an exact shift of the baseline cyclical component.

1.2.2 An Exact An Exact Phase Shift Representation for Leading Cycles

This section will present a simple way to generate an exact shifted version of the common cycle. The section serves as introducing the main innovation of our paper which will then be applied to simulated and real data. But before moving on to the continuous time setting, let us develop the motivation for selecting the framework in continuous time. To do this, let us take as a starting point equation (1.4) where z_t is the series that is governed by the by a cycle that is leading the cyclical component in y_t by $(h+\delta)$ periods where h is an integer and $\delta \in [0, 1)$ is a fractional value. For the coinciding cycle component in y_t of equation (1.1), we choose the compact form of equation (1.3). The integer part of the leading cycle, can be found as a function of $\boldsymbol{\psi}_{t-1}$ by iteratively substituting equation (1.3) on to itslef $h+1$ times similar to what was done in equation (1.5). Hence, the coinciding and the h period leading cycle can both be presented as a function of ψ_{t-1} as follows

$$\begin{aligned} \boldsymbol{\psi}_t &= \rho \boldsymbol{T}(\lambda)\boldsymbol{\psi}_{t-1} + \boldsymbol{\kappa}_t \\ \boldsymbol{\psi}_{t+h} &= \rho^{h+1}\boldsymbol{T}^{h+1}\boldsymbol{\psi}_{t-1} + \sum_{j=0}^{h} \rho^j \boldsymbol{T}^j \boldsymbol{\kappa}_{t+h-j} \end{aligned} \tag{1.7}$$

The second line in equation (1.7) is the leading cycle governing z_t. This cycle will be leading the coinciding cycle ψ_t by h periods. However, to reach the exact leading

cycle of $\psi_{t+h+\delta}$, we need to introduce the concept of stochastic cycles in continuous time.

1.2.3 Stochastic cycles in Continuous Time

In this section we present the continuous time stochastic cycle so that we can get the result of a fractional leading cycle. The result can then be applied for the exact leading cycle of $\psi_{t+h+\delta}$. In Harvey (1990) and Harvey and Stock (1993) stochastic cycles used in business cycle analysis are given with detailed exposition for continuous time processes. We select the same description of these continuous time stochastic cycles for our UCM and then compare this framework to Runstler's method. To give a brief explanation, the continuous time cycle can be represented by a differential equation

$$\mathrm{d}\begin{bmatrix}\psi(t)\\ \psi^*(t)\end{bmatrix} = \begin{bmatrix}\log\rho & \lambda\\ -\lambda & \log\rho\end{bmatrix}\begin{bmatrix}\psi(t)\\ \psi^*(t)\end{bmatrix}\,\mathrm{d}t + \begin{bmatrix}\mathrm{d}\kappa_c(t)\\ \mathrm{d}\kappa_c^*(t)\end{bmatrix} \tag{1.8}$$

where ψ_t, ψ_t^* are continuous time cyclical components with error terms κ_t, κ_t^* being uncorrelated Wiener processes with identical variance of σ_c^*. We leave out the derivation to reach the discrete time representation of equation (1.8), which can be found in details of Harvey (1990). Using the same resource we only write down the general results

$$\begin{bmatrix}\psi_\tau\\ \psi_\tau^*\end{bmatrix} = \rho^{\delta_\tau}\begin{bmatrix}\cos\lambda\delta_\tau & \sin\lambda\delta_\tau\\ -\sin\lambda\delta_\tau & \cos\lambda\delta_\tau\end{bmatrix}\begin{bmatrix}\psi_{\tau-1}\\ \psi_{\tau-1}^*\end{bmatrix} + \begin{bmatrix}\kappa_\tau\\ \kappa_\tau\end{bmatrix} \tag{1.9}$$

where $\delta\tau$ is the fractional time span between two observation points and $\tau = 1, 2, 3, ...T$ is the ordering of irregular spaced intervals and $\kappa_{tau}, \kappa_{tau}^*$ are the discrete time representations of the uncorrelated Wiener processes in equation (1.8) and hence are also serially uncorrelated with $Cov(\kappa_\tau, \kappa_\tau^*)'$ being equal to the following matrix

$$Var\begin{bmatrix}\kappa_\tau\\ \kappa_\tau^*\end{bmatrix} = \frac{\sigma_c^2(1-\rho^{2\delta_\tau})}{\log^{-2}\rho}\boldsymbol{I}_2 \tag{1.10}$$

For simplicity of representation assuming that our data is regularly spaced we do not need to show the subscript τ and set $\delta_\tau = \delta$ in equations (1.9) and (1.10). Referring to section 1.2.1 and following the derivations inPelagatti (2006), we can have a representation of the leading cycle $\psi_{t+h+\delta}$ which has a fractional lead on ψ_{t+h} shown in the following way

$$\boldsymbol{\psi}_{t+h+\delta} = \rho^\delta \boldsymbol{T}^\delta \boldsymbol{\psi}_{t+h} + \boldsymbol{\upsilon}_{t+h+\delta} \tag{1.11}$$

where the error term $\boldsymbol{\upsilon}t + h + \delta$ is a white noise error term with zero mean vector and with variance covariance matrix of

$$\mathbb{E}(\boldsymbol{v}_t \boldsymbol{v}_t^{\mathrm{T}}) = \frac{\sigma_c^2(1-\rho^{2\delta})}{\log \rho^{-2}} \boldsymbol{I}_2 \tag{1.12}$$

which is equation (1.10) with $\delta_\tau = \delta$.

1.2.4 General Results of Continuous Time Dynamical Systems

In this section we present the building blocks of the discrete time model that are obtained from the general results of stochastic dynamical systems. This is because our estimation and inference will be made on a discrete time representation of a continuous time model. These general results are based on the relevant chapters in Harvey (1990) and Harvey and Stock (1993).

We assume that our filtering procedure is conducted on a first order linear process in vector form which is a stochastic differential equation

$$\mathrm{d}\boldsymbol{\alpha}(t) = \boldsymbol{A}\boldsymbol{\alpha}(t)\,\mathrm{d}t + \boldsymbol{R}\,\mathrm{d}\boldsymbol{\mu}(t) \tag{1.13}$$

with A and R being (mxm) and (mxg) non-random matrices, μ is a g-dimensional multivariate Wiener process with covariance matrix equivalent to

$$\mathbb{E}[\,\mathrm{d}\boldsymbol{\mu}(t)\,\mathrm{d}\boldsymbol{\mu}(t)^{'}] = \boldsymbol{Q}\,\mathrm{d}t$$

Equation (1.13) is the continuous time analogue of a state vectors transition equation in discrete time as shown in Harvey (1990). Equation (1.13) has an important interpretation in the sense that it satisfies the following stochastic integral equation

$$\boldsymbol{\alpha}(t) - \boldsymbol{\alpha}(0) = \int_0^t \boldsymbol{A}\boldsymbol{\alpha}(t)\,\mathrm{d}r + \int_0^t \boldsymbol{R}\,\mathrm{d}\boldsymbol{\mu}(r) \tag{1.14}$$

For any initial value $\alpha(0)$, when the above equation is solved, will give the following stochastic difference equation

$$\boldsymbol{\alpha}(t) = \boldsymbol{\alpha}(0)e^{\boldsymbol{A}t} + \int_0^t e^{\boldsymbol{A}(t-z)}\boldsymbol{R}\,\mathrm{d}\boldsymbol{\mu}(z) \tag{1.15}$$

Equation (1.15) in turn will satisfy the stochastic difference equation

$$\boldsymbol{\alpha}(t) = \boldsymbol{\alpha}(t-1)e^{\boldsymbol{A}t} + \int_0^t e^{\boldsymbol{A}(t-z)}\boldsymbol{R}\,\mathrm{d}\boldsymbol{\mu}(z) \tag{1.16}$$

The important advantage of working with dynamical systems in continuous time is that it is able to account for also irregularly spaced data. So, if there is a set of irregularly spaced time sequence t_τ with order subscript $\tau = 1, 2, ..., T$ and the increment between two time points being $\delta_\tau = t_\tau - t_{\tau-1} \geq 0$, then the relation between state vectors at time points t_τ and $t_{\tau-1}$ is

$$\boldsymbol{\alpha}(t_\tau) = e^{\boldsymbol{A}\delta_\tau}\boldsymbol{\alpha}(t_{\tau-1}) + \int_0^{\delta_\tau} e^{\boldsymbol{A}(\delta_\tau - z)}\boldsymbol{R}\,\mathrm{d}\boldsymbol{\mu}(t_{\tau-1} + z) \tag{1.17}$$

Equation (1.17) will provide us the basis of to construct the discrete time representation of the state transition equations. For this, one carries out the following substitutions; $\boldsymbol{T}_\tau = e^{\boldsymbol{A}\delta_\tau}, \boldsymbol{\alpha}_\tau = \boldsymbol{\alpha}(t_\tau)$ and also considers the last term on the right hand side of equation (1.17) as the error term which follows a white noise process with mean vector $\mathbf{0}$ and a covariance matrix[6]

$$\boldsymbol{Q}_\tau = \int_0^{\delta_\tau} e^{\boldsymbol{A}(\delta_\tau - z)}\boldsymbol{R}\boldsymbol{Q}\boldsymbol{R}' e^{\boldsymbol{A}'(\delta_\tau - z)}\,\mathrm{d}z \tag{1.18}$$

to reach the following discrete time representation of the state transition equations

$$\boldsymbol{\alpha}_\tau = \boldsymbol{T}_\tau\boldsymbol{\alpha}_{\tau-1} + \boldsymbol{\mu}_{\tau-1} \tag{1.19}$$

The state vector $\boldsymbol{\alpha}(t)$ can never precisely be observed. The source of information can only be retrieved using observations from the data. These observations in continuous time shown as yt can be defined by a simple dynamic system of equations written compactly in the following form

$$\boldsymbol{y}(t) = \boldsymbol{Z}\boldsymbol{\alpha}(t) + \boldsymbol{\varepsilon}(t) \tag{1.20}$$

with $\boldsymbol{Z}$ being a row vector or matrix according to whether $\boldsymbol{y}(t)$ is a single or a vector of continuous time processes. $\boldsymbol{Z}$ is a factor loading system matrix that assigns which states in $\boldsymbol{\alpha}(t)$ determine the evolution of each process in $\boldsymbol{y}(t)$, $\boldsymbol{\varepsilon}_t$ follows a multivariate white noise process in continuous time with covariance matrix $\boldsymbol{H}$ and uncorrelated with the error term in equation (1.13) such that

$$\mathbb{E}\left[\int_q^p \mathrm{d}\boldsymbol{\mu}(t)\int_m^t \mathrm{d}\boldsymbol{\varepsilon}(t)\right] = 0 \tag{1.21}$$

where $\forall p < q$ and $l < m$.

[6]To reach the covariance matrix in equation (1.18), the result of the multivariate Wiener process are required. For derivations see Harvey (1990) and Harvey and Stock (1993)

In this study we assume that $\boldsymbol{y}(t)$ is only composed of stock variables. Therefore, the continuous time representation in (1.20) can be represented in discrete time form in the following way

$$y_\tau = \boldsymbol{Z}_\tau \boldsymbol{\alpha}_\tau + \boldsymbol{\varepsilon}_\tau \tag{1.22}$$

where $\boldsymbol{\varepsilon}_\tau$ is a vector of white noise error terms and uncorrelated with the error term $\boldsymbol{\mu}_\tau$ in equation (1.19).

1.2.5 Phase Shift Mechanism of Stock Variables

In this section we will introduce the phase shift mechanism which can generate shifted versions of the exact stochastic cycle. Let us consider firstly, that we have m number of stock variables each governed by a stochastic cycle of the form introduced in section 1.2. The evolution of each cycle that end in time lags of $\delta_1, \delta_2, ..., \delta_m$ can be shown as follows

$$\begin{aligned}
\boldsymbol{\psi}(t+\delta_1) &= \rho^{\delta_1}\boldsymbol{T}(\lambda\delta_1)\boldsymbol{\psi}(t) + \boldsymbol{\kappa}(\delta_1) \\
\boldsymbol{\psi}(t+\delta_2) &= \rho^{\delta_2}\boldsymbol{T}(\lambda\delta_2)\boldsymbol{\psi}(t) + \boldsymbol{\kappa}(\delta_2) \\
&\cdots \\
\boldsymbol{\psi}(t+\delta_m) &= \rho^{\delta_m}\boldsymbol{T}(\lambda\delta_m)\boldsymbol{\psi}(t) + \boldsymbol{\kappa}(\delta_m)
\end{aligned} \tag{1.23}$$

where each positive time lag associates with a specific lag number reflecting each distinct cycle. These lags are by no means are an index of order magnitude. To be able to represent a system of stochastic cycles in discrete time state space form, it is required to also provide covariances between the disturbances corresponding to each of the cycles included in equation (1.23). From the equations (1.17)-(1.19) and using some algebra, the covariance between any two stock cycle errors can be summarized as

$$\mathbb{E}[\kappa(\delta_k)\kappa(\delta_l)] = \mathbb{E}[\kappa^*(\delta_k)\kappa^*(\delta_l)] = \frac{\sigma_c^2(1-\rho^{2\delta})}{\log \rho^{-2}} \quad \text{with } \delta = \delta_k \wedge \delta_l, \tag{1.24a}$$

$$\mathbb{E}[\kappa(\delta_k)\kappa^*(\delta_l)] = \mathbb{E}[\kappa^*(\delta_k)\kappa(\delta_l)] = 0. \tag{1.24b}$$

Using the information provided in section 1.2, the set of transition equations that is required can be presented in the following way

$$\begin{aligned}
\varphi_t &= \rho^{\delta}[\cos\lambda\delta \;\; \sin\lambda\delta]\psi_{t-1}^{(0)} + \upsilon_t \\
\boldsymbol{\psi}_t^{(0)} &= \rho\boldsymbol{T}\boldsymbol{\psi}_{t-1}^{(0)} + \boldsymbol{\kappa}_t \\
\psi_t^{(1)} &= \psi_{t-1}^{(0)} \\
&\cdots\cdots \\
\psi_t^{(h+1)} &= \psi_{t-1}^{(h)}
\end{aligned} \tag{1.25}$$

with $Var(\kappa_t)$ is found using the diagonal elements in equation (1.10) and that $Var(\upsilon_t) = Cov(\upsilon_t, \kappa_t)$ is obtained from (1.24a).

Our focus is designing a model that generates exact shifted versions of a common stochastic cycle. The study will provide the theoretical and empirical evidence that equation (1.25) does indeed provide such results. This can easily be checked through the following setting

$$\begin{aligned} \psi_{t-h} &= \psi_t^{(h+1)} = \psi_{t-h}^{(1)} = \psi_{t-h-1}^{(0)} \\ \phi_{t-h} &= \varphi_t = \psi_{t+\delta}^{(1)} = \psi_{t+\delta-1}^{(0)} \end{aligned} \tag{1.26}$$

where the identities in equation (1.26) are h period shifted cycles. More precisely, the setting allows ψ_{t-h} and ϕ_{t-h} to be integer and fractional shifted versions of the stochastic cycle $\psi_{t-1}^{(0)}$.

Using this setting together with the continuous time delay parameter $r > 0$ with $h = \lfloor |r| \rfloor$ and $\delta = |r| - h$, we can apply the stock variable transition equations of (1.23), (1.25) to reconstruct y_t and z_t of equation (1.4) such that

$$\begin{aligned} y_t &= \psi_t^{(h+1)} + \epsilon_t \\ z_t &= \beta\phi_t + \mu_t \end{aligned} \tag{1.27}$$

where β is a scaling factor that aligns the identities in terms of magnitude.

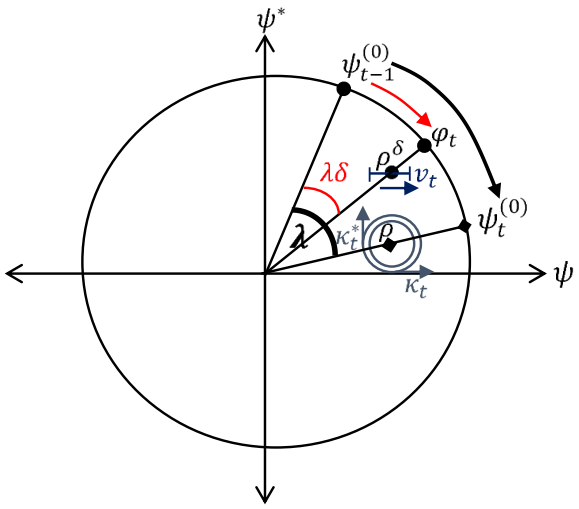
A visual comparison between the introduced method in this paper and the one in Runstler's can provide a general understanding of how the phase shift mechanism works. Figure 1.2 allows a projection of the two competing methods. In Runstler's method the shift is a movement of the original cycle along the circumference of the circle with angular speed of lambda, which ends at the dark square $\psi_t^{(0)}$ (dark arrow). The method fails to take into account the parameter ρ which is an estimated parameter and if had been accounted for would have resulted in the point moving closer to the origin since $\rho < 1$.[7] In our methodology, the movement covers a shorter distance (red arrow) since the angular speed takes into account the fractional shift δ.Further, since the methodology accounts for the parameter ρ^δ, loss of information has decreased considerably with respect to Runstler's case.

1.2.6 Continuous Time Stochastic Trends

Many macroeconomic indicators contain trends that needs to be extracted so that BC analysis can be carried out correctly. This is an important element to account for the underlying dynamics of economic data. In almost all BC analysis that apply UCM, one of the main components of economic variable is a stochastic trend, as it

[7]This was previously stated as the loss of information since the dynamics of the lead-lag relationship of the cycles are lost during the extraction procedure.

Figure 1.2: Geometric Comparison of Phase Shift Methods



Note: $T(\cdot)$ rotates the cycle (sinusoid) along the circle circumference, ρ dampens the effect bringing the final point inside the circle boundary and $(\kappa_t\kappa_t^*)$ are a pair of random errors which are the stochastic components of $\psi_{t-1}^{(0)}, \psi_t^{(0)}, \phi_t$

accommodates the randomness of trends (more specifically their slopes/steepness) associated to it. Since our structural framework is in continuous time, the form defined in Harvey (1990), Harvey and Stock (1993) is presented such that

$$\mathrm{d}\begin{bmatrix}\mu(t)\\ \beta(t)\end{bmatrix} = \begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}\begin{bmatrix}\mu(t)\\ \beta(t)\end{bmatrix}\mathrm{d}t + \begin{bmatrix}\mathrm{d}\zeta(t)\\ \mathrm{d}\eta(t)\end{bmatrix} \tag{1.28}$$

with $\mu(t)$ and $\beta(t)$ showing the evolution of the trend and slope, $\zeta(t)$ and $\eta(t)$ following continuous time white noise processes with variances σ_κ^2 and σ_η^2 respectively. Similar to the continuous time stochastic cycle, estimation and other statistical inference on the model is done using the discrete time representation. Hence the discrete form of (1.28) is

$$\begin{bmatrix}\mu_\tau\\ \beta_\tau\end{bmatrix} = \begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}\begin{bmatrix}\mu_{\tau-1}\\ \beta_{\tau-1}\end{bmatrix}\begin{bmatrix}\zeta_\tau\\ \eta_\tau\end{bmatrix} \tag{1.29}$$

with μ_τ and β_τ both following random walk processes.[8] Setting $\delta_\tau, \sigma_\zeta^2 = 0$, and the initial value of the slope, $\beta_0 = 0$, the trend component will be following an integrated random walk. According to Pelagatti (2015), the formulation of the trend component in this way is fruitful if cyclical components are in the model.

[8]The derivation of the continuous time trend component to its discrete time representation is left out in the study. For further details see Harvey (1990).

1.3 Simulation Exercise and Grid Search

Once the model is built in state space form and since it has a discrete time representation, the Kalman filtering and smoothing recursions can be carried out. The Kalman recursions require initialization of the state vector and its covariance matrix. While Bayesian techniques can be employed to come up with initial values for the state vector, the procedures we have opted to use is through finding the unconditional mean and variances of the state vector. The procedures to find these values for both the stock and flow variable transition equations have been provided in detail in section B and C of the appendix.

In this section a simulation exercise will be conducted, also a numerical maximum likelihood optimization carried out. This is done due to the fact that the Kalman filter produces filtered series of the error values and its covariance matrix which allows a Gaussian likelihood function to be built. From the likelihood function, it is possible we further carry out the estimation of the model parameters using numerical optimization. We aim to show how the likelihood changes over the model's parameter space considering that the phase shift (delay) parameter r is continuous in time. The parameter r is composed of a fractional part δ and an integer part h, which is an important factor that determines the dimensions of the model we estimate in equation (1.25). However, the aim of singling out r in the parameter space and allow it to move continuously, is to understand the behavior of the likelihood function as the delay parameter moves between fractional and integer values.

1.3.1 Simulating time series from State Space Form

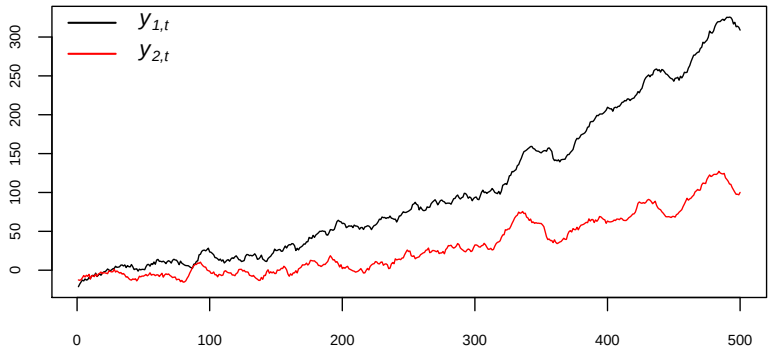
To simulate observations, the simulation smoother based on the works of Durbin and Koopman (2012) was applied. The smoother is an algorithm which first draws a random sequence of disturbances.

It is important to be able to check the model and see if it is able to create random sets of parameter values. Since the model produces shifted versions of coinciding cycles, we will also be able to assess whether the correlation structure amongst the shifted and coinciding cycle is correctly estimated. As is stated in Jarociński (2015), the simulation smoother is a method that provides ways to examine the correlation between unobservable variables. This takes us to the shifted cycles that compose the elements of the unobservable state vector.

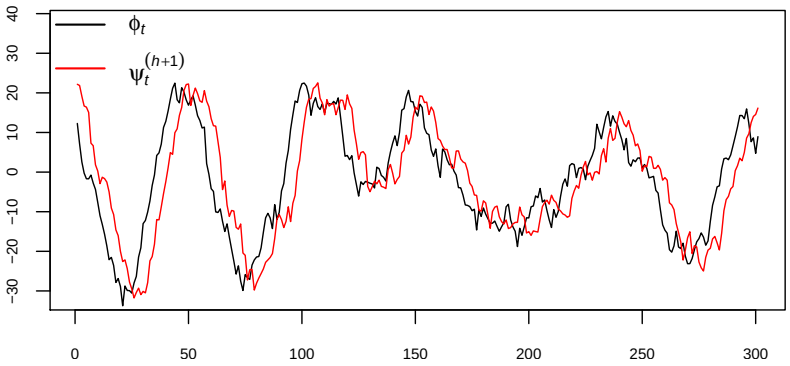
Figure 1.3 show the simulated series and the state cycles that are generated by passing through the simulation smoother on the state space form of stock variable systems. The simulated series on the left hand side of the panel include a stochastic trend which are then filtered out to reach the cyclical components on the right hand side panel. Let us consider that the cycle in $y_{1,t}$ is associated with ϕ_t and the cyclical component in $y_{2,t}$ is associated with $\psi_t^{(h+1)}$. In this setting, $y_{1,t}$ is a leading indicator of ψ_{t-1} by a fractional shift δ, while $y_{2,t}$ is leading ψ_{t-1} by an integer shift of $h+1$.

Figure 1.3: Simulated States and Observations based on Simulation Smoother of Durbin and Koopman (2012)

(a) Simulated Series



(b) Cyclical Components



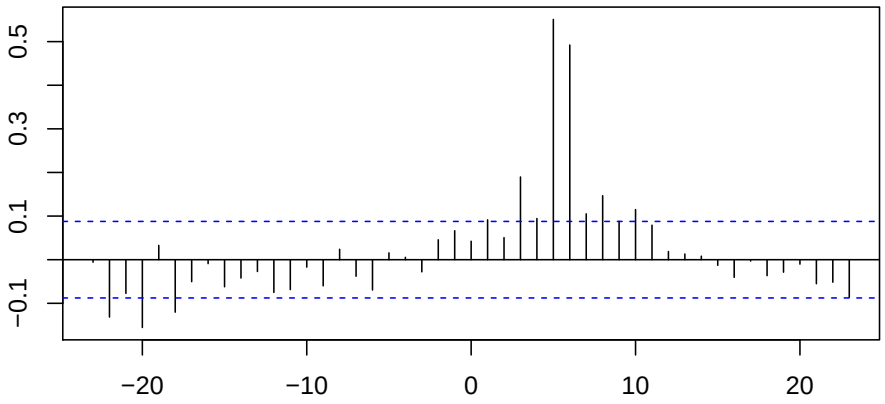
Note: The scaling factor β introduced in equation (1.27) for the stock cycle ϕ_t is set to 1.

One way to understand the level of phase shifts between two analyzed series is to check the cross-correlation function of the $y_{,t}$ pair. In this case the cross correlation between the differenced series $\Delta y_{1,t}$ and $\Delta y_{2,t}$ can give an initial starting point between the relationship of associated cyclical components. The correlation function can be expressed in the following way

$$\begin{aligned}\hat{\rho}_{\Delta y_{1,t},\Delta y_{2,t}}(j) &= \frac{\widehat{Cov}(\Delta y_{1,t}, \Delta y_{2,t-j})}{\sqrt{\widehat{Var}(\Delta y_1)\widehat{Var}(\Delta y_2)}} \\ &= \frac{n^{-1}\sum_{t=j+1}^{n}(\Delta y_{1,t} - \Delta\bar{y}_{1,1:n})(\Delta y_{2,t} - \Delta\bar{y}_{2,1:n})}{\hat{\sigma}_{\Delta y_1,1:n}\hat{\sigma}_{\Delta y_2,1:n}}\end{aligned} \tag{1.30}$$

where $\widehat{Cov}(\cdot), \widehat{Var}(\cdot)$ indicate the sample variance and covariance $\Delta\bar{y}_1, \Delta\bar{y}_2$ the sample means and $\hat{\sigma}_{\Delta y_1}, \hat{\sigma}_{\Delta y_2}$ are the sample standard deviations of $\Delta y_1, \Delta y_2$. If we were to plot the cross correlation function $\hat{\rho}(j)$ against each associated lag j where the sequence $j = -20, -19, ..., 0, 19, 20$; it would be possible to see at which lag the first differenced series have the strongest relation. Figure 1.4 shows the correlogram of the simulated series sample cross correlations. The differenced series for stock variables shows the strongest relation at lags 5 ($\hat{\rho}(5)$) and 6 ($\hat{\rho}(6)$). The cyclical pattern of the correlations is evident and implies a phase shift between the cyclical components. If both the differenced series of $\Delta y_{1,t}$ and $\Delta y_{2,t}$ were on the same phase, the highest cross correlation would have been recorded at lag 0. Hence, it is possible to say that a phase shift is present amongst cycles of the simulated data $y_{1,t}$ and $y_{2,t}$.

Figure 1.4: Cross Correlogram of Simulated Observation Series, First Differenced ($\widehat{\rho}(j)$)



Note: $y-axis$ represents the calculated values for $\widehat{\rho}(j)$), $x-axis$ represents the lag j. Blue lines indicate the 95% confidence intervals.

1.3.2 Log Likelihood Function and Grid Search for Simulated Data

One of the advantages of the Kalman filter is its ability to also provide a loglikelihood function. The Kalman filters recursive application produces for every t, the linear

predictor of the state vector a_t and the one step ahead forecast of observations $\widehat{y}_{t|t-1}$ based on the conditional information from the initial point up to the available information at each of the time points. Each of these sequences can be shown as $\{\boldsymbol{a}_{2|1}, \boldsymbol{a}_{3|2}, ..., \boldsymbol{a}_{t|t-1}\}$ and $\{\widehat{\boldsymbol{y}}_{2|1}, \widehat{\boldsymbol{y}}_{3|2}, ..., \widehat{\boldsymbol{y}}_{t|t-1}\}$ respectively.[9] Obtaining the sequence for $\widehat{\boldsymbol{y}}_{t|t-1}\}$ where $t = 2, ..., n$ simply requires using the first line of equation (1.19) as a linear prediction operator $\mathbb{P}[\boldsymbol{y}_t|\mathbb{I}_{t-1}]$ and replace α_t with appropriate values in the sequence of $\{\boldsymbol{a}_{t|t-1}\}$. Obtaining the sequence of $\{\boldsymbol{a}_{t|t-1}\}$ on the other hand requires an extra update scheme, because with every new prediction error $(\boldsymbol{y}_t - \widehat{\boldsymbol{y}}_{t|t-1})$ provide a clearer picture of the unobserved state. Hence, these errors are input used for the update scheme to correct the next period filter. At this step, the prediction errors can be denoted as $\boldsymbol{v}_t$ and its variance as $\mathbb{E}[\boldsymbol{v}_t \boldsymbol{v}_t^{\mathrm{T}}] = \boldsymbol{F}_t$. Finally, the assumption that the joint density of the data $p(\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_n)$ is a multivariate normal distribution, corresponds to the variable $\boldsymbol{y}_t$ to have a conditional density $p(\boldsymbol{y}_t|\mathbb{I}_{t-1})$ following a normal distribution with mean $\widehat{\boldsymbol{y}}_{t|t-1}$ and variance $\boldsymbol{F}_t$ so that $\boldsymbol{y}_t|\mathbb{I}_{t-1} \sim N\left(\widehat{\boldsymbol{y}}_{t|t-1}^{(\theta)}, \boldsymbol{F}_t^{(\theta)}\right)$.[10] The loglikelihood function can then be computed as

$$\begin{aligned} lnL(\boldsymbol{\theta}; \boldsymbol{y}|\mathbb{I}) &= \sum_{t=1}^{n} lnf(\boldsymbol{\theta}; \boldsymbol{y}_t|\mathbb{I}_{t-1}) \\ &= -\frac{nj}{2} ln(2\pi) - \frac{1}{2} \sum_{t=1}^{n} \left(ln \left| \boldsymbol{F}_t^{(\boldsymbol{\theta})} \right| + \left(\boldsymbol{v}_t^{(\boldsymbol{\theta})} \right)^{\mathrm{T}} \boldsymbol{F}_t^{(\boldsymbol{\theta})} \boldsymbol{v}_t^{(\boldsymbol{\theta})} \right) \end{aligned} \tag{1.31}$$

where j is the number of variables in $\boldsymbol{y}_t$ and the superscripts in $\boldsymbol{v}_t^{(\boldsymbol{\theta})}, \boldsymbol{F}_t^{(\boldsymbol{\theta})}$ indicate that the values are obtained by passing the Kalman filter on the model equation (1.19) where the parameter vector $\boldsymbol{\theta}$ was assigned. Once the log-likelihood function is constructed, we can retrieve maximum likelihood parameter estimates $\widehat{\boldsymbol{\theta}}_{ML}$ by numerical optimization of the log-likelihood based on an objective function.

One analytical approach in finding $\widehat{\boldsymbol{\theta}}_{ML}$ is by a grid search. The main reason of implementing this approach is that when one knows possible values for a parameter within the vector $\widehat{\boldsymbol{\theta}}_{ML}$, in our case this is the delay parameter r, an extensive search can be done over the possible region of values r can take. Through this repeated iteration, the researcher can detect the r value that corresponds to the largest likelihood.

While in section 1.5, using real data, $\widehat{\boldsymbol{\theta}}_{ML}$ will be computed through a general numerical optimization, the motive for the grid search in this section is twofold: first to investigate the log-likelihood behavior when the estimated delay moves along the grid level which includes integer and fractional points; secondly to see if the maximum level for the likelihood function attained along the grid corresponds to the same level found for the correlation amongst shifted cycles. Hence, we are able to analyze if a maximum likelihood is attainable and if so, is the associated delay parameter estimate realistic.

[9] The information on Kalman filter in this section includes only the most essential parts relevant to our study and hence is incomplete. For a complete description see Harvey (1990), Durbin and Koopman (2012), Pelagatti (2015).

[10] $\mathbb{I}_{t-1}$ is the information set that includes all observations up to and including $t-1$ so that $\mathbb{I}_{t-1} = (\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_{t-1})'$.

To start the search, set the parameter space $\Theta = \{r, \boldsymbol{\theta}_0^-\}$ where r defines the grid points; $r = \{1, 1.01, 1.02, ..., 8\}$ and the remaining parameters in equation (1.19) as $\boldsymbol{\theta}_0^- = (e^{\varpi_1}, e^{\varpi_2}, e^{\varpi_3}, e^{\varpi_4}, e^{\varpi_5}, g(\varpi_\rho, 0, 1), g(\varpi_p, 24, 96), \beta_0, e^{\varpi_{10}})'$ includes the initial values of all the model parameters except r.[11] After setting the values, using the construct in equation (1.31), the log-likelihood equation for the grid search can be written as

$$\begin{aligned} lnL(\Theta; \boldsymbol{y}|\mathbb{I}) &= \sum_{t=1}^{n} lnf(\Theta; \boldsymbol{y}_t|\mathbb{I}_{t-1}) \\ &= -\frac{nj}{2} ln(2\pi) - \frac{1}{2} \sum_{t=1}^{n} \left(ln \left| \boldsymbol{F}_t^{(\Theta)} \right| + \left(\boldsymbol{v}_t^{(\Theta)} \right)^{\mathrm{T}} \boldsymbol{F}_t^{(\Theta)} \boldsymbol{v}_t^{(\Theta)} \right). \end{aligned} \tag{1.32}$$

The grid search then is the evaluation of $L(\Theta; \boldsymbol{y}|\mathbb{I})$ for all the possible grid combinations of the values in $\{r\}$ and the arbitrary constant vector $\boldsymbol{\theta}_0^-$.

Previously in Section 1.3.1 it was found that the first differenced simulated data of stock variables registered the highest cross correlation at lag 5. This value can be selected together with the addition of an arbitrary initial value δ_0 so that the initial value of r becomes $r_0 = |j_{max}(\widehat{\rho})| + \delta_0$. By setting the value of r_0 this way, one can take it as the initial value for calculating the maximum likelihood estimation of the model parameters $\widehat{\boldsymbol{\theta}}_{ML}$. We provide in this section of the study, the reasoning of why r_0 is set to $|j_{max}(\widehat{\rho})| + \delta_0$. The conclusion will be that it reflects the best initial estimate given the information content of any observation set.

Figure 1.5 plots the value of the negative log-likelihood with respect to the delay parameter r. In this framework, the grid search is a minimization problem. To gain more insight, it is more appropriate to examine per unit interval and decomposing the delay parameter into its integer part h and the fractional part δ. It can be seen that the likelihood records higher value at the fractional points, topping at $\delta = 0.5$ and declining as it moves away to either integer values of r which lie towards the left or right end of a unit interval corresponding to $\delta = 0$ or 1. The reason for such a behavior can be traced to the stochastic cycle in equation (1.11).

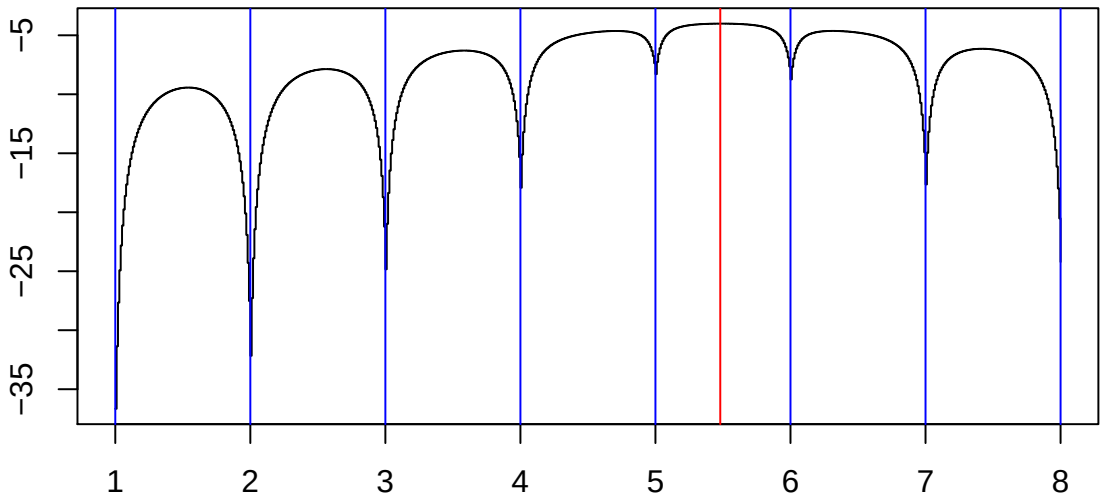
To make matters clearer, let us consider only the first element of $\boldsymbol{\psi}_{t+h+\delta}$, which is denoted as $\psi_{t+h+\delta}$ and is the projection of the compact cycle on to the $x - axis$. To understand the behavior of the log-likelihood, we will need to see the link between $\psi_{t+h+\delta}$ and the stochastic cycles ψ_t and ψ_{t+h}. When $\delta = 0$, the variance of the error υ_δ associated with $\psi_{t+h+\delta}$ is zero so that $\mathbb{V}(\upsilon_\delta) = 0$. This implies that the introduced cycle is not stochastic but deterministic. Notice that υ_δ is the discrete time representation of $\int_0^\delta \kappa(s)ds$ and $\delta = 0$ means that the integral collapses to zero.

On the other extreme, if $\delta = 1$ then the introduced cycle $\psi_{t+h+\delta}$ is identical to both cycles in equation (1.7). In this case the error term υ_δ and the error term κ_t are identical such that $\rho(\upsilon_{t+h+\delta}\kappa_t) = 1$ and the introduced cycle is a shifted version of the common stochastic cycle. Notice also that the continuous time representation of υ_δ when $\delta = 1$ is $\int_0^1 \kappa(s)ds$ which is in turn κ_t.

[11] $g(x, a, b) = a + (b - a)/(1 + e^{-x})$ is a logit function, the expressions ϖ_ρ and ϖ_p are transformed model parameters explained in Appendix.

Lastly, let us also consider the case where $0 < \delta < 1$. More specifically, υ_δ is the discrete time representation of $\int_0^\delta \kappa(s)ds$ which is correlated with $\int_0^1 \kappa(s)ds$. Therefore, the system indeed introduces a distinct stochastic cycle because $0 < \rho(\upsilon_{t+h+\delta}\kappa_t) < 1$ with $\mathbb{V} > 0$. Since the stochastic cycle introduces new information, this raises the degrees of freedom and creates a more flexible model. The more δ is closer to 0.5 the more distinct the new introduced cycle $\psi_{t+h+\delta}$ will be from ψ_t. The distinction of the cycle and in turn the flexibility of the model will reach maximum at $\delta = 0.5$ since at that point $\rho_{min}(\upsilon_\delta\kappa_t)$. The concave shape shape of the negative log-likelihood at unit intervals in Figure 1.5 reflects the fact that the model is less rigid when r takes non-integer values compared to if r was an integer. At every unit interval where $\delta = 0.5$ the likelihood peaks and then descends towards left and right as r moves to integer points where $\delta = 0 or 1$. In addition, Figure 1.5 indicates that there is only a single interval where the log-likelihood reaches maximum and hence a global maximum can be found. Moreover, the value at which $lnL(\Theta; \boldsymbol{y}|\mathbb{I})$ achieves maximum along the grid is consistent with the initial value r_0 that was determined using the information achieved from the data. Hence, it is reasonable to assume that our model construct will be providing reliable estimates when numerical optimization is carried out to find $\widehat{\boldsymbol{\theta}}_{ML}$ through numerical optimization.

Figure 1.5: Grid Search of Simulated Stock Data



Note: $y - axis$ represents negative log-likelihood values, $x - axis$ is the grid for r. The red line indicates the maximum value of $-lnL(\Theta; \boldsymbol{y}|\mathbb{I})$ recorded at $r = 5.5$.

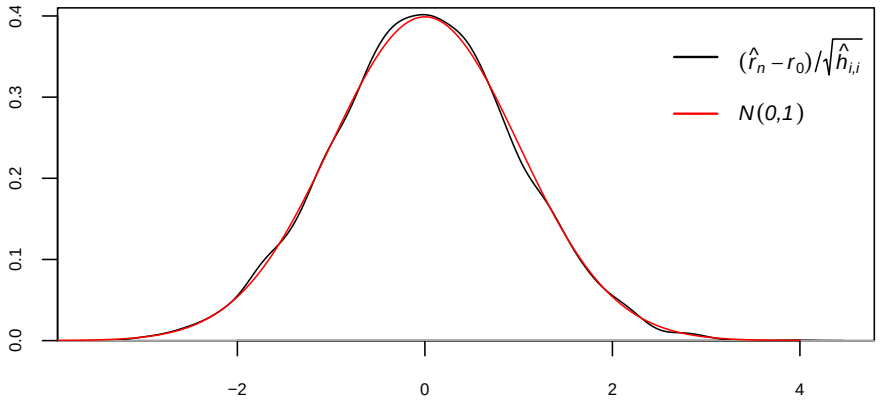
1.4 Sampling Distribution of the Phase Shift Parameter

We evaluate the Maximum Likelihood estimator of the delay parameter denoted as $\widehat{r}_{ML}$ using Monte Carlo simulations. We aim to assess whether $\widehat{r}_{ML}$ is consistent and asymptotically normal. From the general asymptotic concepts for the ML estimator, as $n \longrightarrow \infty$ we can write

$$\sqrt{n}(\widehat{r}_{ML} - r_0) \xrightarrow{d} N\left(0, I_{i,i}^{-1}(r_0)\right) \tag{1.33}$$

where $I_{i,i}^{-1}(r_0)$ is the theoretical variance of $\widehat{r}_{ML}$ and corresponds to the diagonal element associated with r_0 in the inverse of the Fisher Information denoted as $\boldsymbol{I}^{-1}(\boldsymbol{\theta}_0)$ for the vector of model parameters $\boldsymbol{\theta}$. The variance-covariance matrix of the Maximum Likelihood estimator of the parameter vector $\widehat{\boldsymbol{\theta}}_{ML}$ is denoted as $\boldsymbol{I}^{-1}(\theta_{ML})$. It is asymptotically valid to replace the theoretical variance-covariance matrix with its estimator $\mathbb{H}^{-1}(\widehat{\boldsymbol{\theta}}_{ML})$ which is the inverse Hessian of the negative mean log-likelihood evaluated at $\widehat{\boldsymbol{\theta}}_{ML}$. Each diagonal element in $\mathbb{H}^{-1}(\widehat{\boldsymbol{\theta}}_{ML})$ constitutes the estimate of the asymptotic variance of the estimators in $\widehat{\boldsymbol{\theta}}_{ML}$. For $\widehat{r}_{ML}$, the estimate of its asymptotic variance will be the corresponding diagonal element in $\mathbb{H}^{-1}(\widehat{\boldsymbol{\theta}}_{ML})$ defined here as $H_{i,i}^{-1}(\widehat{r}_{ML})$. Figure 1.6 compares the theoretical sampling distribution with density distributions obtained by 1000 Monte Carlo iterations for 500 simulated observations. The results shows that the probability density of the estimator aligns to asymptotic normality.

Figure 1.6: Theoretical Distribution and the Sampling Distributions of the Maximum Likelihood Estimator $\widehat{r}_n$ based on 1000 Monte Carlo iterations



Note: $y - axis$ constitutes the density; $x - axis$ constitutes $\sqrt{n}H(\widehat{r}_n)^{0.5}(\widehat{r}_n - r_0)$ with $r_0 = 2.5$.

The estimates for $\widehat{r}_{ML}$ and their associated standard error estimates $\sqrt{H_{i,i}^{-1}(\widehat{r}_{ML})}$ are averaged over the $n = 1000$ iterations so that $\overline{\widehat{r}}_n = (\widehat{r}_1 + \widehat{r}_2 + ... + \widehat{r}_n)/1000$ and $\sqrt{\overline{H}_{i,i}^{-1}(\widehat{r}_n)} = (\sqrt{H_{i,i}^{-1}(\widehat{r}_1)} + ... + \sqrt{H_{i,i}^{-1}(\widehat{r}_n)})/1000$. Table 1.1 presents these averaged values which are found to lie close to their population equivalents with $\overline{\widehat{r}}_n = 2.543$ and $\sqrt{\overline{H}_{i,i}^{-1}(\widehat{r}_n)} = 1.009$.

Table 1.1: MC averages of $\widehat{r}_n$ estimates and Standard Errors for $n = 1000$ iterations

n	$\overline{\widehat{r}}_n$	Av.St.Error
1000	2.543	1.009
	r_0	Asym. St.Error
-	2.500	1.000

1.5 Data and Empirical Results

The data encompasses three largest EU economies; France, Germany, Italy's industrial production (IP) and Order book series. All the data were collected from the OECD database.[12] It is generally known that IP cycle is a good proxy for the BC. More specifically, the cyclical component is either taken usually as a coinciding indicator or a slightly leading indicator of the BC. The order book series unlike the IP, is not an index but based on a survey data. It is predominantly seen as an indicator that tries to detect turning points of the cycle. Because of the structured questions asked in the survey that constructs this indicator, it is forward looking and hence it should lead the BC. Therefore, when putting both cycles in their state space form, we assume the cyclical component of the Order Book series is the leading cycle with respect to the cyclical component of IP. While the study can be extended to include other macro indicators as in Valle e Azevedo et al. (2006) or even financial indicators as in Koopman et al. (2016), we only concentrate in results of the two macroeconomic indicators. The observations are at monthly frequency with the data period spanning from January 1979 to March 2019 and all data are seasonally adjusted. No transformations such as taking logarithms or first differencing have been done and the application is carried on raw data.

Figure 1.7 shows the time series that are to be analyzed. It is obvious that the IP series exhibits a trend component, while the trend for Order Book is not so evident. Nevertheless, and integrated random walk as explained in equation (1.29) is selected as a trend specification for all the macroeconomic series. Comparing country series, German IP shows the most pronounced upward trend amongst the three countries. With the wake of the global downturn and a sharp decline in industrial production for three of the countries, only German IP series has regained ground and shows a comparably more persistent trend. Regarding the business tendency survey data of the Order book, the similarity of peak and downturns timings for three of the countries is evident. This also reflects the similar risk exposure and market expectations of economic agents in the EU.

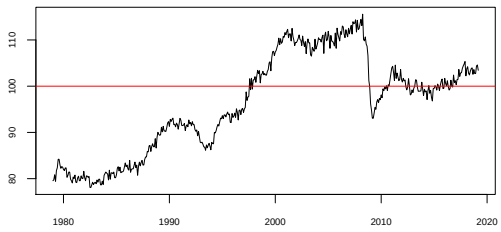
As was done for the simulated data in section 1.3.1, the first differences of IP and OB series cross correlations are examined. The highest lag cross-correlations for each country series pair will be taken with and addition of 0.5 to be added to reach the initial delay parameter r_0. This is required to conduct, the numerical optimization of the log-likelihood. The log-likelihood is constructed through the Kalman recursions as explained briefly in Section 1.3.1 for every time point t. Azavedo (2006) in this context, provide detailed explanation of this optimization process for both the univariate and multivariate UCM. During the optimization, parameter values continue to change as more information from the observation is passed through the Kalman recursions. As a result, the maximum likelihood estimate $\widehat{r}_{ML}$ will differ from the starting value r_0. The depicted cross correlations are presented in Figure 1.8 with the largest positive lagged cross correlations for Germany, France and Italy recorded at lags 1, 1 and 2 in respective order. Large values for cross correlations are seen at negative lags due to the nature of phase

[12] See http://stats.oecd.org/ or a main source of all macroeconomic data at country level.

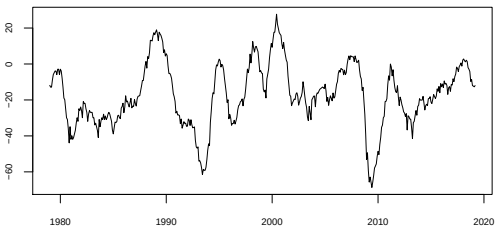
shifts of corresponding cycles however, we restrict ourselves to select the largest cross correlation at positive lags.

Figure 1.7: Seasonally Adjusted Economic Indicators for Germany, France, Italy

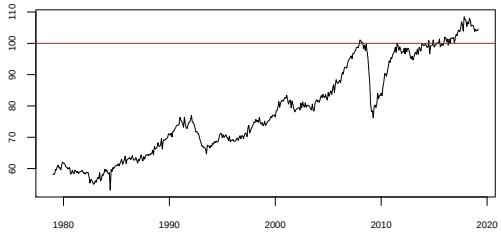
(a) IP: France



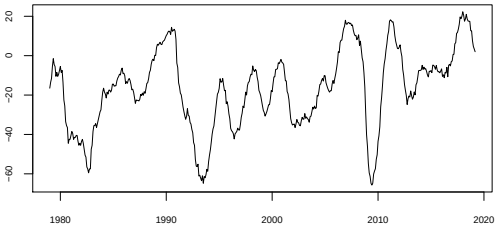
(b) OB: France



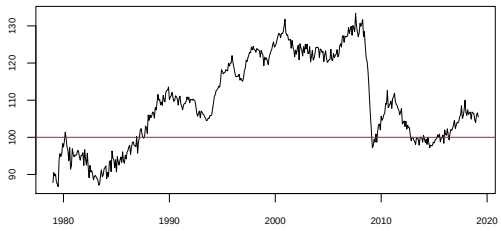
(c) IP: Germany



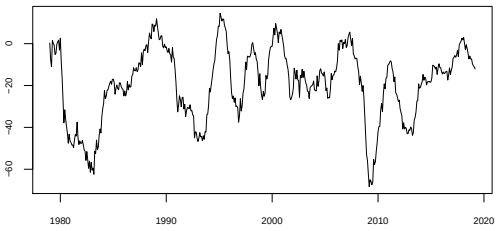
(d) OB: Germany



(e) IP: Italy



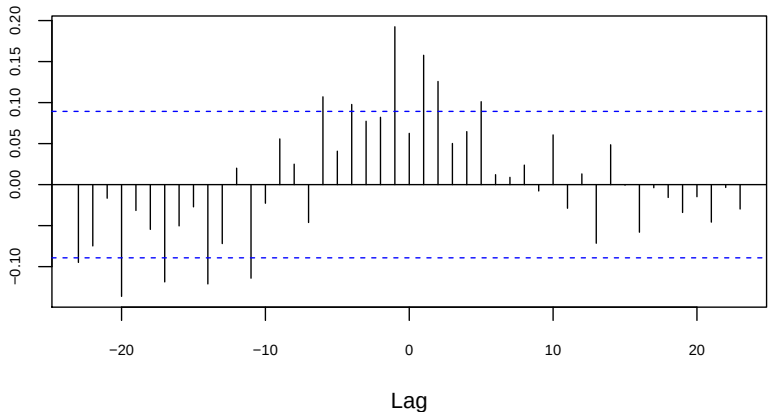
(f) OB: Italy



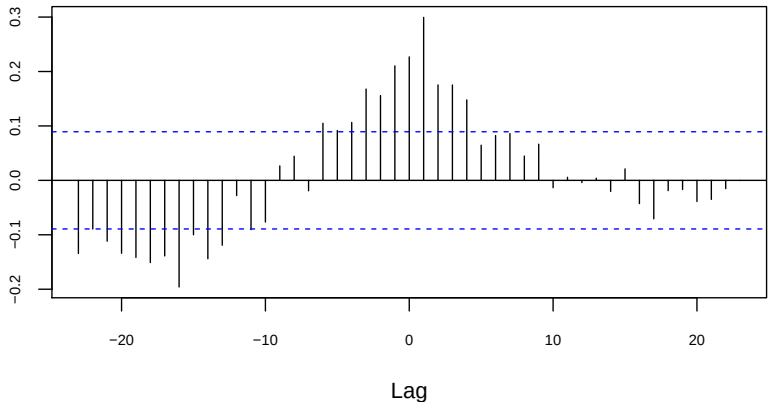
Note: The left side panel shows the IP series, the red line represents $\overline{IP}_{2015} = 100$. The right-side panels are for the order book balance series.

Figure 1.8: Cross Correlogram of IP and Order Book Series, First Differenced ($\widehat{\rho}(j)$)

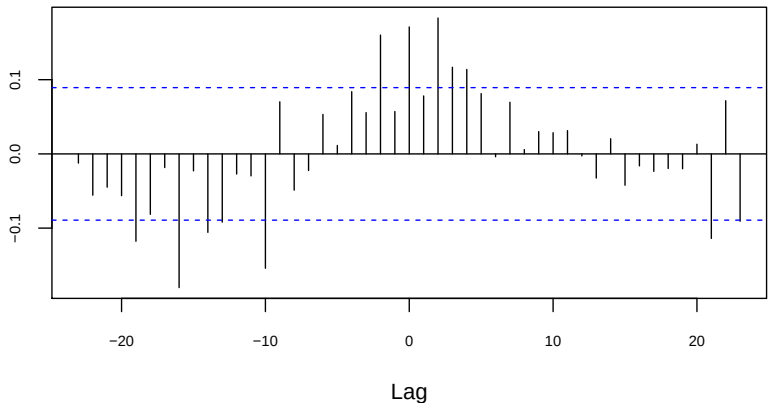
(a) France



(b) Germany



(c) Italy



Note: Blue line's indicate the 95% confidence intervals.

To specify the UCM for stock variables, the choice of components for the IP series and the Order book series for each of the three countries are modelled in the following way

$$\begin{aligned} IP_t &= \mu_{1,t} + \psi_t^{(h+1)} + \varepsilon_{1,t} \\ OB_t &= \mu_{2,t} + \gamma_t + \beta\varphi_t + \varepsilon_{2,t} \end{aligned} \tag{1.34}$$

where $\psi_t^{(h+1)}$ and φ_t are cycles defined in equation (1.25), $\mu_{1,t}$, $\mu_{2,t}$ are the integrated random walk in the form of equation (1.29), γ_t is an idiosyncratic stock cycle in the form of

$$\begin{bmatrix} \gamma_t \\ \gamma_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda & \sin\lambda \\ -\sin\lambda & \cos\lambda \end{bmatrix} \begin{bmatrix} \gamma_{t-1} \\ \gamma_{t-1}^* \end{bmatrix} \begin{bmatrix} \xi_t \\ \xi_t^* \end{bmatrix} \tag{1.35}$$

where ξ_t, ξ_t^*, $\varepsilon_{1,t}$, $\varepsilon_{2,t}$ are white noise stock variable disturbances which are orthogonal to each other and with variances σ_ξ^2, $\sigma_{\varepsilon_1}^2$, $\sigma_{\varepsilon_2}^2$.

Now that the observation equations has been defined, one can proceed in estimating the model parameters $\boldsymbol{\theta}_{ML}$. Under the framework of UCM, this will enable smoothing algorithms to extract the trend, cycles and other components of the stock variables. The Kalman smoothing algorithm, enables the estimation of the state vectors $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..., \boldsymbol{\alpha}_t$ conditional on all the data sample which is the information set $\mathbb{I}_s = (\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_s)$ where $s > t$. Assuming normality of the state vector, the conditional density given $\boldsymbol{\alpha}_t|\mathbb{I}_s \sim N(\boldsymbol{a}_{t|s}, \boldsymbol{P}_{t|s})$ follows $\boldsymbol{a}_{t|s} = \mathbb{E}[\boldsymbol{\alpha}_t|\mathbb{I}_s]$ where $\boldsymbol{P}_{t|s} = \mathbb{V}[\boldsymbol{\alpha}_t|\mathbb{I}_s]$.[13] The term $\boldsymbol{a}_{t|s}$ is the smoothed state at t, which can be calculated for $t = 1, 2, ..., n$ so that we obtain state smoothing vector series $\boldsymbol{a}_{1|s}, \boldsymbol{a}_{2|s}, ..., \boldsymbol{a}_{n|s}$.

The results for the estimated smoothed stock cycles of $\phi_t$and $\psi_t^{(h+1)}$ for each country; Germany, France and Italy are presented in Figure 1.9. It is very clear from three of the panels that the integer part h of the shift, is not a large number as the estimated cycles $\widetilde{\psi}_{t|s}^{(h+1)}$ do not have large phase differences with respect to and $\widetilde{\phi}_{t|s}$. This is an indication that the delay parameter r is not large in number. While the values found for the estimated parameters will be provided in later chapters, for the time being we only focus on a graphical examination. Under the framework of equation (1.34), the smooth estimate of the fractionally shifted cycle $\widetilde{\phi}_{t|s}$ associates with the Order Book series while the integer shifted version $\widetilde{\psi}_{t|s}^{(h+1)}$ relates to the IP series. The assertion that the Order Book series is a leading indicator of the BC can be confirmed especially from the panels for France and Italy. From these panels, it can be seen the smoothed cycle estimate $\widetilde{\psi}_{t|s}^{(h+1)}$ associated to IP has a larger shift to the right wrt $\widetilde{\phi}_{t|s}$ which is the smoothed cycle estimate associated with OB. Hence, we can conclude that, graphically examination implies that our model framework seems to have successfully generated shifted cycles of the exact baseline cycle. More importantly, it could be asserted that the model exploit the

[13]For a complete description of the algorithm see Harvey (1990), Durbin and Koopman (2012) and Pelagatti (2015).

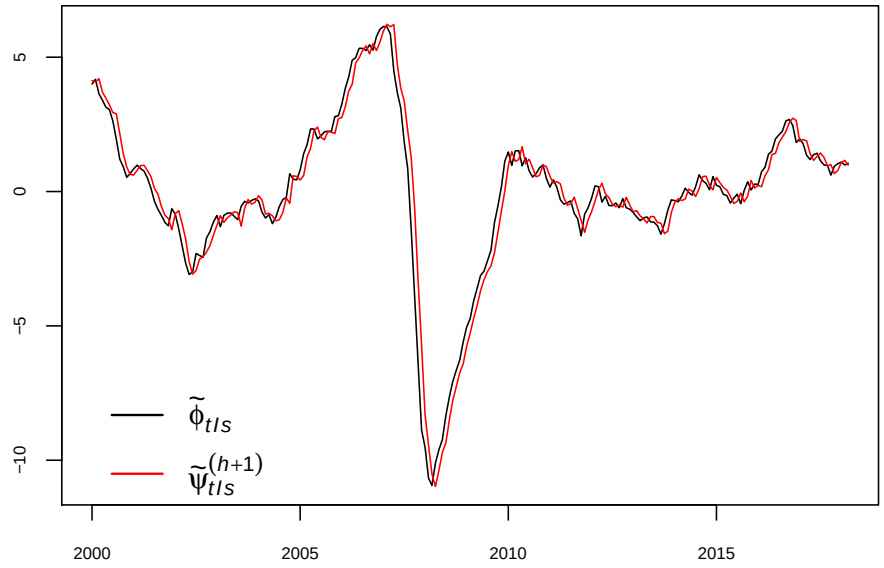
leading/coinciding cycle dynamics inherent in the data series efficiently. Finally, the panels portray BC cycle heterogeneity among 3 major EU economies in terms of peak/downturn sharpness as well as cycle variability.

To check how well the cycles perform in predicting recessions, the smooth estimates of the base cycle $\psi_{t|s}^{(0)}$ for the stock model together with the NBER recession shades are plotted. One can visually inspect the performance of the estimates in this fashion. Figure 1.10 depicts the results for all three countries; France, Germany, Italy. The cycle estimates are able to track every recession period by showing sharp declines as it enters the shaded areas. It can be seen that the turning point for the cycle only takes place towards the end of the recession periods. Hence, smoothed stock cycle estimates present a good tracking record of the recession timings as well as their duration.

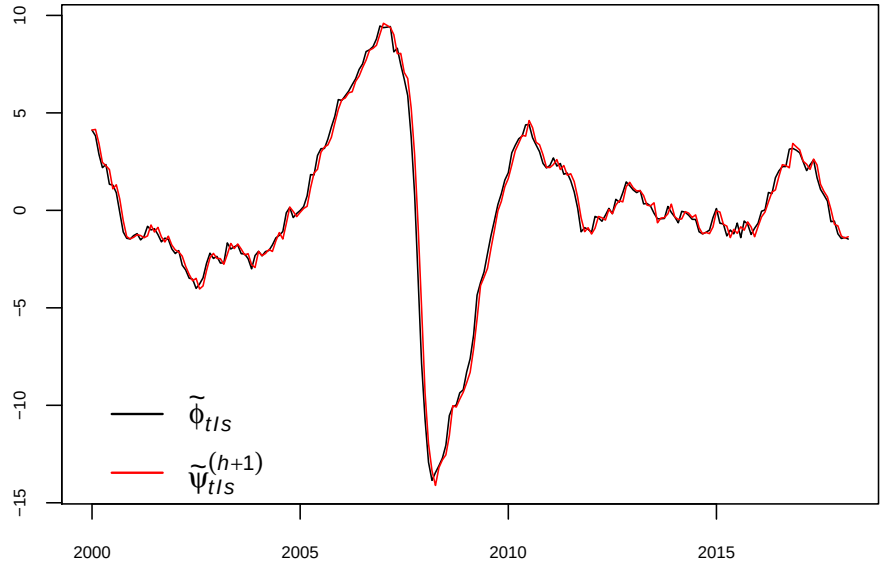
While our study aims to extract exact phase shifted estimates of the BC, as a by-product of the UCM we can extract smooth trend components using the Kalman recursions aswell. We denote the trend components for the IP series $y_{1,t}$ and OB series $y_{2,t}$ as $\mu_{1,t}$ and $\mu_{2,t}$. Figure dd presents the smoothed estimates of the trends. For the IP series shown on the left panels, shows that France, Germany, Italy shows significant changes in the slope, on average produce upward trending estimates of $\widehat{\mu}_{1,t|s}$. For Italy, long periods of a negative slope is calculated during the period of 2008-2015 which produce a hump shaped smooth trend estimate for its IP series. For the Order book series, the extracted smooth trend estimates $\widehat{\mu}_{2,t|s}$ are illustrated on the right hand side panels. The Order book series is cyclical but does not reflect a pronounced upward or downward movement as relevant as the IP series. However, for Germany and Italy, the smoothed component is able to generate trends with changing slope. For France on the other hand, the estimated trend component is fairly flat and does not produce such slope variability in its estimated trend.

Figure 1.9: Estimated Stock Cycle Components: $\widetilde{\phi}_{t|s}$, $\widetilde{\psi}^{(h+1)}_{t|s}$

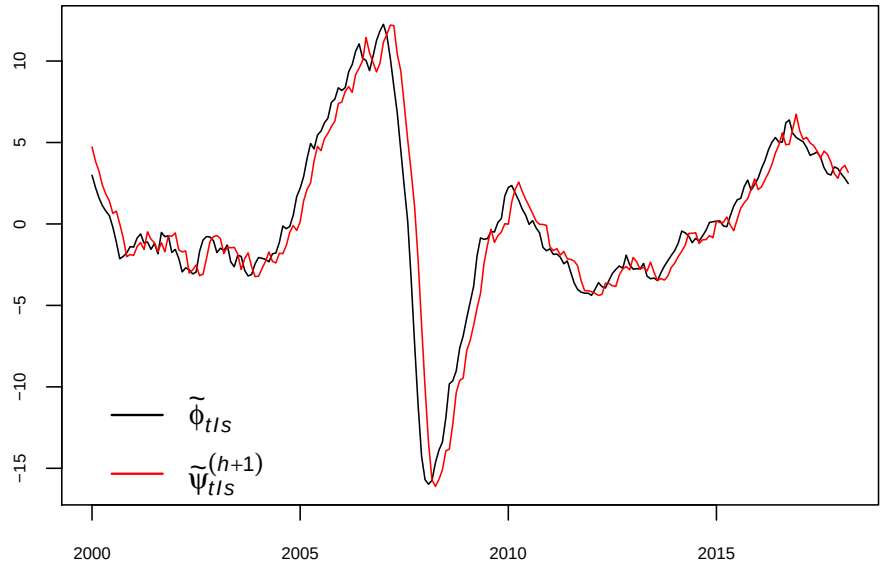
(a) France



(b) Germany



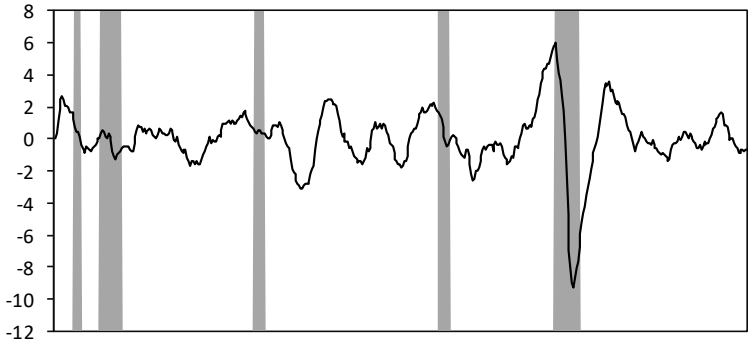
(c) Italy



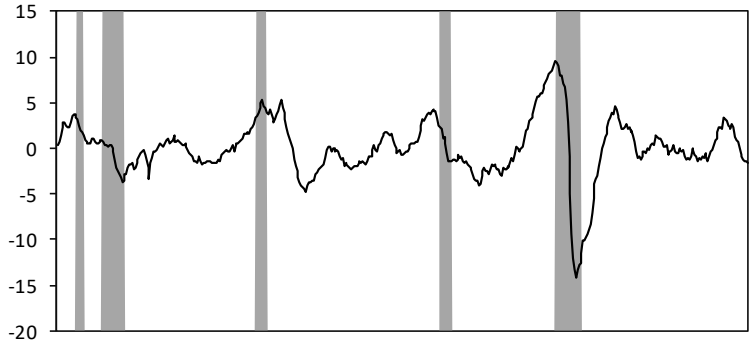
Note: Panels include sample period of 01/1999-03/2019 for clarity.

Figure 1.10: Smoothed Business Cycle Indicator with Recession Dating

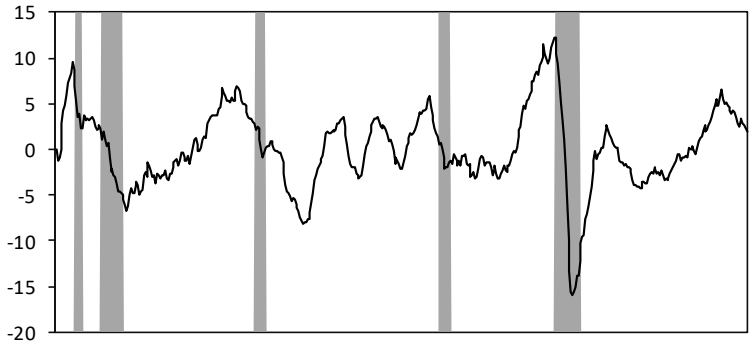
(a) France



(b) Germany



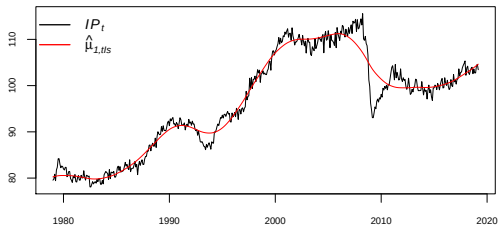
(c) Italy



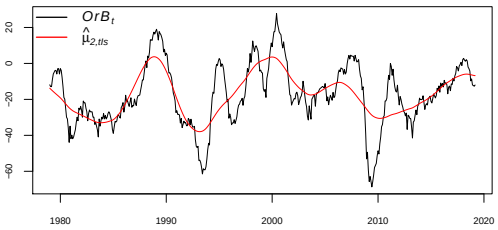
Note: NBER recession datings represented by grey shades were obtained from, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/. Panels (a), (b), (c) corresponds to smoothed common stock cycle estimates of France, Germany, Italy in respective order.

Figure 1.11: Smooth trend estimates of IP and OB: $\widehat{\mu}_{2,t|s}$, $\widehat{\mu}_{1,t|s}$

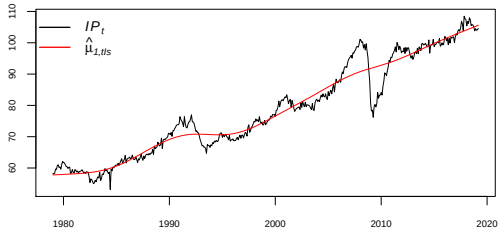
(a) IP: France



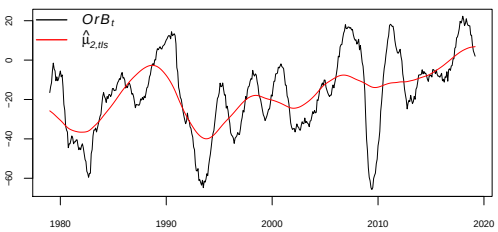
(b) OB: France



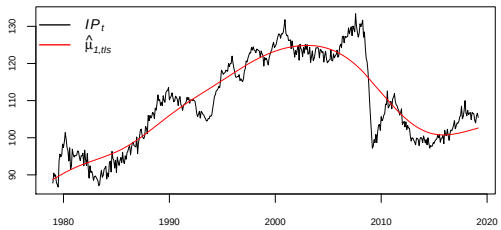
(c) IP: Germany



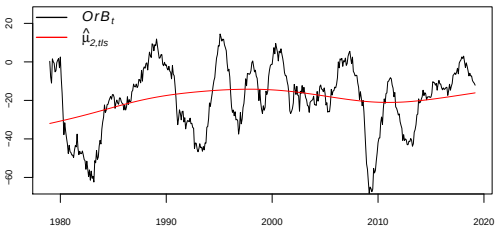
(d) OB: Germany



(e) IP: Italy



(f) OB: Italy



Note: IP_t is total industry production index and OrB_t is the Order Book balance series.

Before presenting the parameter estimates of the stock UCM, we briefly explain starting values chosen for the parameters. For, the parameters σ_{ε_1}, σ_c the choice for initial values are $s_{y_{1,t}}/100$ and $s_{y_{1,t}}/200$ which are the sample variance of $y_{1,t}$ divided by 100 and 200 respectively. For σ_{ε_2}, σ_ξ the initial values are $s_{y_{2,t}}/100$ and $s_{y_{2,t}}/200$ the sample variance of $y_{2,t}$ divided by 100 and 200 respectively. The starting value for r is set at $|\widehat{\rho}(j)| + \delta_0$ according to the positive lag of the highest sample cross correlation between the $\Delta y_{1,t}$ and $\Delta y_{2,t}$ plus $\delta_0 = 0.5$.[14] For the decay parameter $\rho \in (0, 1]$ which determines the persistence of the stochastic cycles ψ_t, γ_t and φ_t; the starting value is set to 0.98. The persistence parameter ρ has a restriction, therefore the necessary anti-transformation would be $\varpi_\rho = \ln[\rho(1-\rho)^{-1}]$ with $\varpi_\rho = 4$.[15] The frequency of the BC determined by λ is selected such that the cycle period roughly lies around 5 years ($\sim$ 60 years). Based on the logit function that determine the period parameter p with limits set between $24 - 96$, it follows that $p = 24 + (96 - 24)/(1 + e^{-\varpi_p})$. By selecting $\varpi_p = 0.5$ it is ensured that p is about $\sim$ 68 months. Using the expression of $\lambda = 2\pi/p$, and plugging in the preset values; the BC frequency can finally be calculated as $\lambda = 0.091$. Lastly, the scaling parameter β is chosen at 1.

Table 1.2 presents the maximum likelihood values of the stock model parameters. Among the three countries, Italy has the highest estimated $\sigma^2_{\varepsilon_1}$ at 0.945. This indicates the irregular component for Italy IP shows greater variation with respect to France and Germany. For the OP series, the highest variation for the irregular component is recorded for France since $\sigma^2_{\varepsilon_2} = 2.176$. Variation in slope components in the trend measured by $\sigma^2_{\eta_1}$ and $\sigma^2_{\eta_2}$ are low for all three countries as expected, but with Italy comparing relatively higher to Germany and France. The estimated value σ^2_c allows us to compute variation in the stochastic cycle ψ_t measured by σ^2_κ through equation (1.24a). Accordingly, σ^2_κ is calculated as 0.061, 0.646 and 0.712 for France, Germany and Italy respectively. This implies that Italy also exhibits the highest variability in its cyclical component. Results for phase shift parameter r shows that Italy leads the BC by 2.3 months followed by France 1.1 months and Germany by less than a month. It should also be noted that despite the shift parameter estimated for Germany being less than an integer value, the model still generates an integer shifted stochastic cycle where $h = 1$. For all countries, the stochastic cycles are very persistent since ρ is estimated to be between $0.971 - 0.982$. The estimated cycle lengths range between 46 to 73 months and therefore the initial value of 68 months seems to be a reasonable starting point. Finally, σ^2_ξ which measures the variance for the stochastic cycle errors of γ_t; are found significantly higher than common cycle variances of σ^2_κ for each country. This may not come as a surprise since the variation in the second stochastic cycle is not accounted by the phase shift mechanism.

[14] See Section 1.3.2 for the reasoning of this choice.

[15] See Appendix Section 1.8.3 for further details.

Table 1.2: ML Estimates of the UCM in State Space Form

	FRANCE		GERMANY		ITALY	
Parameter	Estimate	St.Error	Estimate	St.Error	Estimate	St.Error
$\sigma^2_{\varepsilon_1}$	0.677	0.068	0.518	0.060	0.945	0.120
$\sigma^2_{\varepsilon_2}$	2.176	0.64	0.011	0.047	1.534	0.547
$\sigma^2_{\eta_1}$	0.001	0.001	0.000	0.000	0.001	0.000
$\sigma^2_{\eta_2}$	0.039	0.035	0.020	0.015	0.000	0.000
σ^2_c	0.227	0.060	0.461	0.067	0.976	0.1525
r	1.117	0.251	0.615	0.112	2.319	0.147
ρ	0.971	0.007	0.982	0.005	0.977	0.006
$p = 2\pi/\lambda$	45.536	52.965	53.827	149.208	73.406	85.752
β	2.744	0.578	1.973	0.267	1.578	0.264
σ^2_ξ	4.065	1.235	3.328	0.390	7.010	1.025

Note: St.Error indicate the standard error, computed using the Delta method discussed in Appendix D.

1.6 Extension to higher dimensions

In section 1.2.5, we introduced a single cycle with integer shift h which decided the dimension of the full system. In this section we show how to include many cycles with varying velocities that are each represented by a different corresponding h shift integer. This construct will allow to extend the model defined my the building blocks of equations 1.25, 1.26 and 1.27 in order to incorporate and extra phase shifted cycle. Once such a setting is designed, it will be possible to account for any number of shifted cycles in the system. To simplify matters, consider the case where an additional cycle with a different speed is introduced to the original system.[16] This will require an additional layer of integer shift h_2. In turn the system will posses two common cycles with different delays; h_1 (now defined as the original cycle integer shift) and h_2 (the slower moving cycle integer shift) with $h_2 > h_1$. In the 2-cycle case, the integer shift was presented by its fractional shift counterpart ϕ_t. The fractional shift parameter δ_t over a common cycle gave us the continuous time delay parameter r. To keep the system intact, a second fractional shift is required which will correspond to a second delay parameter r_2. Hence, the original model can be further expanded to accommodate slower cycles by adding new layers of integer shifts; $\{h_2,\ h_3, ...\}$. Each of the new delayed cycles $\{\psi_t^{(h_2)},\ \psi_t^{(h_3)}, ...\}$, will in turn be accompanied by a fractional shifted cycle; $\phi_2,\ \phi_3, ...$ to complete the system. These cycles are fractionally delayed by $\delta_2,\ \delta_3, ...$ over the integer delayed cycles of $\psi_t^{h_2+1},\ \psi_t^{h_3+1}, ...$ and gives the continuous delays $\{r_2,\ r_3, ...\}$ over the baseline ψ_t where $r_i = h_i + \delta_i$.

To understand such a general structure, consider a case where we observe 3 time series $\{y_{it}\}$ with $i = 0, 1, 2$: series y_{1t} lags y_{0t} by $r_1 = h_1 + \delta_1$ and series y_{2i} lags y_{0t} by $r_2 = h_2 + \delta_2$ with h_1 and h_2 positive integers, δ_1 and δ_2 real numbers in $[0, 1)$, and $0 \leq r_1 \leq r_2$:

$$\begin{aligned} y_{0t} &= \mu_{0,t} + \psi_t + \epsilon_{0,t} \\ y_{1t} &= \mu_{1,t} + \gamma_1 \phi_{1,t} + \epsilon_{1,t} \\ y_{2t} &= \mu_{2,t} + \gamma_2 \phi_{2,t} + \epsilon_{2,t} \end{aligned} \tag{1.36}$$

with $\phi_{1,t} = \psi_{t-r_1}$ and $\phi_{2,t} = \psi_{t-r_2}$, from which we have $\phi_{1,t+\delta_1} = \psi_{t-h_1}$ and $\phi_{1,t+\delta_2} = \psi_{t-h_2}$ and $\mu_{i,t}$ is an integrated random walk shown in equation 1.29 for $i = 0, 1, 3$. Equation 1.36 defines the measurement equations to be used in the Kalman recursions.

Assuming ψ_t to be a (continuous time) stochastic cycle, if we want to state the model in state space form we have to write all the relations in a order-1 Markovian form[17]:

[16] The term "*speed*" is used in the context that a cycle leads the other in the event that it has already passed the trajectory of the common cycle. Therefore, speed and delay can be used interchangeably.

[17] The Markovian property refers to the definition whereby if $\boldsymbol{X}_t$ is a stochastic process evolving between finite number of states $s_1, s_2, ..., s_n \in \boldsymbol{S}$ will follow the dependency according to $\boldsymbol{p}(\boldsymbol{X}_{t+1} = s_k | \boldsymbol{X}_t = s_j)$ at each point in time.

$$
\begin{aligned}
\kappa_{t+1} &= WN(0, \sigma_\kappa^2) \\
\kappa_{t+1}^* &= WN(0, \sigma_\kappa^2) \\
\psi_{t+1} &= \rho\cos(\lambda)\psi_t + \rho\sin(\lambda)\psi_t^* + \kappa_t \\
\psi_{t+1}^* &= -\rho\sin(\lambda)\psi_t + \rho\cos(\lambda)\psi_t^* + \kappa_t^* \\
\kappa_{t+1}^{(1)} &= \kappa_t \\
\kappa_{t+1}^{*(1)} &= \kappa_t^* \\
&\vdots \\
\kappa_{t+1}^{(h_2+1)} &= \kappa_t^{(h_2)} \\
\kappa_{t+1}^{*(h_2+1)} &= \kappa_t^{*(h_2)} \\
\psi_{t+1}^{(1)} &= \psi_t \\
\psi_{t+1}^{*(1)} &= \psi_t^* \\
&\vdots \\
\psi_{t+1}^{(h_2+1)} &= \psi_t^{(h_2)} \\
\psi_{t+1}^{*(h_2+1)} &= \psi_t^{*(h_2)} \\
\phi_{1,t+1} &= \rho^{1-\delta_1}\cos(\lambda(1-\delta_1))\psi_t^{(h_1+1)} + \rho^{1-\delta_1}\sin(\lambda(1-\delta_1))\psi_t^{*(h_1+1)} + \nu_{1t} \\
\phi_{2,t+1} &= \rho^{1-\delta_2}\cos(\lambda(1-\delta_2))\psi_t^{(h_2+1)} + \rho^{1-\delta_2}\sin(\lambda(1-\delta_2))\psi_t^{*(h_2+1)} + \nu_{2t}
\end{aligned}
\tag{1.37}
$$

with

$$Var(\nu_{1t}) = Cov(\nu_{1t}, \kappa_t^{(h_1+1)}) = \frac{\sigma_c^2(1-\rho^{2(1-\delta_1)})}{(-2\log(\rho))}$$

and

$$Var(\nu_{2t}) = Cov(\nu_{2t}, \kappa_t^{(h_2+1)}) = \frac{\sigma_c^2(1-\rho^{2(1-\delta_2)})}{(-2\log(\rho))}.$$

Let us see what happens when $h_1 = 0$ and $h_2 = 1$.

$$
\begin{aligned}
\kappa_{t+1} &= WN(0, \sigma_\kappa^2) \\
\kappa^*_{t+1} &= WN(0, \sigma_\kappa^2) \\
\psi_{t+1} &= \rho\cos(\lambda)\psi_t + \rho\sin(\lambda)\psi^*_t + \kappa_t \\
\psi^*_{t+1} &= -\rho\sin(\lambda)\psi_t + \rho\cos(\lambda)\psi^*_t + \kappa^*_t \\
\kappa^{(1)}_{t+1} &= \kappa_t \\
\kappa^{*(1)}_{t+1} &= \kappa^*_t \\
\kappa^{(2)}_{t+1} &= \kappa^{(1)}_t \\
\kappa^{*(2)}_{t+1} &= \kappa^{*(1)}_t \\
\psi^{(1)}_{t+1} &= \psi_t \\
\psi^{*(1)}_{t+1} &= \psi^*_t \\
\psi^{(2)}_{t+1} &= \psi^{(1)}_t \\
\psi^{*(2)}_{t+1} &= \psi^{*(1)}_t \\
\phi_{1,t+1} &= \rho^{1-\delta_1}\cos(\lambda(1-\delta_1))\psi^{(1)}_t + \rho^{1-\delta_1}\sin(\lambda(1-\delta_1))\psi^{*(1)}_t + \nu_{1t} \\
&= \rho^{1-\delta_1}\cos(\lambda(1-\delta_1))\psi_{t-1} + \rho^{1-\delta_1}\sin(\lambda(1-\delta_1))\psi_{t-1} + \nu_{1t} \\
\phi_{2,t+1} &= \rho^{1-\delta_2}\cos(\lambda(1-\delta_2))\psi^{(2)}_t + \rho^{1-\delta_2}\sin(\lambda(1-\delta_2))\psi^{*(2)}_t + \nu_{2t} \\
&= \rho^{1-\delta_2}\cos(\lambda(1-\delta_2))\psi_{t-2} + \rho^{1-\delta_2}\sin(\lambda(1-\delta_2))\psi_{t-2} + \nu_{2t}
\end{aligned}
$$

with

$$
Var(\nu_{1t}) = Cov(\nu_{1t}, \kappa^{(1)}_t) = \frac{\sigma_c^2(1-\rho^{2(1-\delta_1)})}{(-2\log(\rho))}
$$

and

$$
Var(\nu_{2t}) = Cov(\nu_{2t}, \kappa^{(2)}_t) = \frac{\sigma_c^2(1-\rho^{2(1-\delta_2)})}{(-2\log(\rho))}.
$$

Since $\kappa^{(h+1)}_t$ are not contemporaneous with ν_t, we have to rewrite the system to reach a state-space form representation. First of all notice that, if we have know $\kappa^{(h_1+1)}_t$, the conditional distribution of ν_{1t} is Gaussian with first two moments given by

$$
\begin{aligned}
\mathbb{E}\Big(\nu_{1t}|\kappa^{(h_1+1)}_t\Big) &= \frac{Cov\Big(\nu_{1t}, \kappa^{(h_1+1)}_t\Big)}{Var(\kappa^{(h_1+1)}_t)}\kappa^{(h_1+1)}_t = \frac{1-\rho^{2(1-\delta_1)}}{1-\rho^2}\kappa^{(h_1+1)}_t \\
Var\Big(\nu_{1t}|\kappa^{(h_1+1)}_t\Big) &= Var\Big(\nu_{1t}\Big) - \frac{Cov\Big(\nu_{1t}, \kappa^{(h_1+1)}_t\Big)^2}{Var(\kappa^{(h_1+1)}_t)} = \frac{(1-\rho^{2(1-\delta_1)})(\rho^{2(1-\delta_1)}-\rho^2)}{-2\log(\rho)(1-\rho^2)}\sigma_c^2.
\end{aligned}
\tag{1.38}
$$

where equation 1.38 makes use of the first two regression lemmas and with the same results also holding for the pair $\nu_{2t}, \kappa^{(h_2+1)}_t$. Thus, we can rewrite the above system as

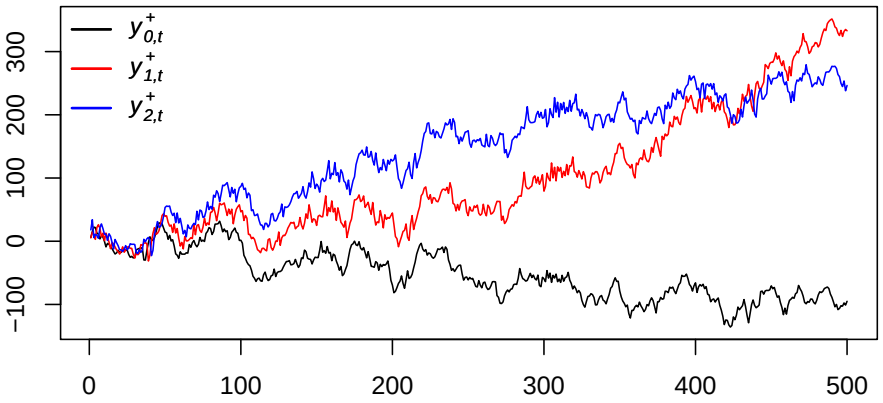
$$
\begin{aligned}
\kappa_{t+1} &= WN(0, \sigma_\kappa^2) \\
\kappa^*_{t+1} &= WN(0, \sigma_\kappa^2) \\
\psi_{t+1} &= \rho\cos(\lambda)\psi_t + \rho\sin(\lambda)\psi^*_t + \kappa_t \\
\psi^*_{t+1} &= -\rho\sin(\lambda)\psi_t + \rho\cos(\lambda)\psi^*_t + \kappa^*_t \\
\kappa^{(1)}_{t+1} &= \kappa_t \\
\kappa^{*(1)}_{t+1} &= \kappa^*_t \\
&\vdots \\
\kappa^{(h_2+1)}_{t+1} &= \kappa^{(h_2)}_t \\
\kappa^{*(h_2+1)}_{t+1} &= \kappa^{*(h_2)}_t \\
\psi^{(1)}_{t+1} &= \psi_t \\
\psi^{*(1)}_{t+1} &= \psi^*_t \\
&\vdots \\
\psi^{(h_2+1)}_{t+1} &= \psi^{(h_2)}_t \\
\psi^{*(h_2+1)}_{t+1} &= \psi^{*(h_2)}_t \\
\phi_{1,t+1} &= \rho^{1-\delta_1}\cos(\lambda(1-\delta_1))\psi^{(h_1+1)}_t + \rho^{1-\delta_1}\sin(\lambda(1-\delta_1))\psi^{*(h_1+1)}_t + \frac{1-\rho^{2(1-\delta_1)}}{1-\rho^2}\kappa^{(h_1+1)}_t \\
&\quad + WN\left(0, \frac{(1-\rho^{2(1-\delta_1)})(\rho^{2(1-\delta_1)}-\rho^2)}{-2\log(\rho)(1-\rho^2)}\sigma_c^2\right) \\
\phi_{2,t+1} &= \rho^{1-\delta_2}\cos(\lambda(1-\delta_2))\psi^{(h_2+1)}_t + \rho^{1-\delta_2}\sin(\lambda(1-\delta_2))\psi^{*(h_2+1)}_t + \frac{1-\rho^{2(1-\delta_2)}}{1-\rho^2}\kappa^{(h_2+1)}_t \\
&\quad + WN\left(0, \frac{(1-\rho^{2(1-\delta_2)})(\rho^{2(1-\delta_2)}-\rho^2)}{-2\log(\rho)(1-\rho^2)}\sigma_c^2\right)
\end{aligned}
\tag{1.39}
$$

Equation 1.39 provides the transition equations which allow the system be written in discrete time form and compute both the Kalman filter and smoothed estimates of the model states. We can now apply the above construct to simulate the data given the above model specification. To asses the model, initially we use simulated data and evaluate whether an extra delayed cycle can be generated. We use the simulation steps outlined in Durbin and Koopman (2002). Accordingly, a sequence of the state vectors $\boldsymbol{\alpha}_t$ and the observations $\boldsymbol{y}_{it}$ can be simulated from the joint density $\boldsymbol{\alpha}^+, \boldsymbol{y}^+ \sim f(\boldsymbol{\alpha}, \boldsymbol{Y}) = f(\boldsymbol{\alpha})f(\boldsymbol{Y}|\boldsymbol{\alpha})$. The first part of the product is the model density which follows from the transition equation 1.42, with simulated values plugged in recursively such that $\boldsymbol{\alpha}$. This allows us to find the simulated state vector $\boldsymbol{\alpha_t^+}$ using the state errors generated from a Gaussian distribution; $\boldsymbol{\eta_t^+} \sim N(\boldsymbol{0_{[3]}}, \boldsymbol{Q_t})$. In a similar fashion, the measurement error is generated according to the distribution $\boldsymbol{\epsilon_t^+} \sim N(\boldsymbol{0_{[3]}}, \boldsymbol{H_t})$ with $\boldsymbol{H_t} = \boldsymbol{I_{[3]}}$. This allows us to retrieve the simulated observations from the simulated signal and noise: $\boldsymbol{y_t^+} = \boldsymbol{\Phi_t^+} + \boldsymbol{\epsilon_t^+}$, with signal is $\boldsymbol{\Phi_t^+} = \boldsymbol{Z_t}\boldsymbol{\alpha_t^+}$.[18] Figure 1.12 presents these steps which produces the simulated observations. According to equation 1.36, we have introduced an integrated random walk for each series in addition to common cycle with phase variants and error terms

[18] See Appendix 1.8.4 for the specifications of the system matrices.

all of which is evident from the evolution paths of $\boldsymbol{y_{it}^{+}}$ series.

Figure 1.12: Simulated unconditional observations for 3 cycle system; $\boldsymbol{y_t^+ = \theta_t^+ + \epsilon_t^+}$ with $\boldsymbol{\theta_t^+ = Z_t \alpha_t^+}$ and $\boldsymbol{\epsilon_t^+} \sim N(\mathbf{0_{[3]}}, \boldsymbol{I_{[3]}})$



Note: For a detailed explanation of the unconditional observation and state simulation algorithm see Durbin and Koopman (2002).

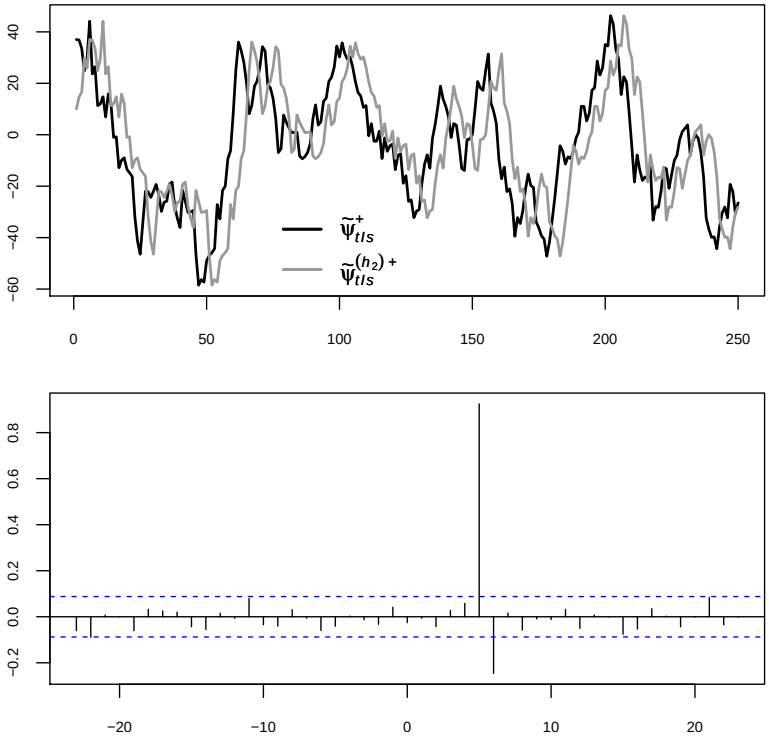
From the simulated observations, we compute the conditional states such that $\tilde{\boldsymbol{\alpha}}^{+}_{\boldsymbol{t|s}} = \mathbb{E}(\alpha_t = \alpha_t^+ | y_t = y_t^+)$ and the cross correlograms between Δy_{it}^+ series. In panel (a) of figure 1.13, total phase difference between the slowest cycle $\boldsymbol{\psi_t^{(h_2+1)+}}$ and the coinciding cycle $\boldsymbol{\psi_t^+}$ is highlighted. The delay between the two cycles can be measured using the cross correlogram between lags of the leading indicator Δy_{0t}^+ and current value of the lagging indicator Δy_{2t}^+. Hence we measure the correlation $\hat{\rho}\left(\Delta y_{1,t}^+, \Delta y_{0,t-j}^+\right)$ against each j^{th} lag of $\Delta y_{0,t-j}^+$. The maximum correlation is reached at lag 4 such that $j_{max}(\widehat{\rho}) = 4$ meaning that the total integer delay can be set at $h_2 = 4$. Since r_1 and r_2 are delays in continuous time, the fractional portion of phase shifts can be further analysed. This is possible as we set δ_1 and δ_2 at arbitrary fraction levels of 0.5 each. In panel (b) and (c), the fractional shifted cycles ϕ_{1t}^+ and ϕ_{2t}^+ are each delayed over $\boldsymbol{\psi_t^{(h_1+1)+}}$ and $\boldsymbol{\psi_t^{(h_2+1)+}}$. As is expected, there is a strong alignment with the path of the integer shifts albeit in the presence of discrepancies. As was postulated in section 1.3.2, these patterns arise from the influence of fractional delays (δ) which generate unique cycles as they move away from their interval limits of 0 and 1.[19] This confirms the assertion that fractionally shifted cycles ϕ_{1t} ϕ_{2t} are similar cycles to integer shifted cycle $\boldsymbol{\psi}_t$ up to an extent, such that they are similar in frequency and amplitude but differ in terms of their stochastic features that arise from the error components presented in equation 1.39.

In section 1.3.2 we looked in to how model flexibility and degrees of freedom were related. More specifically, it was asserted that the fractional part of the delay parameter δ, had an impact on the model flexibility. This impact was reflected by the behaviour of the log likelihood function $lnL(\Theta; \boldsymbol{y}|\mathbb{I})$ computed from equation 1.32. When we consider a second fractional delay, the grid search of the likelihood will reveal a findings that further strengthens our first conclusion in section 1.3.2 where we stated that degrees of freedom was not just linked with the number of free observations. This result has relevance to the assertions in Janson et al. (2012) which states that degrees of freedom is not a static concept within a single model

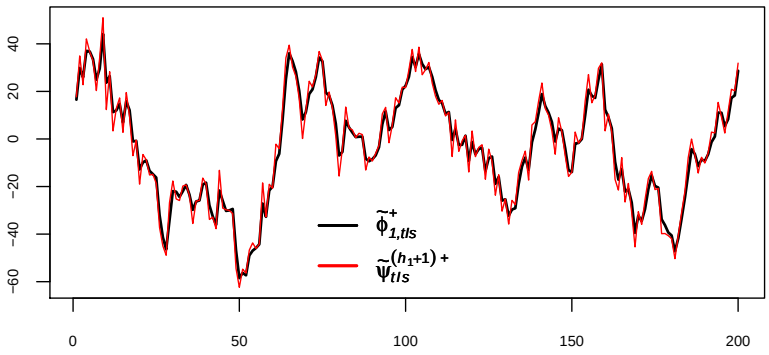
[19]Parameter values that form the parameter vector $\boldsymbol{\theta}$ were set either intuitively or from initial information obtained from the data. They do not reflect MLE estimates at this stage of the study.

Figure 1.13: Similar cycles based on; $\tilde{\boldsymbol{\alpha}}^{+}_{\boldsymbol{t|s}}$, $\boldsymbol{\eta}^{+}_{\boldsymbol{t}} \sim N(\mathbf{0}_{[3]}, \boldsymbol{Q}_{\boldsymbol{t}})$, $f(\boldsymbol{\alpha}) : \boldsymbol{T}_{\boldsymbol{t}}\boldsymbol{\alpha}^{+}_{\boldsymbol{t}} + \boldsymbol{R}_{\boldsymbol{t}}\boldsymbol{\eta}^{+}_{\boldsymbol{t}}$

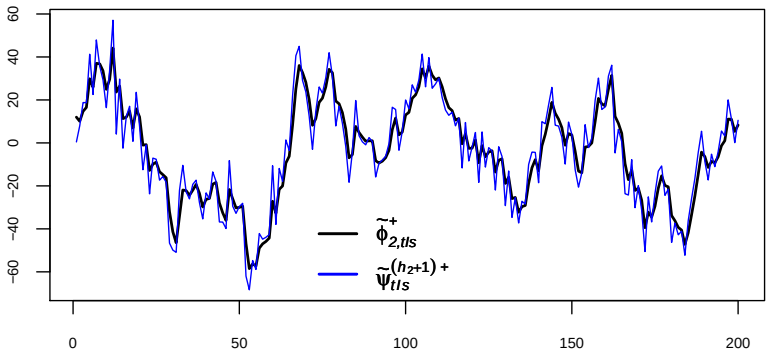
(a) Baseline and $h_2 = j_{max}\{\hat{\rho}(\Delta y^{+}_{0,t}, \Delta y^{+}_{2,t-j})\}$ shifts



(b) $h_1 = j_{max}\{\hat{\rho}(\Delta y^{+}_{1,t}, \Delta y^{+}_{0,t-j})\}$, $\delta_1 = 0.5$)



(c) $h_2 = h_1 + j_{max}\{\hat{\rho}(\Delta y^{+}_{2,t}, \Delta y^{+}_{1,t-j})\}$, $\delta_2 = 0.5$)



Note: Smoothed estimates of simulated states, requires the simulated unconditional observations in figure 1.12.

but rather a concept which has changing features with respect to the complexity of the model structure. In figure 1.14, we see the shape of the log likelihood when we allow one of the delay parameters to move freely and setting the other delay parameter together with all the other model parameters which form the parameter vector $\boldsymbol{\theta_0^-}$ at a constant value. The only difference of the parameter space θ of section 1.32, is that in the 3 cycle model we include parameters $\sigma^2_{\epsilon_2}$. $\sigma^2_{\eta_2}$, $\sigma^2_{\nu_1}$, $\sigma^2_{\nu_2}$, γ_1, γ_2 and the delay parameter r_δ where $r_2 = r_1 + r_\delta$. In panel (a) and (b) we let one of the delay parameters to move freely while the other is set at a fractional value 2.5. In panel (c) on the other hand we let r_1 move freely while setting r_δ at an integer of 2. In the first two panels, the negative log likelihood ($-lnL(\Theta; \boldsymbol{y}|\mathbb{I})$) is indeed minimized at the theoretical value of 2.5 where the integer value 2 is what the correlogram of our data had indicated and while $\delta = 0.5$ is where ϕ_{1t} has the lowest correlation with the exact integer shifted cycle of ψ_{h_1+1}. This low correlation causes the rigidness of the model to decline while increasing model flexibility. In turn $-lnL(\Theta; \boldsymbol{y}|\mathbb{I})$ jumps when r_1 hits 2.5. However, unlike in the 2 cycle case, the likelihood not only drops at the end and start of unit intervals $(0, 1)$ but also at each mid interval point of 0.5 through out the grid space $\{0, ..., 8\}$, excluding point 2.5. The reason for this is that, the second delay parameter which equals $r_2 = r_1 + r_\delta$ hits integral values every time r_1 has a fractional part of 0.5 since we initially set $r_2 = r_1 + 2.5$. Similarly in panel (b), when r_1 is set at 2.5 and r_δ is allowed to moves freely along the grid line, the minimization of the likelihood is achieved at point 2.5. Therefore, one of the fractional shifted cycles; ϕ_{2t} disappear and align with $\psi_t^{(5)}$ with a correlation of $\widehat{\boldsymbol{\rho}}\Big(\phi_{2t}, \psi_t^{(h1+(h_2-h_1)+1)}\Big) = 1$, while $\widehat{\boldsymbol{\rho}}\Big(\phi_{1t}, \psi_t^{(s)}\Big) < 1$ for $\forall s \in [0, 1, ..(h_2 + 1)]$. The perfect correlation is achieved at every mid interval point along the grid due to the exact reasons in panel (a). In panel (c) on the other hand, setting $r_\delta = 2$ results in achieving the highest model flexibility when $r_1 = 2.5$ with no jumps at mid intervals. This result comes from the fact that the model is able to generate two unique stochastic cycles in addition to the integer shifted cycles psis since $\widehat{\boldsymbol{\rho}}\Big(\phi_{1t}, \psi_t^{(h_1+1)}\Big) < 1$ and $\widehat{\boldsymbol{\rho}}\Big(\phi_{1t}, \psi_t^{(h_2+1)}\Big) < 1$ with correlations being at the lowest possible levels at mid interval levels. Setting $r_\delta = 2$ (an integer value) gives the model the highest flexibility as both r_2 and r_1 become a fraction. By doing so, the abrupt movements seen at mid unit intervals are eliminated and the likelihood behaviour converges to the two cycle model represented in figure 1.5. [20]

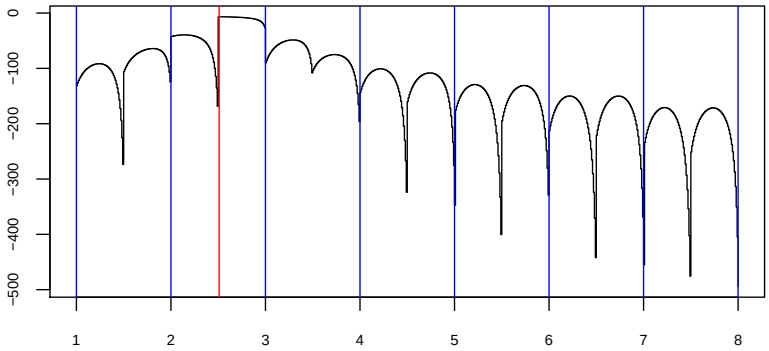
Since we have proved that our model successfully generates fractionally and integer shifted cycles, it is possible to move to an economic application. We consider the extraction of the common economic cycles for Italy by extending the bi-variate monthly data of Industrial Production (IP_t) and Order Book (OB_t) and add a third time series of the quarterly GDP growth rate $\Delta_{-4}gdp_t$. The series measures the percentage change of the quarterly GDP level (value added in Euro terms) series from the same quarter in the previous year.[21] This series will lag both IP_t and OB_t cycle components and hence allow the system to determine the parameters related to the cycles $\psi_t^{(h2)}$ and ϕ_{2t} which are h_2, r_δ and therefore r_2. As a quarterly series, we consider that the Δ_{-4} is a specific case of the seasonal MA(1) polynomial $b(L) = 1-$

[20] For a further analysis on the model, see Appendix 1.8.5 MCMC estimate distribution of the the delay parameter.

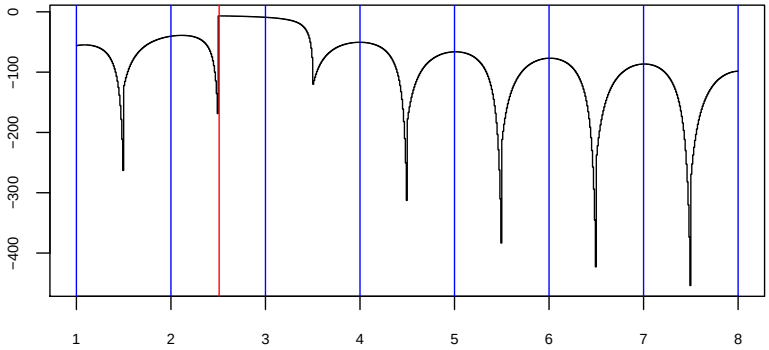
[21] Data source of Italy GDP QoQ is https://data.oecd.org/.

Figure 1.14: Grid Search of shift parameters, r_1, r_δ

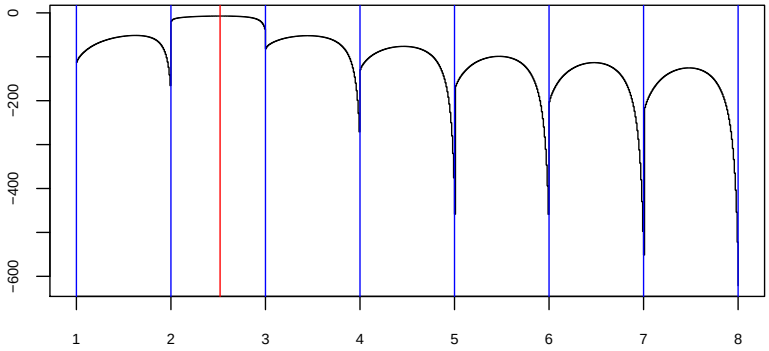
(a) $-lnL(\Theta; \boldsymbol{y}|\mathbb{I})$ and $\Theta = \{r_1, \boldsymbol{\theta}_0^-\}$ when $\delta_2^0 = 0.5$



(b) $-lnL(\Theta; \boldsymbol{y}|\mathbb{I})$ and $\Theta = \{r_\delta, \boldsymbol{\theta}_0^-\}$ when $\delta_1^0 = 0.5$



(c) $-lnL(\Theta; \boldsymbol{y}|\mathbb{I})$ and $\Theta = \{r_1, \boldsymbol{\theta}_0^-\}$ when $\delta_2^0 = 0$



Note: $y - axis$ represents negative log-likelihood values, $x - axis$ is the grid for r_1 in panel (a) and (c), while r_δ in (b). The red line indicates the maximum value of $-lnL(\Theta; \boldsymbol{y}|\mathbb{I})$ recorded at 2.5 for r_1 according to panels (a), (c) or for or r_δ according to panel (b).

$b_1 L^4$ where $b_1 = 1$. Furthermore we make use of the Kalman smoother as a gap filling technique by estimating $a = \mathbb{E}(\tilde{\alpha}_{t|n}|y = \Delta_{-4} gdp_n)$ as a univariate case. Panel (a) in figure 1.15 indicates that the interpolation method of the Kalman smoother results in the transformation of the quarterly data to a monthly frequency.[22] The data

[22]Interpolation indicates the smoothing of the low frequency data at the missing points to their

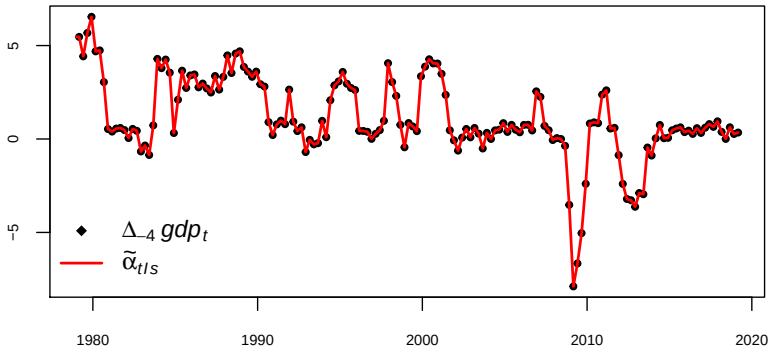
points are linked without the exclusion of a single observation and therefore evading the possibility of any information loss during the process. In panel (b) and (c) we examine the cross correlations between the first difference of the series to select the best arrangement for the ordering of lead/lag cycle dependency. In panel (c) of figure 1.8, it was already determined that ΔIP_t series lags ΔOB_t which we reinstate in panel (b). The $\Delta\Delta_{-4}gdp_t$ also lags the ΔOB_t series by a similar margin at lag $j = 2$. This indicates that there is a coinciding relationship between cyclical oscillations of IP_t and $\Delta_{-4}gdp_t$. The cross correlation findings which is not presented here comply with the notion of coinciding cycles between the two series since the maximum correlation is identified at lag zero: $j_{max}\{\hat{\rho}(\Delta\Delta_{-4}gdp_t, \Delta IP_{t-j})\} = 0$. Nevertheless, our model structure requires that we declare an ordering classification between the multivariate series. Therefore, while acknowledging that in some empirical work IP and GDP cycle may infact coincide, we apply our economic intuition, where we assert that cycles in the industrial sector leads the growth cycle of GDP. From this initial setting, we justify our ordering classification to be in the form of: $OB_t \xrightarrow{lead} IP_t \xrightarrow{lead} \Delta_{-4}gdp_t$. From equation 1.36, we can then insert the series such that $\{y_{0t} = OB_t,\ y_{1t} = IP_t,\ y_{2t} = \Delta_{-4}gdp_t\}$. Table 1.3 presents the details of the ML estimates of the 3-cycle UCM model which allows us to uncover the lead/lag dynamics of the 3-cycle system. We find the first lead of the common cycle in OB_t (ψ_t) over the common cycle in IP_t (ϕ_{1t}) that is measured by $\hat{r}_1$ to be roughly 1.3 months while the lead over $\Delta_{-4}gdp_t$ growth cycle stands at 1 quarterly period with $\hat{r}_2 = 3$. The lead of oscillations in IP_t over $\Delta_4 gdp_t$ cycles stands at 1.6 observation periods which implies that our findings support the initial assumption that indicators measuring economic activity in the industrial sector may be a relevant candidate for predicting economic cycles for the rest of the economy. Furthermore since $\hat{delta}_1 = 0.3435$ and $\hat{delta}_2 = 0.6565$, the model generates the two unique stochastic cycles associated with the fractional shifts $\tilde{\phi}_{1t}$ and $\tilde{\phi}_{2t}$. Table 1.15 also gives details on the statistical features of oscillations embedded in the system. Since we have only estimated cycles with equal frequency $(\hat{\lambda})$, all cycles regardless of whether they are fractional or integer shifts are classified as similar cycles in the sense that they have equal persistency levels in terms of their stochastic features and have also equal wave lengths. The estimated persistency measure with $\hat{\rho} = 0.9672$ suggests strong persistency that translates into smoothness of the identified cycles. The estimated wave periodicity at 30.3 suggests that our cycles are within the business cycle range of 2 to 8 years. Furthermore, the estimated scale factors estimated at $\hat{\gamma}_1 = 0.188$ and $\hat{\gamma}_2 = 0.016$ are well below the unit factor suggesting that the common oscillations within industrial production indicator and GDP growth cycles are moderated versions of the common cycle.

Figure 1.16 reflects on how the integer and fractional delays are embodied over the set of similar cycles. Using the MLE estimates of table 1.3 we extract smoothed estimates based on Kalman filter recursive equations (see Appendix 1.8). Once the cycles have been identified, a graphical visualization provides the details of the cycle paths. In panel (a) the total integer shift of $\hat{h}_2 + 1$ over the coinciding common cycle $\tilde{\psi}_{t|n}$ is displayed. The calculated $1-$quarter lag is evident and presents itself as an exact shift by $\tilde{\psi}_{t|n}^{(\hat{h}_2+1)}$. It is evident that the total shift has not altered in any way the

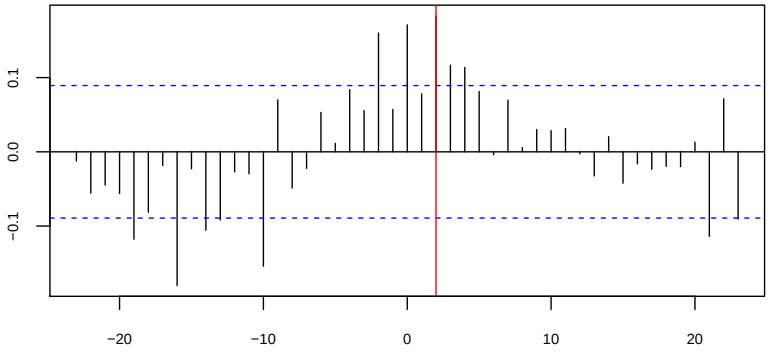
closest leftward and rightward non-missing points.

Figure 1.15: Quarterly GDP growth rate as a 3^{rd} cycle indicator (Italy)

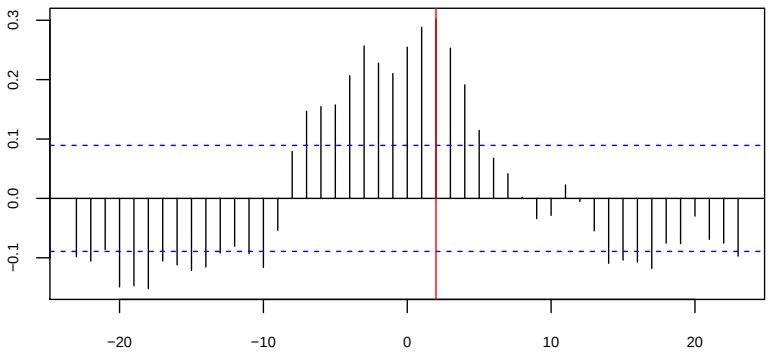
(a) Gap fill with Kalman Filter: Monthly & quarterly $\Delta_{-4}gdp_t$



(b) Cross correlogram, Italy: ΔIP_t vs. ΔOB_{t-j}



(c) Cross correlogram, Italy: $\Delta\Delta_{-4}gdp_t$ vs. ΔOB_{t-j}



Note: $\Delta_{-4}x_t=(1-L^4)x_t = x_t - x_{t-4}$ indicates the quarterly seasonal differencing. Blue horizontal lines indicate the estimated 95% confidence intervals on the ACF. Red vertical lines indicate the j^{th} lag which maximizes $\hat{\boldsymbol{\rho}}(\Delta y_{1t}, \Delta y_{0,t-j})$ for (b) and $\hat{\boldsymbol{\rho}}(\Delta y_{2,t}, \Delta y_{0,t-j})$.

shape of the coinciding cycle and the de-synchronization exhibits a perfect shift of the original cycle. The result can be extended to all the other intermediate integer steps where $h < \hat{h}_2+1$. This allows us to verify that the phase shift mechanism introduced in section 1.6 can be interpreted as an optimal filtering procedure as it accounts for the full lead/lag dynamics amongst the multivariate data set while keeping the common cycle perfectly intact. Panel (b) shows the effect of the non-integer portion

Table 1.3: ML Estimates of the 3-cycle model parameters

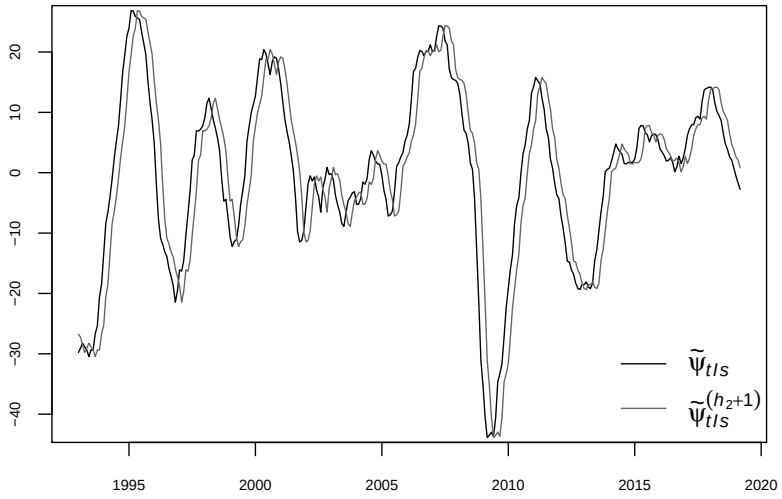
Parameter	Estimate	St.Error
$\sigma^2_{\epsilon_1}$	4.5175	0.0278
$\sigma^2_{\epsilon_2}$	2.2965	0.0108
$\sigma^2_{\epsilon_3}$	0.0000	-
$\sigma^2_{\eta_1}$	0.0002	-
$\sigma^2_{\eta_2}$	0.0041	0.0001
$\sigma^2_{\eta_3}$	0.0727	0.0002
σ^2_c	0.4593	0.0047
r_1	1.3435	0.0000
r_δ	1.6565	0.0000
ρ	0.9672	0.0003
$p = 2\pi/\lambda$	46.0693	1.6471
γ_1	0.1880	0.0007
γ_2	0.0160	-

Note: St.Error indicate the standard error, computed using the Delta method discussed in Appendix D.

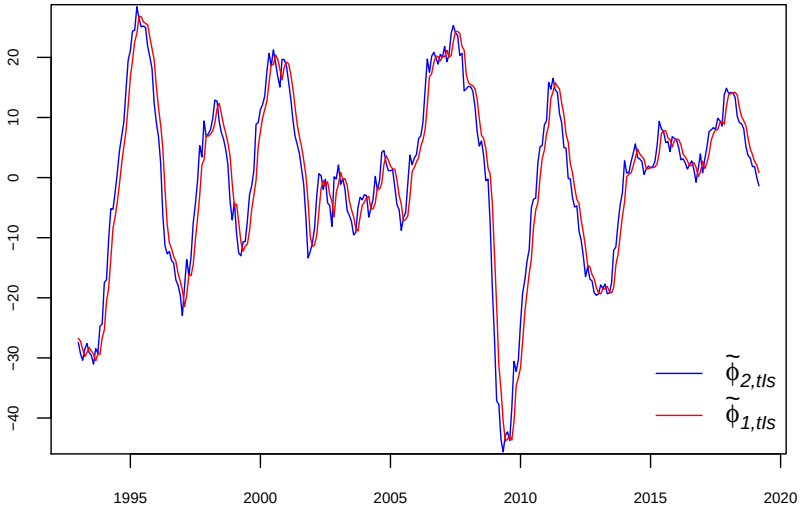
of the delay parameters; $\hat{\delta}_1$ and $\hat{\delta}_2$, that create new stochastic cycles which differ from the integer shifts $\{\tilde{\psi}_{t|n}, \tilde{\psi}^{[1]}_{t|n}, ..., \tilde{\psi}^{(\hat{h}_2+1)}_{t|n}\}$ only through error components. The fractional shifts allow the system to capture a more detailed mapping of cycles via fractional phases with visually negligible differences in their cyclical evolution. These differences which were also recorded in panels (b) and (c) of figure 1.13 is enough to lower the correlation of fractional and integer cycles and increase the model flexibility that translates into an increase in the degrees of freedom. Therefore the model achieves a statistically satisfying position. This strengthens the proposition that the filtering procedure introduced in this study for generating similar stochastic cycles is statistically optimal and capable of fully exploiting the complete information set of the lead/lag dynamics inherent in the data set. Furthermore, the inclusion of new cycles despite increasing the system dimension and complicating matters does not hamper the capability of the filtering technique with regards to reflecting and identifying information on cycle synchronicity contained within each individual data.

Figure 1.16: Cycle estimates and phase shift analysis

(a) Total integer shift, $\hat{h}_2 + 1 = 3$ over coinciding indicator



(b) Fractional shifts; ($\hat{\delta}_1 = 0.3435$, $\hat{\delta}_2 = 0.6565$)



Note: The scaling factors γ introduced in equation 1.36 for the individual cycles are excluded in the graphs. To increase the discrepancy of the plotted series, we start at 01/1993.

1.7 Conclusion

In this study we propose an optimal filtering technique applied to macroeconomic indicators in order to estimate exact cycles shifted in continuous time. Previously, works of Rünstler (2004), Valle e Azevedo et al. (2006), Koopman et al. (2016) amongst others have applied the shifting mechanism with the aim of exploiting further information from the underlying BC dynamics of leading/lagging indicators. However, the methodologies cannot be an optimal filtering technique due to it failing to take into account the information content of the data. Despite generating a shifted version of the BC, the shifted cycle exclude both the new error sequence and the model parameters which are themselves conditioned on the data. Such representation is sub-optimal considering the amount of information loss during the process.

Runstler,'s method does not restrict the phase shift (delay parameter) to only integer values. The parameter is allowed to freely move in continuous time. In this context, one can use continuous time framework to connect the phase shift parameter with the continuous time stochastic cycle representation. By using results of continuous time dynamical systems in Harvey (1990), Harvey and Stock (1993); a formula for the continuously shifted BC's is presented. Further, the new formulation incorporates a new error process for the associated shift cycle distinct form the original BC and parameter values that rely on data. As a result, the shifted stochastic cycle show closer features to the original BC compared to shifted cycles generated in Runstler work. Our study, carefully describes the mechanism that produces shifted cycles and presents the state space form representation in discrete time so that it is ready for statistical inference. We apply initially a case for a bi-variate system where one series is assigned to an integer shift and the other to a fractional shift. After providing the simulation results and an economic application, the model is extended to encompass a third phase shift dynamic which also forms the basis for all higher order phase dynamics.

From evidence provided by literature summarized in Sections 1.1-1.1.1, the choice of decomposing the data to its trend, cyclical and irregular components is the UCM. The two main reasons asserted for the choice was firstly the UCM did not fall into drawbacks of the commonly used filter techniques of HP BK, CF; and secondly it allowed the simultaneous decomposition of the data into its main components. Our study builds the UCM with a continuous time delay parameter in state space form. The state space form is a linear discrete time representation and is essential in estimating smooth estimates of cyclical components by applying Kalman recursions. To evaluate the model, the study first carries out a simulation exercise on the phase shift mechanism. The exercise reveals that the model can indeed generate phase shifted but exact versions of a common cycle from stock data. Also as a buy product of the Kalman filter, the log-likelihood function can be constructed from the prediction error and its computed variance. This makes it possible to evaluate the behavior of the log-likelihood with respect to a single continuous time parameter grid. Our findings suggest that when the parameter grid is at fractional values the model flexibility is enhanced as the fractionally phase shifted cycle generated by the model produces a new cycle. This increases the degrees of freedom and in turn

leads to higher log-likelihood levels. On the other hand, at integer ends, the parameter grid calculates lower likelihood values as either the phase does not produce any stochastic cycle, or the new cycle is identical the stochastic cycles in the system. This behavior creates a concave shape for log-likelihood at unit intervals and allows the log-likelihood level to have a global maximum. However, when an additional dimension of phase dynamics is introduced, the peak of the likelihood drops sharply at mid unit intervals due to the dependency of several continuous delay parameters. To uncover the link between model flexibility and degrees of freedom where the concept has also been conveyed in Janson et al. (2012), we set the model back to its original form i.e. we allow only the first continuous delay (the delay associated to the least non-synchronized cycle) to have a fractional delay and set all the dependent delays associated to slower cycles to integer values. This allows the extended model to converge back to the original 2 cycle model in terms of its likelihood pattern in the grid space of the first delay parameter. Therefore we conclude by stating that the model retains its optimal features even when the complexity of the phase dynamics is extended.

Our study includes two different empirical analysis. In the first part we use of a set of macroeconomic indicators: Industrial production index and Order Book retail volume (a survey based indicator) for three EU major economies: France, Germany, Italy. Smoothed cycle and trend estimates of the data show differences in terms of length, frequency, phase shifts with respect to each country common cycle. The models are able to produce exact fractional and integer shifted cycles to reveal and hence fully exploit the leading cycle dynamics contained in the data. Estimated common cycles are able to track well the recession periods. However, cycle parameters of estimated period values and phase shifts as well as variances change considerably between countries. In the second part of empirical analysis we introduce the quarterly GDP growth cycle for Italy which is transformed to a monthly frequency by applying the gap filling ability of the Kalman smoother. We add this indicator to the IP and Order Book indicator for Italy and apply the 3-cycle model to reveal the ordering assignment of lead/lag relationship for the three series. The model reveals results that are inline with the economic intuiting which assigns an ordering arrangement where Order Book series leads the IP index which in-turn leads the growth cycle of the full scale economy for Italy.

Findings of our empirical analysis shed light in to the heterogeneity of the delay patterns for major EU economies. The non-synchronicity of economic cycles across the three countries indicate that European Central Bank linked policies may require a higher level of coordination with regional monetary authorities and should take into account structural imbalances amongst EU member states. Further more, in the presence of cyclical heterogeneity between country level economic activity, a single "one fit for all" strategy may come as counter productive policy implementation for the Euro-zone.

In the second empirical analysis, revealing the ordering of economic indicators can provide policy makers an opportunity to fine tune specifically the timings of monetary policy and its guidance. Such an improvement by revealing the ordering of indicators related to economic activity could correct repercussions of the time inconsistency of monetary policy. A second possibility of identifying the ordering of

economic indicators is that the information could expose spillover effects of economic activity at sector level. This is especially relevant when specific sectors emerge as main engines of economic growth and spearhead the economy at certain periods of its growth path.

References

Basistha, A. and Nelson, C. R. (2007). New measures of the output gap based on the forward-looking new keynesian phillips curve. *Journal of Monetary Economics*, 54(2):498–511.

Başţürk, N., Çakmakli, C., Ceyhan, S. P., and Van Dijk, H. K. (2014). Posterior-predictive evidence on us inflation using extended new keynesian phillips curve models with non-filtered data. *Journal of Applied Econometrics*, 29(7):1164–1182.

Bates, S. and Ndiaye, C. T. (2014). Economic growth from a structural unobserved component modeling: The case of senegal. *Economics Bulletin*, 34(2):951–965.

Baxter, M. and King, R. G. (1999). Measuring business cycles: approximate band-pass filters for economic time series. *Review of Economics and Statistics*, 81(4):575–593.

Chen, X. and Mills, T. C. (2012). Measuring the euro area output gap using a multivariate unobserved components model containing phase shifts. *Empirical Economics*, 43(2):671–692.

Christiano, L. J. and Fitzgerald, T. J. (2003). The band pass filter. *international economic review*, 44(2):435–465.

Claessens, S., Kose, M. A., and Terrones, M. E. (2012). How do business and financial cycles interact? *Journal of International economics*, 87(1):178–190.

Cogley, T. and Nason, J. M. (1995). Effects of the hodrick-prescott filter on trend and difference stationary time series implications for business cycle research. *Journal of Economic Dynamics and control*, 19(1-2):253–278.

Creal, D., Koopman, S. J., and Zivot, E. (2010). Extracting a robust us business cycle using a time-varying multivariate model-based bandpass filter. *Journal of applied econometrics*, 25(4):695–719.

De Haan, J., Inklaar, R., and Jong-A-Pin, R. (2008). Will business cycles in the euro area converge? a critical survey of empirical research. *Journal of economic surveys*, 22(2):234–273.

Delle Monache, D. and Harvey, A. C. (2011). The effect of misspecification in models for extracting trends and cycles. Technical report, Working Papers EWP 2011/013, Euroindicators.

Doménech, R. and Gomez, V. (2006). Estimating potential output, core inflation, and the nairu as latent variables. *Journal of Business & Economic Statistics*, 24(3):354–365.

Drehmann, M., Borio, C. E., and Tsatsaronis, K. (2012). Characterising the financial cycle: don't lose sight of the medium term!,. Technical report, BIS Working Paper, 380.

Durbin, J. and Koopman, S. J. (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika*, 89(3):603–616.

Durbin, J. and Koopman, S. J. (2012). *Time series analysis by state space methods*. Oxford university press.

Galati, G., Hindrayanto, I., Koopman, S. J., and Vlekke, M. (2016). Measuring financial cycles in a model-based analysis: Empirical evidence for the united states and the euro area. *Economics Letters*, 145:83–87.

Gomez, V. (2001). The use of butterworth filters for trend and cycle estimation in economic time series. *Journal of Business & Economic Statistics*, 19(3):365–373.

Gonzalez, R. B., Lima, J., Marinho, L., et al. (2015). Business and financial cycles: an estimation of cycles' length focusing on macroprudential policy. *Central Bank of Brazil Working Papers*, (385).

Grant, A. L. and Chan, J. C. (2017). Reconciling output gaps: Unobserved components model and hodrick–prescott filter. *Journal of Economic Dynamics and Control*, 75:114–121.

Greene, W. H. (2003). Econometric analysis 7th edition international edition.

Harvey, A. (2011). Modelling the phillips curve with unobserved components. *Applied Financial Economics*, 21(1-2):7–17.

Harvey, A. and Stock, J. H. (1993). Estimation, smoothing, interpolation and distribution for structural time series models in continuous time. *Models, Methods and Applications of Econometrics*, 1:55–70.

Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter*. Cambridge university press.

Harvey, A. C. and Jaeger, A. (1993). Detrending, stylized facts and the business cycle. *Journal of applied econometrics*, 8(3):231–247.

Harvey, A. C. and Trimbur, T. M. (2003). General model-based filters for extracting cycles and trends in economic time series. *Review of Economics and Statistics*, 85(2):244–255.

Harvey, A. C., Trimbur, T. M., and Van Dijk, H. K. (2007). Trends and cycles in economic time series: A bayesian approach. *Journal of Econometrics*, 140(2):618–649.

Hodrick, R. J. and Prescott, E. C. (1997). Postwar us business cycles: an empirical investigation. *Journal of Money, credit, and Banking*, pages 1–16.

Hwa Oh, K. and Zivot, E. (2006). The clark model with correlated components. *Available at SSRN 877398*.

Ivanov, L. T. (2015). Is" the ideal filter" really ideal: the usage of frequency filtering and spurious cycles. *South-Eastern Europe Journal of Economics*, 3(1).

Janson, L., Fithian, W., and Hastie, T. (2012). Effective degrees of freedom: a flawed metaphor. *Biometrika*, 99(1):1–8.

Jarociński, M. (2015). A note on implementing the durbin and koopman simulation smoother. *Computational Statistics & Data Analysis*, 91:1–3.

Kamber, G., Morley, J., and Wong, B. (2018). Intuitive and reliable estimates of the output gap from a beveridge-nelson filter. *Review of Economics and Statistics*, 100(3):550–566.

Koopman, S. J. and Azevedo, J. V. E. (2008). Measuring synchronization and convergence of business cycles for the euro area, uk and us. *Oxford Bulletin of Economics and Statistics*, 70(1):23–51.

Koopman, S. J., Lit, R., and Lucas, A. (2016). Model-based business cycle and financial cycle decomposition for europe and the us. *TI Discussion Series*, (16-051/IV).

Larsson, G. and Tamas, V. (2012). *Comparison of Dettrending Methods*. PhD thesis, Uppsala University.

Mitra, S. and Sinclair, T. M. (2012). Output fluctuations in the g-7: An unobserved components approach. *Macroeconomic dynamics*, 16(3):396–422.

Moës, P. (2012). Multivariate models with dual cycles: implications for output gap and potential growth measurement. *Empirical Economics*, 42(3):791–818.

Mohr, M. F. (2005). A trend-cycle (-season) filter. *ECB Working Paper*, page 499.

Nelson, C. R. and Kang, H. (1981). Spurious periodicity in inappropriately detrended time series. *Econometrica: Journal of the Econometric Society*, pages 741–751.

Nilsson, R. and Gyomai, G. (2011). Cycle extraction: A comparison of the phase-average trend method, the hodrick-prescott and christiano-fitzgerald filters.

Orphanides, A. and Norden, S. v. (2002). The unreliability of output-gap estimates in real time. *Review of economics and statistics*, 84(4):569–583.

Pedregal, D. J. et al. (2001). Some comments on the use and abuse of the hodrick-prescott filter. In *Review on Economic Cycles, International Association of Economic Cycles*. Citeseer.

Pelagatti, M. M. (2006). Estimation and optimal filtering for a common stochastic cycle shifted in continuous time. *in Società Italiana di Statistica: Atti della XLIII Riunione Scientifica*.

Pelagatti, M. M. (2015). *Time series modelling with unobserved components*. CRC Press.

Perron, P. and Wada, T. (2009). Let's take a break: Trends and cycles in us real gdp. *Journal of monetary Economics*, 56(6):749–765.

Perron, P. and Wada, T. (2016). Measuring business cycles with structural breaks and outliers: Applications to international data. *Research in Economics*, 70(2):281–303.

Planas, C., Rossi, A., and Fiorentini, G. (2008). Bayesian analysis of the output gap. *Journal of Business & Economic Statistics*, 26(1):18–32.

Proietti, T. (2006). Trend–cycle decompositions with correlated components. *Econometric Reviews*, 25(1):61–84.

Rünstler, G. (2004). Modelling phase shifts among stochastic cycles. *The Econometrics Journal*, 7(1):232–248.

Rünstler, G. and Vlekke, M. (2015). Business and financial cycles: an unobserved components models perspective. Technical report, European Central Bank.

Schenk-Hoppé, K. R. (2001). Economic growth and business cycles: A critical comment on detrending time series. *Studies in Nonlinear Dynamics & Econometrics*, 5(1).

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

Van Aarle, B., Kappler, M., Sachs, A., Seymen, A., and Weyerstrass, K. (2008). Study on economic integration and business cycle synchronization. *Mannheim: Centre for European Economic Research (ZEW) and Institute for Advanced Studies (IHS)*.

Van der Vaart, A. W. (2000). *Asymptotic statistics*, volume 3. Cambridge university press.

Woitek, U. et al. (1998). *A note on the Baxter-King filter*. Department of Economics, University of Glasgow.

1.8 APPENDIX

1.8.1 Initialization of the Model

To carry out statistical inference and filtering procedures, it is required that $\alpha(t)$ is stationary. Recalling the regularly spaced continuous state vector of equation (1.15) in section 1.2.4, to satisfy stationary of $\alpha(t)$, the matrix $\boldsymbol{A}$ should have roots where

the real parts of each root should be negative. This also implies through the discrete time representation that the matrix exponential $e^{\boldsymbol{A}} = \boldsymbol{T}$ should have characteristic roots (eigenvalues) lying outside (inside) the unit circle. If the condition is satisfied, then for $z \to -\infty, e^{\boldsymbol{A}z} \to 0$. Hence $\boldsymbol{\alpha}(t)$ in equation (1.15) can be shown as

$$\boldsymbol{\alpha}(t) = \int_{-\infty}^{t} e^{\boldsymbol{A}(t-z)} \boldsymbol{R}\boldsymbol{\mu}(z)\, \mathrm{d}z \tag{1.40}$$

where the mean of $\alpha(t)$ can be calculated as $\mathbb{E}(\boldsymbol{\alpha}(t)) = \mathbf{0}$. To find the covariance matrix of $\boldsymbol{\alpha}(t)$, set $\delta_\tau = 0$ and change integral limits in equation (1.18) to $z \in [0, -\infty)$ so that

$$\mathbb{V}[\boldsymbol{\alpha}(t)] = \int_{0}^{-\infty} e^{-\boldsymbol{A}z} \boldsymbol{R}\boldsymbol{Q}\boldsymbol{R}' e^{-\boldsymbol{A}'z}\, \mathrm{d}z \tag{1.41}$$

These first two moments of $\alpha(t)$ can now be taken as the initial conditions to be used in the Kalman filter algorithm. Hence, using the discrete form state vector $\boldsymbol{\alpha}_\tau$, its linear estimator $\boldsymbol{a}_\tau$ will have the following initial conditions for its mean and variance: $\boldsymbol{a}_{1|0} = \mathbf{0}$ and $\boldsymbol{P}_{1|0} = \mathbb{V}[\boldsymbol{\alpha}(t)]$ respectively.

The Kalman filter algorithm requires an initial value for the mean vector and another for the covariance matrix of the state vector estimates denoted as; $\{\boldsymbol{a}_{1|0}, \boldsymbol{P}_{1|0}\}$. Once there is a value for the pair, under normality conditions, the filter recursions can be carried. The unobservable state vector $\boldsymbol{a}lpha_t$ contains the stochastic trend and cyclical components. The trend components are non-stationary as they follow a random walk process while the stochastic cycles are stationary. In the presence of non-stationarity components in the state vector $\boldsymbol{\alpha}_t$, the application of diffuse initial values are required. Section **??** will provide a brief computation in achieving the the diffuse values under non-stationary conditions.

1.8.2 Building the System Matrices

This section presents, the system matrices used to build the UCM and the computations required to achieve the initial value estimates for the state vector mean and covariance matrices.

Consider the state space form applied to the real data

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{Z}_t \boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t \\ \boldsymbol{\alpha}_{t+1} &= \boldsymbol{T}_t + \boldsymbol{R}_t \boldsymbol{\eta}_t \end{aligned} \tag{1.42}$$

where $\boldsymbol{y}_t = (y_{1,t}, y_{2,t})'$, $\boldsymbol{\varepsilon}_t = (\varepsilon_{1,t}, \varepsilon_{2,t})'$ are $2x1$ vectors of the observation series and error series; $\boldsymbol{\alpha}_{t+1}$ is $(h+1)x1$ vector of state variables; $\boldsymbol{\eta}_t$ is a $5x1$ vector of corresponding errors. Finally, $\boldsymbol{R}_t$, $\boldsymbol{T}_t$, $\boldsymbol{Z}_t$ are the system matrices of dimensions

$(h+10)x5$, $(h+10)x(h+10)$, $2x(h+10)$ in respective order. The error terms $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\eta}_t$ are serially uncorrelated sequences with a $2x1$ vector of $\mathbf{0}$ and covariance matrices of $\boldsymbol{H}_t$ and $\boldsymbol{Q}_t$ respectively. Lastly, both error terms are each serially uncorrelated for all t, with $t = 1, \ldots, n$.

It should be noted that, not all the states within $\boldsymbol{\alpha}_{t+1}$ have a disturbances term assigned in the state error vector $\boldsymbol{\eta}_t$. In this case, their corresponding value in $\boldsymbol{\eta}_t$ are each zero. To make the state error vector $\boldsymbol{\eta}_t$ more compact such that it only contains non-zero elements, it is pre-multiplied by matrix $\boldsymbol{R}_t$. The elements in $\boldsymbol{R}_t$ can either be 1 which indicates non-zero variances for the associated state variable errors; or zero if the error term is non-degenerate.

The system matrices and vectors of the compact form in 1.42 can be explicitly written in the following manner

$$\boldsymbol{\alpha}_t = \begin{pmatrix} \mu_{1,t} & \beta_{1,t} & \mu_{2,t} & \beta_{2,t} & \gamma_t & \gamma_t^* & \varphi_t & \psi_t^{(0)*} & \psi_t^{(0)} & \psi_t^{(1)} & \psi_t^{(2)} & \cdots & \psi_t^{(h)} & \psi_t^{(h+1)} \end{pmatrix}'$$

$$\boldsymbol{T} = \begin{array}{c} \\ \mu_1 \\ \beta_1 \\ \mu_2 \\ \beta_2 \\ \gamma \\ \gamma^* \\ \varphi \\ \psi^{(0)*} \\ \psi^{(0)} \\ \psi^{(1)} \\ \psi^{(2)} \\ \vdots \\ \psi^{(h)} \\ \psi^{(h+1)} \end{array} \begin{array}{c} \begin{array}{ccccccccccccc} \mu_1 & \beta_1 & \mu_2 & \beta_2 & \gamma & \gamma^* & \varphi & \psi^{(0)*} & \psi^{(0)} & \psi^{(1)} & \psi^{(2)} & \cdots & \psi^{(h)} & \psi^{(h+1)} \end{array} \\ \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & c & s & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & -s & -c & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & c^+ & s^+ & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & c & -s & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & s & c & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \end{array}$$

where $c = \rho \cos \lambda$, $s = \rho \sin \lambda$, $c^+ = \rho^\delta \cos(\delta\lambda)$, $s^+ = \rho^\delta \sin(\delta\lambda)$

$$\boldsymbol{\eta}_t = \begin{pmatrix} \eta_{1,t} & \eta_{2,t} & \xi_t & \xi_t^* & \upsilon_t & \kappa_t & \kappa_t^* \end{pmatrix}'$$

$$\boldsymbol{R} = \begin{array}{c} \\ \mu_1 \\ \beta_1 \\ \mu_2 \\ \beta_1 \\ \gamma \\ \gamma^* \\ \varphi \\ \psi^{(0)*} \\ \psi^{(0)} \\ \psi^{(1)} \\ \psi^{(2)} \\ \vdots \\ \psi^{(h)} \\ \psi^{(h+1)} \end{array} \begin{array}{c} \begin{array}{ccccccc} \eta_{1,t} & \eta_{2,t} & \xi_t & \xi_t^* & \upsilon_t & \kappa_t & \kappa_t^* \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

$$\boldsymbol{Q} = \begin{array}{c} \\ \mu_1 \\ \beta_1 \\ \mu_2 \\ \beta_2 \\ \gamma \\ \gamma^* \\ \varphi \\ \psi^{(0)*} \\ \psi^{(0)} \\ \psi^{(1)} \\ \psi^{(2)} \\ \vdots \\ \psi^{(h)} \\ \psi^{(h+1)} \end{array} \begin{array}{c} \begin{array}{ccccccccccccccc} \mu_1 & \beta_1 & \mu_2 & \beta_2 & \gamma & \gamma^* & \varphi & \psi^{(0)*} & \psi^{(0)} & \psi^{(1)} & \psi^{(2)} & \cdots & \psi^{(h)} & \psi^{(h+1)} \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \sigma^2_{\eta_1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \sigma^2_{\eta_2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma^2_{\gamma} & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma^2_{\gamma} & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{\upsilon} & 0 & \sigma^2_{\upsilon} & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{\kappa} & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{\upsilon} & 0 & \sigma^2_{\kappa} & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \end{array}$$

$$\boldsymbol{Z} = \begin{array}{c} \\ y_1 \\ y_2 \end{array} \begin{array}{c} \begin{array}{ccccccccccccc} \mu_1 & \beta_1 & \mu_2 & \beta_2 & \gamma & \gamma^* & \varphi & \psi^{(0)*} & \psi^{(0)} & \psi^{(1)} & \psi^{(2)} & \cdots & \psi^{(h)} & \psi^{(h+1)} \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & \chi & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \end{array}$$

$$\boldsymbol{H} = \begin{bmatrix} \sigma^2_{\varepsilon_1} & 0 \\ 0 & \sigma^2_{\varepsilon_2} \end{bmatrix}$$

Initial Vector and Covariance Matrices for Kalman Filter Recursions

It was previously discussed that, the initialization for the mean of the state vector is to set the mean value for each state variable to zero, such that

$$\boldsymbol{a}_{1|0} =$$

$$\begin{array}{c} \begin{array}{cccccccccccccc} \mu_1 & \beta_1 & \mu_2 & \beta_2 & \gamma & \gamma^* & \varphi & \psi^{(0)*} & \psi^{(0)} & \psi^{(1)} & \psi^{(2)} & \cdots & \psi^{(h)} & \psi^{(h+1)} \end{array} \\ \left[\begin{array}{cccccccccccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{array}\right] \end{array} \tag{1.43}$$

The first four states in $\boldsymbol{\alpha}_t$ corresponds to nonstationary state variables. The rest of the variables constitute the stationary part of the vector. With no further information regarding the state variables, both the stationary and non-stationary states should be grouped and considered separately. Using the exposition in Pelagatti (2015), we separate the state vector into stationary and non-stationary sub vectors, denoting them as $\boldsymbol{\alpha}_t^*$ and $\boldsymbol{\alpha}_t^+$. For the time being, let us only be concerned with the stationary sub-vector. Setting the mean of the state vector equal to $\mathbf{0}_{(4x4)}$, it is possible to write the following

$$\begin{aligned} \boldsymbol{a}^* &= \mathbb{E}[\boldsymbol{\alpha}_{t+1}^*] = \mathbb{E}[\boldsymbol{T}^*\boldsymbol{\alpha}_t^* + \boldsymbol{R}^*\boldsymbol{\eta}_t^*] \\ &= \boldsymbol{T}^*\boldsymbol{a}^* \\ &= \mathbf{0}_{(4x4)} \end{aligned} \tag{1.44}$$

with T^*, R^*, $\boldsymbol{\eta}_t^*$ corresponding to to the submatrices and sub-vectors that contain only the associated values of the stationary states. The covariance matrix of $\boldsymbol{\alpha}_{t+1}^*$ is then equal to

$$\begin{aligned} \boldsymbol{P}_1^* &= \mathbb{E}[\boldsymbol{\alpha}_{t+1}^* - \boldsymbol{a}^*][\boldsymbol{\alpha}_{t+1}^* - \boldsymbol{a}^*]^\top \\ &= \mathbb{E}[\boldsymbol{T}^*\boldsymbol{\alpha}_t^* + \boldsymbol{R}^*\boldsymbol{\eta}_t^* - \boldsymbol{a}^*][\boldsymbol{T}^*\boldsymbol{\alpha}_t^* + \boldsymbol{R}^*\boldsymbol{\eta}_t^* - \boldsymbol{a}^*]^\top \\ &= \boldsymbol{T}^*\boldsymbol{P}_1^{*}\boldsymbol{T}^{*\top} + \boldsymbol{R}^*\boldsymbol{Q}^*\boldsymbol{R}^{*\top} \end{aligned} \tag{1.45}$$

The next step is to apply the rules of the vectorization operator vec and the Kronecker product $\otimes$ so that matrix $\boldsymbol{P}_1^*$. is vectorized and its solution can be found. Using the known property that for any unknown matrix $\boldsymbol{X}$ and known matrices $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$; if $\boldsymbol{AXB} = \boldsymbol{C}$, the following identity holds

$$\operatorname{vec}(\boldsymbol{AXB}) = \operatorname{vec}(\boldsymbol{C}) = (\boldsymbol{B}^\top \otimes \boldsymbol{A}) \operatorname{vec}(\boldsymbol{X}) \tag{1.46}$$

Substituting (1.45) into (1.46), we can obtain the solution for $\boldsymbol{P}_1^*$ by solving the following equation

$$\text{vec}(\boldsymbol{P}_1^*) = (\boldsymbol{I}_{(h+10)^2} - \boldsymbol{T}^* \otimes \boldsymbol{T}^{*\top})^{-1} \text{vec}(\boldsymbol{R}^* \boldsymbol{Q}^* \boldsymbol{R}^{*\top}) \tag{1.47}$$

where the matrix $(\boldsymbol{I}_{(h+10)^2} - \boldsymbol{T}^* \otimes \boldsymbol{T}^{*\top})$ should be non-singular. The stationarity of the matrix can be evaluated by checking the matrix $\boldsymbol{T}^* \otimes \boldsymbol{T}^{*\top}$ which has eigenvalues equal to the product of the eigenvalues of matrix $\boldsymbol{T}^*$. As was previously stated in Appendix 1.8.1, for stationarity of the state vector $\boldsymbol{\alpha}_{t+1}$, it is required that $\boldsymbol{T}$ has to have eigenvalues inside the unit circle. Since the sub state vector $\boldsymbol{\alpha}_t^*$ is stationary, this implies that $\boldsymbol{T}_*$ has characteristic roots (eigenvalues) outside (inside) the unit circle. This ensures that $(\boldsymbol{I}_{(h+10)^2} - \boldsymbol{T}^* \otimes \boldsymbol{T}^{*\top})$ has eigenvalues which are also inside the unit circle. This is because each of its eigenvalues are equal to 1 minus the eigenvalues of matrix $\boldsymbol{T}^* \otimes \boldsymbol{T}^{*\top}$ which are smaller than unit value.

Lastly, the calculated elements for $\boldsymbol{P}_1^*$ which are all stacked into one vector as a result of the vectorization process in (1.47), are placed back into a square matrix with the same dimension.

To complete the covariance matrix of the complete state vector $\boldsymbol{\alpha}_{t+1}$, we also require covariance matrix for the non-stationary state sub-vector $\boldsymbol{\alpha}_t^+$. Our model incorporates only 4 non-stationary states. However, there is no further information regarding the non-stationary states, and therefore it is reasonable to make use the diffuse initialization. In this method, the researcher selects the most obvious values as estimates for the mean vector and covariance matrix. Denoting the mean and the covariance matrix of $\boldsymbol{\alpha}_t^+$ as $\mathbb{E}[\boldsymbol{\alpha}_t^+] = \boldsymbol{a}^+$ and $\boldsymbol{P}^+$, the diffuse priors will be $\boldsymbol{a}_{1|0}^+ = \boldsymbol{0}$, $\boldsymbol{P}^+ = \varsigma \boldsymbol{I}_4$ where $\varsigma \longrightarrow \infty$.

With these results, the full covaraince matrix of $\boldsymbol{\alpha}_{t+1}$ which is $\boldsymbol{P}_{1|0}$ can be initialized. Plugging in the matrices $\boldsymbol{P}_1^*$ and $\boldsymbol{P}^+$ in block diagonal form while also respecting the order of the stock variables in $\boldsymbol{\alpha}_{t+1}$, the following covaraince matrix can be written

$$\boldsymbol{P}_{1|0} = \left[\begin{array}{c|c} \boldsymbol{P}^+ & \boldsymbol{0}_{4x(h+6)} \\ \hline \boldsymbol{0}_{(h+6)x4} & \boldsymbol{P}_1^* \end{array}\right]$$

writing the covariance matrix more explicitly we have the following

$$\boldsymbol{P}_{1|0} =$$

	μ_1	β_1	μ_2	β_2	γ	γ^*	φ	$\psi^{(0)*}$	$\psi^{(0)}$	$\psi^{(1)}$	$\psi^{(2)}$	$\cdots$	$\psi^{(h)}$	$\psi^{(h+1)}$
μ_1	∞	0	0	0	0	0	0	0	0	0	0	$\cdots$	0	0
β_1	0	∞	0	0	0	0	0	0	0	0	0	$\cdots$	0	0
μ_2	0	0	∞	0	0	0	0	0	0	0	0	$\cdots$	0	0
β_2	0	0	0	∞	0	0	0	0	0	0	0	$\cdots$	0	
γ	0	0	0	0	$p_{1,1}$	0	0	0	0	0	0	$\cdots$	0	0
γ^*	0	0	0	0	0	$p_{2,2}$	0	0	0	0	0	$\cdots$	0	0
φ	0	0	0	0	0	0	$p_{3,3}$	$p_{3,4}$	$p_{3,5}$	$p_{3,6}$	$p_{3,7}$	$\cdots$	$p_{3,h+9}$	$p_{3,h+10}$
$\psi^{(0)*}$	0	0	0	0	0	0	$p_{4,3}$	$p_{4,4}$	0	0	0	$\cdots$	0	0
$\psi^{(0)}$	0	0	0	0	0	0	$p_{5,3}$	0	$p_{5,5}$	$p_{5,6}$	$p_{5,7}$	$\cdots$	$p_{5,h+9}$	$p_{5,h+10}$
$\psi^{(1)}$	0	0	0	0	0	0	$p_{6,3}$	0	$p_{6,5}$	$p_{6,6}$	$p_{6,7}$	$\cdots$	$p_{6,h+9}$	$p_{6,h+10}$
$\psi^{(2)}$	0	0	0	0	0	0	$p_{7,3}$	0	$p_{7,5}$	$p_{7,6}$	$p_{7,7}$	$\cdots$	$p_{7,h+9}$	$p_{7,h+10}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$\psi^{(h)}$	0	0	0	0	0	0	$p_{h+9,3}$	0	$p_{h+9,5}$	$p_{h+9,6}$	$p_{h+9,7}$	$\cdots$	$p_{h+9,h+9}$	$p_{h+9,h+10}$
$\psi^{(h+1)}$	0	0	0	0	0	0	$p_{h+10,3}$	0	$p_{h+10,5}$	$p_{h+10,6}$	$p_{h+10,7}$	$\cdots$	$p_{h+10,h+9}$	$p_{h+10,h+10}$

(1.48)

where $p_{i,j}$ are elements of $\boldsymbol{P}_1^*$ with $i, j = 1, ..., h+6$.

1.8.3 Delta Method for State Space parameters

The delta method allows for the computation of standard errors when they can not be computed directly from the Hessian matrix which serves as the covariance matrix of the parameter vector.

Using the results provided in Van der Vaart (2000), Greene (2003) and Pelagatti (2015); this section will give a brief summary of the delta method and how it relates to the study.

Consider that the asymptotic distribution of the estimated parameters $\widehat{\boldsymbol{\theta}}_{ML}$, which converges to a normal distribution ($asn \longrightarrow \infty$) with the following mean and covariance matrix

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_{ML} - \boldsymbol{\theta}_0) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}_{ML}}) \tag{1.49}$$

where $\boldsymbol{\theta}_0$ is the unobservable true population parameter vector and $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}_{ML}}$ is the asymptotic variance of the vector of $\widehat{\boldsymbol{\theta}}_{ML}$ which is equivalent to the inverted Fisher information matrix

$$\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}_{ML}} = \boldsymbol{I}^{-1}(\boldsymbol{\theta}_0) = \left\{ -\mathbb{E}\left[\frac{\partial^2 lnL(\boldsymbol{\theta}; \boldsymbol{y}|\mathbb{I})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \right\}^{-1} \tag{1.50}$$

Practically, $\boldsymbol{I}^{-1}(\boldsymbol{\theta}_0)$ is not evaluated since the operator $-\mathbb{E}[\cdot\,]$ indicates the expected values for $\boldsymbol{\theta}_0$, which is unknown. An alternative specification to compute the expected inverted Fisher $\boldsymbol{I}^{-1}(\boldsymbol{\theta}_0)$ is to replace it with its observed value at $\widehat{\boldsymbol{\theta}}_{ML}$ so that we have $\widehat{\boldsymbol{I}}^{-1}(\widehat{\boldsymbol{\theta}}_{ML})$ and using the Hessian of the log-likelihood. Plugging inside the squared brackets in (1.50) with the Hessian matrix of the log likelihood evaluated at $\boldsymbol{\theta}_0$ we have

$$\widehat{\boldsymbol{I}}^{-1}(\boldsymbol{\theta}_0) = \left\{-\mathbb{E}\left[\frac{1}{n}\mathbb{H}(\boldsymbol{\theta}_0)\right]\right\}^{-1} = \left\{-\mathbb{E}\left[\frac{1}{n}\frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top}\sum_{t=1}^{n} lnL(\boldsymbol{\theta};\boldsymbol{y}_t|\mathbb{I}_{t-1})\right]_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right\}^{-1}. \tag{1.51}$$

Secondly, calculating the Hessian of the log-likelihood by replacing θ_0 with $\widehat{\boldsymbol{\theta}}_{ML}$ in (1.51) we reach the following

$$\widehat{\boldsymbol{I}}^{-1}(\widehat{\boldsymbol{\theta}}_{ML}) = \left\{-\mathbb{E}\left[\frac{1}{n}\mathbb{H}(\widehat{\boldsymbol{\theta}}_{ML})\right]\right\}^{-1} = \left\{-\mathbb{E}\left[\frac{1}{n}\frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top}\sum_{t=1}^{n} lnL(\boldsymbol{\theta};\boldsymbol{y}_t|\mathbb{I}_{t-1})\right]_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_{ML}}\right\}^{-1}. \tag{1.52}$$

Hence, we reach the sampling distribution of $\widehat{\boldsymbol{\theta}}_{ML}$ shown as

$$\widehat{\boldsymbol{\theta}}_{ML} \approx N\left(\mathbf{0}, \mathbb{H}^{-1}(\widehat{\boldsymbol{\theta}}_{ML})\right) = N\left(\mathbf{0}, n^{-1}\widehat{\boldsymbol{I}}^{-1}(\widehat{\boldsymbol{\theta}}_{ML})\right) \tag{1.53}$$

Consider the score of the log-likelihood at $\widehat{\boldsymbol{\theta}}_{ML}$ which is $\left[\frac{\partial}{\partial\boldsymbol{\theta}}\sum_{t=1}^{n} lnL\boldsymbol{\theta};\boldsymbol{y}_t|\mathbb{I}_{t-1}\right]_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_{ML}}$. When the score of the log-likelihood at $\widehat{\boldsymbol{\theta}}_{ML}$ is equal to zero, the Hessian of the log-likelihood $\mathbb{H}(\widehat{\boldsymbol{\theta}}_{ML})$, must be negative semi-definite to achieve a global maxima for the log-likelihood. Secondly, as has been carried out in this study, parameter restrictions require a mapping scheme that uses the functions of parameters, such that $\boldsymbol{\theta} = \boldsymbol{g}(\boldsymbol{\varpi})$ where $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_p)'$, $\boldsymbol{\varpi} = (\varpi_1, \varpi_2, ..., \varpi_p)'$ and the vector dimensions of $px1$ and $kx1$ may not be equal.

In both the above cases, the standard errors of the parameters in $\widehat{\boldsymbol{\theta}}_{ML}$ cannot be directly retrieved by simply taking the square root of diagonal elements in $\mathbb{H}^{-1}(\widehat{\boldsymbol{\theta}}_{ML}$. The delta method allows a way to overcome this obstacle by using the vector of the parameters $\boldsymbol{\varpi}$ which have been exposed to the transformation function $\boldsymbol{g}(\cdot\,)$. The delta method requires finding the Jacobian matrix which we denote as $\boldsymbol{J}$ which its form will not be given here explicitly but is defined as the partial derivatives of $\boldsymbol{g}(\boldsymbol{\varpi})$ with respect to the elements in $\boldsymbol{\varpi}$.

In our model our variance estimates (θ_i) where $i = 1, 2, ..., 6; 10$ can be found from the transformation $\theta_{\varpi_i} = g_1(\varpi_i) = e^{\varpi_i}$, where ϖ_i represents one of the following; $ln(\sigma^2_{\varepsilon_1}), ln(\sigma^2_{\varepsilon_2}), ln(\sigma^2_{\eta_1}), ln(\sigma^2_{\eta_2}), ln(\sigma^2_c), ln(r), ln(\sigma^2_\xi)$. The parameter p for the cycle period is found through $p = \theta_p = g_2(\varpi_p) = 24+(72)/(1+e^{-\varpi_p})$, the decay parameter

ρ by using logit transformation of $\rho = \theta_p = g_3(\varpi_p) = 1/(1 + e^{-\varpi_p})$ and the scale parameter β which is straightforward since $\theta_\beta = g_4(\varpi_\beta) = \varpi_\beta = \beta$. The Jacobian $\boldsymbol{J}(\hat{\boldsymbol{\varpi}})$ is then a diagonal matrix with the derivative of each $\boldsymbol{g}(\hat{\boldsymbol{\varpi}})$ with respect to the parameters in $\boldsymbol{\varpi}$.

After computing $\boldsymbol{J}(\hat{\boldsymbol{\varpi}})$, the distribution assumption in (1.49) can be made for the estimated transformation parameter $\boldsymbol{\varpi}$ such that

$$\sqrt{n}(\hat{\boldsymbol{\varpi}} - \boldsymbol{\varpi}_0) \xrightarrow{d} N\left(\mathbf{0}, \mathbb{H}^{-1}(\hat{\boldsymbol{\varpi}})\right) \tag{1.54}$$

Setting $\widehat{\boldsymbol{\theta}} = \boldsymbol{g}(\hat{\boldsymbol{\varpi}})$, the delta method implies that [23]

$$\sqrt{n}(\boldsymbol{g}(\hat{\boldsymbol{\varpi}}) - \boldsymbol{g}(\boldsymbol{\varpi}_0)) \xrightarrow{d} N\left(\mathbf{0}, \boldsymbol{J}^\top(\hat{\boldsymbol{\varpi}})\mathbb{H}^{-1}(\hat{\boldsymbol{\varpi}})\boldsymbol{J}(\hat{\boldsymbol{\varpi}})\right) \tag{1.55}$$

The standard errors for the parameters in $\widehat{\boldsymbol{\theta}}$ are then equal to the square root of the diagonal elements in $\boldsymbol{J}^\top(\hat{\boldsymbol{\varpi}})\mathbb{H}^{-1}(\hat{\boldsymbol{\varpi}})\boldsymbol{J}(\hat{\boldsymbol{\varpi}})$.

1.8.4 3 Cycle State Space Form

The SSF representation of the 3 cycle UCM model defined by the system of equations 1.37 and 1.39 is composed of the following system vectors and matrices to build the measurement and state update equations of the Kalman filter:

$$\boldsymbol{\alpha_t} = (\boldsymbol{\alpha_t^{(\mu,\beta)}}\ \boldsymbol{\alpha_t^{\kappa}}\ \boldsymbol{\alpha_t^{\kappa^\dagger}}\ \boldsymbol{\alpha_t^{\psi}}\ \boldsymbol{\alpha_t^{\psi^\dagger}}\ \boldsymbol{\alpha_t^{\phi}})'$$

$$\boldsymbol{\alpha^{(\mu,\beta)}} = (\mu_{1t}\ \beta_{1t}\ \mu_{2t}\ \beta_{2t}\ \mu_{3t}\ \beta_{3t})'$$

$$\boldsymbol{\alpha^{\kappa}} = (\kappa_t\ \kappa_t^*)'$$

$$\boldsymbol{\alpha^{\kappa^\dagger}} = (\kappa_t^{(1)}\ \kappa_t^{*(1)}\ \kappa_t^{(2)}\ \kappa_t^{*(2)}, ..., \kappa_t^{(h_2+1)}\ \kappa_t^{*(h_2+1)})'$$

$$\boldsymbol{\alpha^{\psi}} = (\psi_t\ \psi_t^*)'$$

$$\boldsymbol{\alpha^{\psi^\dagger}} = (\psi_t^{(1)}\ \psi_t^{*(1)}\ \psi_t^{(2)}\ \psi_t^{*(2)}, ..., \psi_t^{(h_2+1)}\ \psi_t^{*(h_2+1)})'$$

[23] The covariance matrix for $\boldsymbol{g}(\hat{\boldsymbol{\varpi}})$ is found first by approximating $\boldsymbol{g}(\hat{\boldsymbol{\varpi}})$ from the first two terms of the Taylor series expansion; $\boldsymbol{g}(\hat{\boldsymbol{\varpi}}) \approx \boldsymbol{g}(\boldsymbol{\varpi}_0) + J^\top(\hat{\boldsymbol{\varpi}})(\hat{\boldsymbol{\varpi}} - \boldsymbol{\varpi}_0)$ and then using (1.54), to approximate the asymptotic variance-covariance matrix of $\sqrt{n}(\boldsymbol{g}(\hat{\boldsymbol{\varpi}}) - \boldsymbol{g}(\boldsymbol{\varpi}_0))$ such that $\mathbb{V}\left[p\text{-}\lim_{n\to\infty} n^{1/2}\left(\boldsymbol{g}(\hat{\boldsymbol{\varpi}}) - \boldsymbol{g}(\boldsymbol{\varpi}_0)\right)\right] \approx J^\top(\hat{\boldsymbol{\varpi}})\mathbb{H}^{-1}(\hat{\boldsymbol{\varpi}})\boldsymbol{J}(\hat{\boldsymbol{\varpi}})$.

$$\boldsymbol{\alpha^{\phi}} = (\phi_{1t} \;\; \phi_{2t})'$$

$$\boldsymbol{T_t} = \begin{bmatrix} \boldsymbol{T_t^{(1)}} & \boldsymbol{T_t^{(2)}} & \boldsymbol{T_t^{(3)}} & \boldsymbol{T_t^{(4)}} & \boldsymbol{T_t^{(5)}} \end{bmatrix}$$

$$\boldsymbol{T_t^{(1)}} = \begin{bmatrix} \boldsymbol{I_{[3]}} \otimes \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \\ \boldsymbol{0_{[(10+4h_2)\times 6]}} \end{bmatrix}$$

$$\boldsymbol{T_t^{(2)}} = \begin{bmatrix} \boldsymbol{0_{[7\times(2+2h_2)]}} \\ \boldsymbol{I_{[2+2h_2]}} \\ \boldsymbol{0_{[(6+2h_2)\times(2+2h_2)]}} \end{bmatrix}$$

$$\boldsymbol{T_t^{(3)}} = \begin{bmatrix} \boldsymbol{0_{[(16+4h_2)\times 2]}} \end{bmatrix}$$

$$\boldsymbol{T_t^{(4)}} = \begin{bmatrix} \boldsymbol{0_{[(12+2h_2)\times(2+4h_2)]}} \\ \boldsymbol{I_{[2+4h_2]}} \\ \boldsymbol{0_{[2\times(2+4h_2)]}} \end{bmatrix}$$

$$\boldsymbol{T_t^{(5)}} = \begin{bmatrix} \boldsymbol{0_{[(16+4h_2)\times 4]}} \end{bmatrix}$$

$\boldsymbol{T_t}$ is a square matrix with dimension $(16 + 4h_2) \times (16 + 4h_2)$, where we plug in specific values or functions which are denoted as $\boldsymbol{T_{i,j,t}}$ and corresponds to its i^{th} row and j^{th} column such that

$$\boldsymbol{T}_{11+2h_2,7,\boldsymbol{t}} = 1,\ \boldsymbol{T}_{12+2h_2,7,\boldsymbol{t}} = 1,\ \boldsymbol{T}_{11+2h_2,7,\boldsymbol{t}} = 1,\ \boldsymbol{T}_{15+4h_2,9+2h_2,\boldsymbol{t}} = 1,\ \boldsymbol{T}_{16+4h_2,9+2h_2,\boldsymbol{t}} = 1$$

$$\boldsymbol{T}_{11+2h_2,11+2h_2,\boldsymbol{t}} = \rho\cos(\lambda),\ \boldsymbol{T}_{11+2h_2,12+2h_2,\boldsymbol{t}} = \rho\sin(\lambda)$$

$$\boldsymbol{T}_{12+2h_2,11+2h_2,\boldsymbol{t}} = -\rho\sin(\lambda),\ \boldsymbol{T}_{11+2h_2,11+2h_2,\boldsymbol{t}} = \rho\cos(\lambda)$$

$$\boldsymbol{T}_{15+4h_2,11+2(h_1+h_2),\boldsymbol{t}} = \rho^{(1-\delta_1)}\cos(\lambda(1-\delta_1))$$

$$\boldsymbol{T}_{15+4h_2,12+2(h_1+h_2),\boldsymbol{t}} = \rho^{(1-\delta_1)}\sin(\lambda(1-\delta_1))$$

$$\boldsymbol{T}_{16+4h_2,11+4h_2,\boldsymbol{t}} = \rho^{(1-\delta_2)}\cos(\lambda(1-\delta_2))$$

$$\boldsymbol{T}_{16+4h_2,12+4h_2,\boldsymbol{t}} = \rho^{(1-\delta_2)} \sin(\lambda(1-\delta_2))$$

$$\boldsymbol{\eta_t} = (\eta_{1t} \;\; \eta_{2t} \;\; \eta_{3t} \;\; \kappa_t \;\; \kappa_t^* \;\; \nu_{1t} \;\; \nu_{2t})'$$

$$\boldsymbol{R_t} = \left[\boldsymbol{I}_{[3]} \otimes \begin{pmatrix} 0 & 1 \end{pmatrix}'\right] \oplus \boldsymbol{R_t^{\phi}}, \quad \boldsymbol{R_t^{\phi}} = \begin{bmatrix} \boldsymbol{0}_{[(6+4h_2)\times 2]} \\ \boldsymbol{I}_{[2]} \end{bmatrix}$$

$$\boldsymbol{Q_t} = diag\begin{pmatrix} \sigma_{\eta_1}^2 & \sigma_{\eta_2}^2 & \sigma_{\eta_3}^2 & \sigma_{\kappa}^2 & \sigma_{\kappa}^2 & \sigma_{\nu_1}^2 & \sigma_{\nu_2}^2 \end{pmatrix}$$

$$\boldsymbol{Z_t} = \left[\boldsymbol{I}_{[3]} \otimes \begin{pmatrix} 1 & 0 \end{pmatrix} \quad \boldsymbol{0}_{[(5+2h_2)\times 3]} \quad \boldsymbol{Z_t^{\psi}} \quad \boldsymbol{Z_t^{\gamma_1}} \quad \boldsymbol{Z_t^{\gamma_2}}\right]$$

$$\boldsymbol{Z_t^{\psi}} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}'$$

$$\boldsymbol{Z_t^{\gamma_1}} = \begin{pmatrix} 1 & \gamma_1 & 0 \end{pmatrix}'$$

$$\boldsymbol{Z_t^{\gamma_1}} = \begin{pmatrix} 0 & 0 & \gamma_2 \end{pmatrix}'$$

$$\boldsymbol{H_t} = diag\begin{pmatrix} \sigma_{\epsilon_0}^2 & \sigma_{\epsilon_1}^2 & \sigma_{\epsilon_2}^2 \end{pmatrix}$$

$$\boldsymbol{\hat{\alpha}_{1|0}} = \boldsymbol{0}_{[16+4h2]}$$

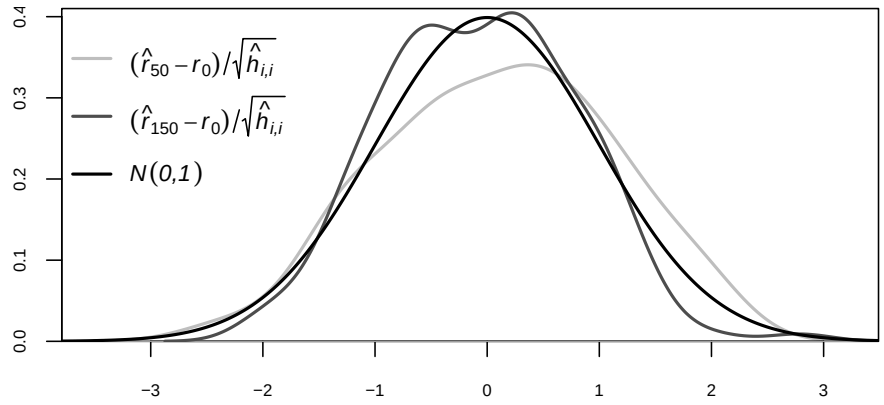
$$\boldsymbol{P_{1|0}} = \tau \boldsymbol{I}_{[6]} \oplus \boldsymbol{P^*}$$

where $\boldsymbol{P^*}$ is computed from equation 1.45.

1.8.5 MCMC estimate of the delay parameter in 3-cycle model

In the framework of section 1.4, we display the MCMC estimate related to the maximum likelihood estimate of $\hat{r}_1$. We try to assess whether the parameter distribution converges to normality when the number of iterations are small. According to figure 1.17, with a population size of $N = 500$, convergence of the distribution of the ML estimator $\hat{r}_1$ to normality is not achieved at small number of iterations (at $n = 50$ and $n = 160$). Nevertheless, there still seems to be a gradual convergence towards normality. Furthermore, from table 1.4, the average value of the 160 MCMC simulated values calculated at 2.4975, comes fairly close to the true population parameter value of 2.5.

Figure 1.17: Theoretical Distribution and the Sampling Distributions of the Maximum Likelihood Estimator $\widehat{r}_1^{(n)}$ based on 160 MCMC iterations



Note: $y - axis$ constitutes the density; $x - axis$ constitutes $\sqrt{n}H(\widehat{r}_{1_{(n)}})^{0.5}(\widehat{r}_{1_{(n)}} - r_{1_0})$ with $r_{1_{(0)}} = 2.5$.

Table 1.4: MC averages of $\widehat{r}_{1_{(n)}}$ for $n = 160$ iterations

n	$\overline{\widehat{r}}_{1_{(n)}}$
160	2.4975
	$r_{(0)}$
-	2.500

Chapter 2

Non-parametric and parametric multivariate approaches in constructing common economic cycles

2.1 Introduction

Understanding the features of cyclical patterns in an economy is vital for designing reliable macroeconomic policy. This has made business cycle analysis a prominent line of activity in macroeconomic research. In this context there has been growing literature in the methodologies that are able to extract information content related to business cycles from economic and financial data sets. While extraction of smooth cyclical components from data sets is in itself complicated and may lead to sub-optimal results (see; Nelson and Kang (1981), Mohr (2005), Pelagatti (2006), Ivanov (2015), Hamilton (2018)), the application of alternative techniques have been continuously introduced in to econometric literature.

This study implements the SSA methodology developed by Golyandina et al. (2001) and its multivariate extension introduced by Golyandina et al. (2013) and Hassani and Mahmoudvand (2013). This methodology which has recently been introduced to econometric literature is less known compared to the UCM methodology of Harvey and Jaeger (1993) and the HP-filter of Hodrick and Prescott (1997), the BK-filter of Baxter and King (1999), CF filter of Christiano and Fitzgerald (2003) and their many alternatives which have gained wide usage in both engineering and economic practice. In contrast to the UCM technique, the SSA does not rely on any theoretical economic model and different than the stated filters, its algorithm does not dependent on any parameters. More specifically, the input for its algorithm does not rely on random variables but rather on just two predetermined input parameters; the window length L which determines the structure of the data matrix input. More on the mentioned parameters and the choice of values will be conveyed in the following sections.

The way the SSA defines time series as an additive model of linear combinations of sine-cosine waves, makes it a meaningful task to overview spectral characteristics of the decomposed products. After implementing the algorithm, the study examines the periodogram for each of the extracted periodic components to understand spectral properties over the frequency spectrum close to the business cycle frequency. These components of macroeconomic data are expected to be amplitude modulated cyclical functions and their periodograms will be asymptotically a suitable spectral measure, considering that they are sampled over a large finite time interval. For finite stationary series, having de-trended series allows for pin-pointing the dominating frequencies of each time series. These cyclical components can then be used for comparison with corresponding cycles of equal frequency in terms of their phase and to confirm whether the time series are coinciding or non-synchronous in terms of their phases.

The following section provides a general literature survey of some of the economic and financial applications of the SSA

The paper contributes to business cycle literature by providing a straight forward criteria for building and tracking business cycle indicators using the non-parametric filtering technique of SSA. The three main criteria applied to the SSA extracted components are as follows

- they are common for each individual time series of the multivariate data set in question, in other words they contribute to the cyclical evolution of every series in the data
- in the frequency domain, each component peaks within the frequency spectrum characterized by the business cycle frequency range limits of 8-years and 1.5-years
- the components are grouped as pairs since they represent a sinusoidal pair: one component is associated with the sine part and the other the cosine part of the sinusoid

Once the selection is made through the above guidelines, each time series in the data set is reconstructed according to the selected components by applying the SSA algorithm. The reconstructed series is then chosen according to which one is able to best track the real business cycle indicators (the EuroCoin and the QoQ GDP growth). In terms of application, our paper shares a parallel aim of papers Altissimo et al. (2000), (2006) and de Carvalho and Rua (2017). These papers try to construct regional or national level business cycle indicators when there are competing time series variables in the data set that could play the role of a business cycle tracker once filtered. These time series contain cyclical components common amongst each other and therefore provide a more precise definition of a business cycle in this respect. In terms of methodology Altissimo et al. (2000) uses a combination of filters and also turning point analysis to construct a business cycle index for fluctuations in Italian economic activity. (2006) and de Carvalho and Rua (2017) on the other hand use a multivariate filter to construct a cycle indicator. For the former, a the parametric model of Unobserved Component Model (UCM) is applied for constructing an EU wide indicator while for the later, the non-parametric SSA is used to construct a US based cycle indicator. In this framework, our paper blends the multivariate SSA used in de Carvalho and Rua (2017) but with a different grouping strategy such that components aggregated can be defined as sinusoids. On the other hand, we also try to follow the structure and results achieved by Valle e Azevedo et al. (2006); cycle indicators for the EU region. By doing this, we examine whether non-parametric filtering techniques can achieve a successful tracking of cyclical fluctuations so that it can be seen as an alternative to parametric filtering methods. The study also includes an empirical analysis using parametric based filters. We construct common business cycles using two multivaraite UCM alternatives for extracting common cycles and and give a comparison with SSA based filter results as a robustness check.

2.2 Literature

Previously introduced by papers of Broomhead and King (1986), Broomhead et al. (1987), Vautard and Ghil (1989) for describing dynamical systems; the SSA has been used as an alternative non-parametric filtering technique in modeling physical phenomenon in various engineering applications ranging from meteorology (Palus and

Novotna (2004)), hydrology (Marques et al. (2006)), geophysics (Allen and Smith (1997), Schoellhamer (2001)), image processing (Rodriguez-Aragon and Zhigljavsky (2010)), amongst others. On the other hand, the literature regarding SSA applications to economics has been gradually on the rise only in recent years. Early studies of SSA that included economic time series is found in the monographs of Golyandina et al. (2001). The papers by Hassani et al. (2009) and Patterson et al. (2011) show the implementation of SSA to Industrial sector data which contain periodic time varying amplitudes. They are able to provide filtering and forecasting results which are able to outperform more commonly used ARMA procedures. Hassani and Thomakos (2010) provide an overall evaluation of the tools that can be implemented by both SSA and its multivariate version (M-SSA) to be used in economic applications. Hassani et al. (2010a) tests the Efficient Market Hypothesis for a set of exchange rate series through the application of a M-SSA. Hassani et al. (2010b) examine SSA and GARCH methods performances of noise reduction capabilities using stock market time series. They are able to present evidence that the SSA is superior to GARCH in terms of filtering a deterministic component that is able to track stock price behavior. Another study that looks into noise reduction of SSA methodology is carried out by Hassani et al. (2011b). Beneki et al. (2012) studies the UK tourism arrival data which contain a complex trend as well as amplitude modulated seasonal and periodic patterns. The authors show that the SSA can regroup sub-components corresponding to these movements. [1] de Carvalho et al. (2012) provide a detailed analysis of the similarities between SSA and band-pass filters of BK and CK. The study carries out a real time prediction exercise where results indicate that when the data of interest is subject to revisions, SSA emerges as a more reliable methodology compared to the classical band-pass filters.

Hassani et al. (2013) carry out a comparative analysis of extraction and prediction between the non parametric models of 1D-SSA and M-SSA and the structural parametric models of Atheson - Ohanion (Atkeson et al. (2001)) based on a dynamic linear regression and the Phillips curve based Dynamic Factor Model (Stock and Watson (2002)) using the forward looking inflation rates for a set of countries. Taking a random walk model as a benchmark, the authors suggest from their findings that the non-parametric models are superior for identification and prediction. Gonzalez et al. (2015) identify financial cycles computed from the Credit-to-GDP gap indicator using SSA and the Unobserved Component models. They apply spectral techniques to identify frequency ranges to be able confirm that financial cycles on average have higher periodicity compared to business cycles. Silva and Hassani (2015) demonstrate the application of SSA to trade data when recession based structural breaks are present in the time series. [2] Leon (2015) presents results from cross correlation and spectral measures for SSA and univariate based cycle estimates of the EU zone GDP series. His results reveal that at business cycle frequency, cycle estimates show higher level of periodicity and steeper recession phases when compared to the standard univariate filters commonly used in economics. Hassani et al.

[1] The authors refer to Ghil et al. (2002), which states that the SSA is a flexible model since it does not require any assumptions such as normality, stationarity and linearity. However its results could require further spectral techniques for the sake of robustness.

[2] Refer to Hyndman et al. (2007) for a detail discussion and Hyndman and Athanasopoulos (2018) for the implementation of optimal ARIMA, Exponential Smoothing and Neural Networks models.

(2015) applies the optimal-SSA (Hassani et al. (2012)) for tourist arrival series to compare predictability performance with respect to the forecasting approaches in Silva and Hassani (2015). The authors highlight the efficiency of the optimal SSA in separating signal from noise such and how this results in improved forecasting outcomes.[3,4,5] Skare and Buterin (2015) evaluate the complicated patterns underlying a single unemployment time series through a multivariate set of indicators relevant to the labour market. Authors, assume that the employment follows a fractionally integrated series with long memory in its mean reverting process. Their results conclude that SSA is able to recover forces that contribute to the evolution of unemployment from the multivariate data set. Sella et al. (2016) present a study that shows how the SSA can reveal the dominating cyclical components for a set of economic indicators. The study is further extended to a set of countries and includes spectral analysis which is able to identify SSA based cycle estimates with similar frequency features in the aim of uncovering cross-country wide business cycle phenomenons within the Euro area. de Carvalho and Rua (2017) tries to identify a business cycle inherent in the US output gap using 1D-SSA and M-SSA but by altering the grouping stage with a Fisher -g test. This extra step allows an alternative identification and re-grouping of the similar reconstructed components prior to the grouping step. M-SSA proves to be a better option in terms of detecting the stage of the US economy compared to the 1D-SSA. Silva et al. (2018a) applies the optimization framework of Hassani et al. (2012) to vector and recurrent versions of M-SSA.[6] Using a set of country level industrial production data, the algorithm for the optimized multivariate model is presented as a new technique capable of identifying leading indicators. An interesting M-SSA application is by de Carvalho and Martos (2020) who aim to track multiple polling data prior to Brexit election period. The M-SSA is able to give a close approximation of the general voting intention up to election day. One recent work which studies the performance of SSA approximation when mixed frequency data is involved is by Silva et al. (2018b) which presents M-SSA approach for the prediction of low frequency (quarterly) macro indicators from high frequency (monthly) indicators. Similarly, Hassani et al. (2019) find the M-SSA methodology produces accurate estimates for quarterly GDP estimates obtained from higher frequency monthly indicators. Regarding financial cycles, Skare and Porada-Rochoń (2019) employ the M-SSA to track credit and housing market cycles both of which are determinants of financial cycles. By isolating these underlying periodic components from relevant market variables, the study traces which of the financial sectors dictate the long and short-run oscillations of the financial cycles. Škare and Porada-Rochoń (2020) examines the interactions between financial and business cycles via the SSA and M-SSA approaches. The authors analyse synchronicity using cross spectral measures. Complemented by cross spectral measures, the SSA is seen to clearly identify cycles near the frequency range of both the business and financial cycles.

[3]Optimization is carried out within the framework of Hassani et al. (2012) where a forecast based lost function is minimized according to the window length L and the number of distinct eigenvalues r of the symmetric matrix at the SSA decomposition stage

[4]Refer to Golyandina et al. (2001) for a technical evaluation of the recurrent-SSA algorithm

[5]See Saayman and de Klerk (2019) for an M-SSA application of tourist arrival data results.

[6]Refer to Ghodsi et al. (2018) for a description and comparative simulation analysis for the vector (general) and recurrent forms of SSA.

2.3 Aim and Scope of the Study

As is seen from literature summarized in section 2.2, there is vast amount of evidence which is continuing to grow and show the successful applications of SSA and its modified versions in addressing many questions relevant in theoretical and applied economics. In this context, we would like to exploit the technique further as an emerging alternative in our analysis of business cycles.

In terms of the SSA methodology, most papers have concentrated in comparisons with econometric models rather than explicitly formalizing the description of the components to be decomposed and approximated. Papers by Thomakos et al. (2013), Thomakos (2008), Hassani and Thomakos (2010), de Carvalho et al. (2012), de Carvalho and Rua (2017), Thomakos and Hassani (2020) convey a clearer formalization prior to carrying out the SSA filtering procedure. Bringing the representations they have used, we also gain a better perspective to the reader on where our strategy stands amongst the applied literature at present. For example, Thomakos (2008), Hassani and Thomakos (2010) provide a form for the trend extraction by predetermining the data generating process of the stochastic trend, such that for a time series sequence $\{y_t\}_{t=1}^{n}$;

$$\begin{aligned} y_t &= x_t + \epsilon_t; \;\; y_t = \gamma t + \sum_{i=1}^{t} e_i \\ x_t &= \gamma + x_{t-1} + e_t \end{aligned} \tag{2.1}$$

The leading eigentriplets is then associated with the stochastic trend in equation 2.1. Similarly, in Hassani and Thomakos (2010) a common stochastic trend can also be identified by the SSA such that:

$$\begin{aligned} y_t &= \alpha\mu_t + \epsilon_t \\ x_t &= \mu_t + e_t \\ \mu_t &= \mu_{t-1} + \varepsilon_t \end{aligned} \tag{2.2}$$

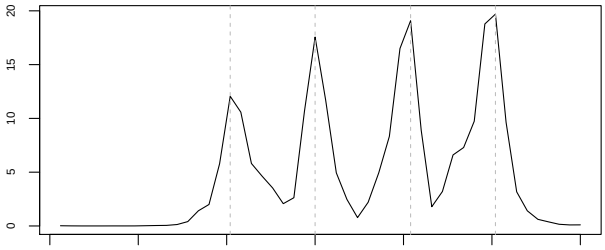
From a linear combination of the equations in 2.2, allows facilitates the representation of the common component which can then be identified and be extracted. However this form lacks the identification of a common cycle component. In Hassani and Thomakos (2010) and de Carvalho and Rua (2017), authors are able to define common cyclical components in the data generating process. More generally the specification is made in the following manner

$$\begin{aligned} y_{1t} &= \gamma_1 + R_1 \sin(2\pi\omega_1 t) + R_2 \sin(2\pi\omega_2 t) + R_3 \sin(2\pi\omega_3 t) + \epsilon_{1t} \\ y_{2t} &= \gamma_2 + R_1 \sin(2\pi\omega_1 t) + R_2 \sin(2\pi\omega_2 t) + \epsilon_{2t} \\ y_{3t} &= \gamma_3 + R_1 \sin(2\pi\omega_1 t) + R_2 \sin(2\pi\omega_2 t) + \epsilon_{3t} \end{aligned} \tag{2.3}$$

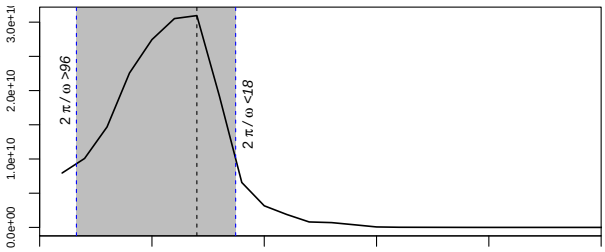
where R_i is the amplitude, ω_i is the frequency of the corresponding sinusoids. In equations 2.3, the time series only contain a limited amount of frequencies. If the

Figure 2.1: Smoothed Periodograms of SSA filtered Tourism and GDP data

(a) SSA filtered UK Tourism series (monthly)



(b) SSA filtered EU GDP series (quarterly)



Note: Both panels includes filtered series from a univariate-SSA with $L = 12$. In panel (a) the UK tourism series is the seasonal component based on 6 components; $5 - 11$ all outside the BC frequency spectrum. Panel (b) is the quarterly GDP growth series; based on components $2 - 4$ all of which fall with in the BC frequency spectrum indicated by the grey area..

data contained seasonal periodicity then such a specification would be meaningful. On the other hand, when a business cycle component is inherent in the data then a spectrum of frequencies would be necessary to represent the data in the frequency domain. Figure 2.1 shows the discrepancies between the periodograms of two different types of data; one with a seasonality component and one with a business cycle component. The smoothed periodogram of UK tourism data contain a predominant seasonality pattern and panel (a) gives a good example of the behaviour of its spectral density over a frequency band. Notice the sharp peaks with respect to panel (b) which shows the periodogram of a business cycle component filtered from the GDP series. As can be seen, the peak is not concentrated and dispersed along the neighborhood spectrum restricted by the limits of the business cycle frequency (1.5 years and 8 years). If our filter is able to capture all the frequency spectrum with-in the shaded area and the corresponding components, then we would have a realistic representation of a business cycle indicator.

In this framework, for a larger data set, it is possible to modify equation 2.3 along the above lines, so that it includes higher linear combination of harmonic components such that

$$
\begin{aligned}
y_{1t} &= \mu_{1t} + \sum_{i=1}^{m} \left[a_i^{(1)} \cos \lambda_i t + b_i^{(1)} \sin \lambda_i t \right] + \epsilon_{1t} \\
y_{2t} &= \mu_{2t} + \sum_{i=1}^{m} \left[a_i^{(2)} \cos \lambda_i t + b_i^{(2)} \sin \lambda_i t \right] + \epsilon_{2t} \\
&\dots \\
y_{pt} &= \mu_{pt} + \sum_{i=1}^{m} \left[a_i^{(p)} \cos \lambda_i t + b_i^{(p)} \sin \lambda_i t \right] + \epsilon_{pt} \\
\mu_{1t} &= \mu_{1,t-1} + \beta_{1,t-1} + e_{1t} \\
\beta_{1t} &= \beta_{1,t-1} + \upsilon_{1t} \\
&\dots \\
\mu_{pt} &= \mu_{p,t-1} + \beta_{p,t-1} + e_{pt} \\
\beta_{pt} &= \beta_{p,t-1} + \upsilon_{pt}
\end{aligned}
\tag{2.4}
$$

where $\epsilon_{1t}, ..., \epsilon_{pt}, \upsilon_{1t}, ..., \upsilon_{pt}$ are each white noise sequences and similarly $e_{1t}, ..., e_{pt}$ are error sequences with but with $\sigma^2_{e_1} = ... = \sigma^2_{e_p} = 0$, $\lambda_i \in \mathbb{R}$ are the corresponding frequencies of the harmonic component. For the coefficients $a_i^{(j)}$, $b_i^{(j)}$, where $j = 1, ..., p$, we use basic principles of cyclical time series. The sine/cosine pairs in the above system of equations share one to one correspondence with the sine wave representation of

$$
R_{ji} \sin(\lambda_i t + \phi_{ji}) \tag{2.5}
$$

where $R_{ji} \in \mathbb{R}$ is the amplitude of the harmonic wave, $\phi_i \in [-\pi,\ \pi]$ is the phase of the wave both of which satisfy the following principles:

$$
\begin{aligned}
a_i^{(j)} &= R_{ji} \sin \phi_{ji} & b_i^{(j)} &= R_{ji} \cos \phi_{ji} \\
R_{ji} &= \sqrt{\left(a_i^{(j)}\right)^2 + \left(b_i^{(j)}\right)^2} & \phi_{ij} &= arctan \frac{a_i^{(j)}}{b_i^{(j)}}
\end{aligned}
\tag{2.6}
$$

In our study we look into two main issues that can be traced using both equations 2.3-2.6. Firstly, we try to collect all frequency pairs (sine/cosine waves) such that $2\pi/96 < \lambda_1, ..\lambda_j < 2\pi/18$. Secondly, since cyclical pairs can be represented by equation 2.5, it would be possible to also examine lead lag relationship amongst the common cyclical components associated with each individual time series in our multivariate data. A final note on equation 2.3 will be the remaining components that fall outside our business cycle definition but inherent in the data. These are the trend, higher frequency waves and noise. Regarding trend like behaviour, these can be passed on to components: $\mu_{1t}, ..., \mu_{pt}$ and be treated as stochastic trends or low frequency components below business cycle frequency limits.[7] For high frequency

[7]Our study does not examine commonality in the trend component which can refer to a cointegration relationship in the data set if the trend is stochastic.

periodic patterns which can refer to seasonality, irregularity, error terms, all these can be attributed to $\epsilon_{1,t}, ..., \epsilon_{p,t}$.[8] Therefore, we are able to model a complex periodic series associated to the business cycle as a linear combination of sine/cosine pairs where each of the sine/cosine pairs are common to all the time series of a multivariate data set.

All in all, our study partially rely on frequency analysis which gives a wider perspective to the researcher when observing harmonic components of macroeconomic indicators. It is also a much more convenient way of expressing phase differences and similarity of cycles. Hence, our results can give further information content emerging through decomposition of time series and their analysis through the frequency domain that may not be obvious from the analysis of time domain properties of cycles extracted from parametric filtering processes.

2.4 Singular Spectrum Analysis and Constructing a Business Cycle Indicator

The SSA methodology has both univariate and multivariate versions.[9] The multivariate versions of H-MSSA (horizontal multivariate SSA) and V-MSSA (vertical multivariate SSA) can also further be adopted to accommodate unbalanced or missing data sets (the series differ in length). In this study, we select the V-MSSA firstly because it is one of the multivariate versions of the SSA which can deal with a system of time series data but simply because of the way the data matrix (the first step of the SSA algorithm) is organized. It is easier to envisage each univariate data being stacked vertically rather than in a horizontal fashion when forming the multivariate data matrix. The structure of the data matrix will be further explained in section 2.4.1. Hassani and Mahmoudvand (2018), Golyandina (2020) indicate that there is no difference between H-MSSA and V-MSSA other than computational aspects. Despite no clear difference between the two multivariate versions, in this study we set our framework to V-MSSA.

There are two separate stages within the SSA where both stage require two steps each to be completed. Each of these steps allows for decomposing and identification of a long term trend, oscillating components and noise. The following sections will give the description of these stages for the single (1D-SSA) and multivariate (H-MSSA) cases. The section incorporates some of the framework and definitions in Golyandina et al. (2001), Hassani et al. (2009), Golyandina et al. (2018), Hassani and Mahmoudvand (2018). For the sake of completeness and to show the steps in eventually achieving a cycle indicator, the following section will give a brief summary of the algorithm to be used.

[8] Since the data we examine are seasonally adjusted, ϵit's can be made to account for noise and all periodic patterns with frequencies beyond the upper limit of the business cycle limit.

[9] For a further discussion on the variations of SSA (oblique and derivative), see Golyandina and Shlemov (2013).

2.4.1 Forming and Decomposing a Data Matrix for multivariate macroeconomic data

We assume a multivariate data set formed by sequences of individual time series data; $\mathbb{F}_t = \left(\mathbb{F}_t^{(1)}, \mathbb{F}_t^{(2)}, ..., \mathbb{F}_t^{(p)}\right)$, where each series has the following sequence $\mathbb{F}^{(j)} = \{f_1^{(j)}, f_2^{(j)}, ..., f_N^{(j)}\}$, $j = 1, .., p$. In terms of economic realizations it would be more meaningful to write $\mathbb{F}_t = \{\mathbb{F}_t\}_{t=1}^{N}$.[10]

The decomposition stage (the first stage of SSA algorithm) starts with the embedding step and followed by the decomposition step explained in the next sections. The embedding results in a mapping of the data sequence to a matrix form with K number of $P \times L$ sized column vectors, where $K = N - L + 1$, L is an integer and also named as the window length which is defined to be $2 \leq L \leq N/2$. The output from this step is an $pL \times K$ matrix with Hankel properties and is named as the trajectory matrix $\boldsymbol{X} = (x_{i,j})_{i,j}^{pL,K}$, shown explicitly as

$$\boldsymbol{X} = \boldsymbol{J}_{MD}\left(\mathbb{F}_t\right) =$$

$$\begin{bmatrix} x_{11} & x_{12} & x_{31} & \cdots & x_{1K} \\ x_{21} & x_{22} & x_{32} & \cdots & x_{2K} \\ x_{31} & x_{22} & x_{23} & \cdots & x_{2K} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{L1} & x_{L2} & x_{L3} & \cdots & x_{LK} \\ & & \cdot\cdot & & \\ x_{L+1,1} & x_{L+1,2} & x_{L+1,3} & \cdots & x_{L+1,K} \\ x_{L+2,1} & x_{L+2,2} & x_{L+2,3} & \cdots & x_{L+2,K} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{2L,1} & x_{2L,2} & x_{2L,3} & \cdots & x_{2L,K} \\ & & \cdot\cdot & & \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{(n-1)L+1,1} & x_{(n-1)L+1,2} & x_{(n-1)L+1,3} & \cdots & x_{(n-1)L+1,K} \\ x_{(n-1)L+2,1} & x_{(n-1)L+2,2} & x_{(n-1)L+2,3} & \cdots & x_{(n-1)L+2,K} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{nL,1} & x_{nL,2} & x_{nL,3} & \cdots & x_{nL,K} \end{bmatrix} = \begin{bmatrix} f_1^{(1)} & f_2^{(1)} & f_3^{(1)} & \cdots & f_K^{(1)} \\ f_2^{(1)} & f_3^{(1)} & f_4^{(1)} & \cdots & f_{K+1}^{(1)} \\ f_3^{(1)} & f_4^{(1)} & f_5^{(1)} & \cdots & f_{K+2}^{(1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_L^{(1)} & f_{L+1}^{(1)} & f_{L+2}^{(1)} & \cdots & f_N^{(1)} \\ & & \cdot\cdot & & \\ f_1^{(2)} & f_2^{(2)} & f_3^{(2)} & \cdots & f_K^{(2)} \\ f_2^{(2)} & f_3^{(2)} & f_4^{(2)} & \cdots & f_{K+1}^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_L^{(2)} & f_{L+1}^{(2)} & f_{L+2}^{(2)} & \cdots & f_N^{(2)} \\ & & \cdot\cdot & & \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ f_1^{(p)} & f_2^{(p)} & f_3^{(p)} & \cdots & f_K^{(p)} \\ f_2^{(p)} & f_3^{(p)} & f_4^{(p)} & \cdots & f_{K+1}^{(p)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_L^{(p)} & f_{L+1}^{(p)} & f_{L+2}^{(p)} & \cdots & f_N^{(p)} \end{bmatrix} \tag{2.7}$$

where $\boldsymbol{J}_{MD}$ is a linear mapping operator mapping each of the p time series in the $\mathbb{R}^N$ space onto the $L \times K$ matrix space and vertically stacking them on top one by one.[11] The mapping basically selects a window size parameter L and slides the window size through the span of each data with length Ns. The vertically stacked trajectory matrix in equation 2.7 is then written more compactly in the following way

[10] While it is possible to work with different data length, we assume that each individual time series $\mathbb{F}^{(j)}$ has equal length N.

[11] pL is a summation such that $\sum_{i=1}^{p} L_i = \sum_{i=1}^{p} L = pL$ since all series are of equal length.

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}^{(1)} \\ .. \\ \boldsymbol{X}^{(2)} \\ .. \\ \vdots \\ .. \\ \boldsymbol{X}^{(p)} \end{bmatrix} \tag{2.8}$$

Each of the blocks $\boldsymbol{X}^{(j)}$ have dimensions $L \times K$ and represent one of the time series data $\mathbb{F}^{(j)}$ in matirx space. Through the operator $\boldsymbol{J}_{MD}$, every $\boldsymbol{X}^{(j)}$ is transformed to be of Hankel structure. This basically means that for each skew diagonals in $\boldsymbol{X}^{(j)}$ the elements remain constant as it ascends from the upper right side to the lower left side of the matrix.[12] I final note for the embedding step would be on how the value of the window parameter L is determined.

Before moving to the next step, an important note would be made on the parameter L. According to Hassani et al. (2009) L should take into account the frequency of the data and select a value that is a multiple to the known periodicity of the time series in question. For example, a quarterly series which happens to be influenced by annual length seasonality should have a window length parameter multiplicative of 4 and a minimum value that can capture the full periodicity inherent in the evolutionary process of the series. In de Carvalho and Rua (2017), $L = 32$ as they aim to estimate output gap based on GDP series which is in quarterly frequency and having a periodicity of the business cycle ($\sim$ 8-years) which translates to 32 quarters. On the other hand Golyandina (2010) states that for separability of the signal-noise component, L should be as close enough to $N/2$ or at least only a few periods away if the period length is not large and when the determinant period is known to the researcher. Regarding separability, the theoretical paper by Hassani et al. (2011a) and a simulation study by Mahmoudvand et al. (2013) provide a second strategy. According to these papers, the best signal-noise separability is achieved by simply selecting $L = median\{1, ..., N\}$.[13] Overall, taking these in to consideration, we select the window length rule according to Hassani and Thomakos (2010) and Golyandina (2010). Our reasoning is that the signal reconstruction approach in these papers is a similar purpose of our paper which aims to construct a common signal (the business cycle).

The next step is to decompose $\boldsymbol{X}$ such that one can retrieve its eigen-vectors and eigen-values. Since $\boldsymbol{X}$ is not usually a positive symmetric matrix (mostly $L \neq K$), the positive symmetric matrix $\boldsymbol{X}\boldsymbol{X}^{\top}$ can allow computing eigen-vectors and eigen-values such that

[12] If we were only dealing with a single time series, than the trajectory matrix would not be stacked and would therefore have a Hankel structure.

[13] Separability can be tested using the filtered based correlation measure coined as $w-$correlation; see Hassani et al. (2009) for a detailed definition.

$$\begin{aligned}\underset{(pL\times pL)}{\boldsymbol{X}\boldsymbol{X}^{\top}} &= \begin{bmatrix}\boldsymbol{X}^{(1)}\\ \cdot\cdot\\ \boldsymbol{X}^{(2)}\\ \cdot\cdot\\ \vdots\\ \cdot\cdot\\ \boldsymbol{X}^{(p)}\end{bmatrix}\left[\boldsymbol{X}^{(1)\top}:\boldsymbol{X}(2)\top:\cdots:\boldsymbol{X}^{(p)\top}\right]\\ &= \begin{bmatrix}\boldsymbol{X}^{(1)}\boldsymbol{X}^{(1)\top}:\boldsymbol{X}^{(1)}\boldsymbol{X}^{(2)\top}:\cdots:\boldsymbol{X}^{(1)}\boldsymbol{X}^{(p)\top}\\ \boldsymbol{X}^{(2)}\boldsymbol{X}^{(1)\top}:\boldsymbol{X}^{(2)}\boldsymbol{X}^{(2)\top}:\cdots:\boldsymbol{X}_2\boldsymbol{X}_p^{\top}\\ \cdot\cdot\quad:\quad\cdot\cdot\quad:\cdots:\quad\cdot\cdot\\ \vdots\quad\quad\vdots\quad\quad\ddots\quad\quad\vdots\\ \cdot\cdot\quad:\quad\cdot\cdot\quad:\cdots:\quad\cdot\cdot\\ \boldsymbol{X}^{(p)}\boldsymbol{X}^{(1)\top}:\boldsymbol{X}^{(p)}\boldsymbol{X}^{(2)\top}:\cdots:\boldsymbol{X}^{(p)}\boldsymbol{X}_p^{\top}\end{bmatrix}\end{aligned} \tag{2.9}$$

where one can compute the eigen-value and eigen-vectors of $\boldsymbol{X}\boldsymbol{X}^{\top}$ from the diagonal blocks in equations 2.9. However, the SSA methodology decomposes the trajectory matrix to its eigen-value/vectors by applying the Singular Value Decomposition (SVD) so that

$$\underset{(pL\times K)}{\boldsymbol{X}} = \underset{(pL\times pL)}{\boldsymbol{U}} \times \underset{(pL\times K)}{\boldsymbol{Q}} \times \underset{(K\times K)}{\boldsymbol{V}^{\top}} \tag{2.10}$$

where the columns of $\boldsymbol{U}$ and rows of $\boldsymbol{V}^{\top}$ contain the left and right singular-vectors (eigenvectors) while $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r \geq 0$ are the singular-values (square root of the eigenvalues) of $\boldsymbol{X}$ ($\boldsymbol{X}\boldsymbol{X}^{\top}$). The compact representation of the SVD in equation 2.10 is a sum of rank-1 matrices which also constitute the SVD of the vertical blocks $\boldsymbol{X}^{j}$'s. These rank-1 matrices can each be found separately by taking the corresponding columns (U_i), row ($V_i^{\top}$) and singular-values (σ_i) in equation 2.10 so that

$$\boldsymbol{X}_i = \sigma_i U_i V_i^{\top}, \qquad (\boldsymbol{X}_i = 0 \quad if \quad \sigma_j = 0) \tag{2.11}$$

with $2 \leq i \leq pL$. Taking the cumulative of all $\boldsymbol{X}_i$'s, the following expansion can be reached

$$\begin{aligned}\boldsymbol{X} &= \sum_{i=1}^{pL}\boldsymbol{X}_i\\ &= \sum_{i=1}^{pL}\sigma_i U_i V_i^{\top}\end{aligned} \tag{2.12}$$

Equation 2.12 is an alternative representation of 2.7 where the trajectory matrix has been expanded/decomposed into rank-1 components with similar dimensions

of $pL \times K$. Another way at looking to the expansion of equation 2.12 is that for every rank 1 matrix $\boldsymbol{X}_i$, there associates an eigentriple $\{\sigma_i, U_i, V_i\}$ where its multiplicative sum can compute back the original data matrix $\boldsymbol{X}$. The columns of the orthogonal matrix $\mathbf{U}$ form the orthonormal basis of the column space ($\mathbb{R}^{pL}$) of the trajectory matrix ($\boldsymbol{X}$) while $\mathbf{V}$ form the basis of the row space ($\mathbb{R}^{K}$). The pL length columns are composed of p number of L length sub-columns, where each of these sub-columns can be traced back to the individual time series of the multivariate data set $\{f_{i+1}^{(j)}, f_{i+2}^{(j)}, ..., f_{i+L}^{(j)}\}_{i=0}^{K-1}$ for $j = 1, .., p$. [14] This explanation will be important later on in the empirical analysis stage which we will further evaluate the sub columns of the singular-vectors.

To sum up the overall significance of this stage of the algorithm, we can say that that the main intuition behind the SVD step, is to break the data into its main arguments. By doing so, we are at the first step to distinguish between signal and noise of a given time series data.

2.4.2 Reconstructing and approximating components for a set of related macroeconomic data

The second stage is composed of two steps: the grouping and diagonal averaging. This stage leads to isolate the different characteristics of the data series. We consider the decomposed matrix components provided in section 2.4.1 as input in the final phase of the algorithm: reconstruction and estimation. To reconstruct the main components, the grouping step requires an evaluation of the rank 1 matrices ($\boldsymbol{X}_i$) followed by the separation of the matrices into sub-groups that do not intersect. This kind of separation is done via indices $I_1, I_2, ..., I_l$, where $l \leq pL$. A generic sub-index is formed by $I_r = i_1, ...i_r$ such that $i_i \in \boldsymbol{S} = 1, 2, ..., r_x$ with the index satisfying the properties that $\forall\, I_r \subset S$ and $I_j \cap I_l = \emptyset$ for $j \neq l$. Previously stated in section 2.3, generally speaking the elementary matrices do not contain all the necessary information for a given sub-component of a time series (i.e. the trend, oscillations, noise, etc). They require to be represented by a collection of elementary components. As a result the grouping step incorporates and links the unit rank matrices to form a meaningful feature of the data. Since all the neccessary definitions have been made, we can proceed to formulate the grouping step as

$$\begin{aligned} \boldsymbol{X} &= \boldsymbol{X}_1 + \boldsymbol{X}_2 + \boldsymbol{X}_3 + ... + \boldsymbol{X}_{r_x} \\ &= \sum_{i \in I_1} \boldsymbol{X}_i + \sum_{i \in I_2} \boldsymbol{X}_i + ... + \sum_{i \in I_l} \boldsymbol{X}_i \\ &= \boldsymbol{X}_{I_1} + \boldsymbol{X}_{I_2} + ... + \boldsymbol{X}_{I_l} \\ &= \sum_{s=1}^{l} \boldsymbol{X}_{I_s} \end{aligned} \tag{2.13}$$

Regarding equation 2.13, a simple example can be considered for a better understanding. Assume the researcher evaluates that the multivariate system can be

[14] Notice that the columns of $\boldsymbol{U}$ will be spanning the entire data set.

described by 2 types of components; a long-term trend and a seasonal oscillation which are associated with the indices $\{i_1, i_2, ..., i_j\} = I_1$ and $\{i_{j+1}, i_{j+2}, ..., i_{l-u}\} = I_2$ which are computed by the SVD step. Similarly, the rest of the matrices not included in I_1 and I_2 are then assumed not to be defining the multivariate data and represented by the remaining components $\{i_{l-u+1}, i_{l-u+2}, ...i_{pl}\} = I_3$. In this context, the indices can be split in to the main groups which can be passed on to the matrix structures; the trend: $\boldsymbol{X}_{I_1}$, the seasonal pattern $\boldsymbol{X}_{I_2}$ and the error: $\boldsymbol{X}_{I_3}$. In this case; $\boldsymbol{X} = \boldsymbol{X}_{I_1} + \boldsymbol{X}_{I_2} + \boldsymbol{X}_{I_3}$. The two extreme examples for the grouping stage would be complete separability at one end and inseparability of components at the other end. In the former, every $\boldsymbol{X}_i$ can be represented by a single index such that $I_1 = \{1\}, I_2 = \{2\}, ..., I_l = \{r_x\}$. For the later case, a time series can be represented by a single component constructed by all the components; $I_1 = \{1, 2, ..., L\}$.

To see how well the grouping strategy is realized in the time domain, we need to remap the grouped matrices into multivariate time series sets. However, we should take note that the unit rank matrices $\boldsymbol{X}_j$ and as a result their grouped versions will not have identical elements on their skew-diagonal entries. Therefore, they are not of a Hankel structure. This obstructs the SSA algorithm to reconstruct matrix elements back into multivariate time series sets. To guarantee a reconstruction process, the grouped matrices $\boldsymbol{X}_{I_s}$ or the elementary matrices before the grouping stage $\boldsymbol{X}_j$ are needed to be Hankelized. To formulate this step, consider the element of a grouped matrix is $\{x\}_{i,j}^{pL,K}$. The diagonlization of its associated matrix involves the averaging of every element along its skewed diagonal $i + j = k + 2$, $\forall k$ where $0 \leq k \leq pL + K - 2$. The operator for this transformation can be denoted as $\boldsymbol{J}_v$. In this framework the following expression can be reached

$$\begin{aligned} \boldsymbol{J}_v \boldsymbol{X} &\equiv \boldsymbol{J}_v \left(\sum_{s=1}^{l} \boldsymbol{X}_{I_s} \right) \\ \boldsymbol{X} &\equiv \boldsymbol{J}_v \boldsymbol{X}_{I_1} + \boldsymbol{J}_v \boldsymbol{X}_{I_2} + ... + \boldsymbol{J}_v \boldsymbol{X}_{I_l} \\ &\equiv \tilde{\boldsymbol{X}}_{I_1} + \tilde{\boldsymbol{X}}_{I_2} + ... + \tilde{\boldsymbol{X}}_{I_l} \\ &\equiv \sum_{s=1}^{l} \tilde{\boldsymbol{X}}_{I_s} \end{aligned} \tag{2.14}$$

using the principles; that $\boldsymbol{J}_v \boldsymbol{X} = \boldsymbol{X}$ since $\boldsymbol{X}$ is a Hankel matrix which means its Hankelization is itself, and the linearity in $\boldsymbol{J}_v$ such that for any rectangular matrix $\boldsymbol{Y}$, $\boldsymbol{Z}$ the following holds: $\boldsymbol{J}_H(\boldsymbol{Y} + \boldsymbol{Z}) = \boldsymbol{J}_H \boldsymbol{Y} + \boldsymbol{J}_H \boldsymbol{Z}$.[15]

The resulting $\tilde{\boldsymbol{X}}_{I_s}$'s will have elements $\{\tilde{x}_{yz}^{I_s(j)}\}_{y,z=1}^{L,K}$ that remain constant along the skew diagonals at every block where $1 \leq j \leq p$. Every diagonal will correspond to a single observation of the component belonging to one of the individual time series data $\tilde{\mathbb{F}}_t^{(j,s)}$ associated with the j'th block. This brings us to the final step of the SSA, where it is now possible to transform matrices $\tilde{\boldsymbol{X}}_{I_s}$'s as sub-components of the estimates of the original data $\mathbb{F}_t^{(s)}$'s using the reverse operator $\boldsymbol{J}_{1D}^{-1}$. The complete set of approximated components can be achieved by

[15]For a detalied explanation on the rules that are used for the diagonalization process of a vertically stacked group matrix $\tilde{\boldsymbol{X}}_{I_s}$, see section 2.8.1.

$$\boldsymbol{J}_{MD}^{-1}\left(\sum_{s=1}^{l}\tilde{\boldsymbol{X}}_{I_s}\right)=\sum_{j=1}^{p}\sum_{s=1}^{l}\tilde{\mathbb{F}}_t^{(j,s)} \tag{2.15}$$

where $1 \leq j \leq p$ is the number of time series, $1 \leq s \leq l$ is the number of sub-components and $t = \{1, 2, ..., N\}$ is the length of the time series components. Hence, the time series approximations can be summarized in the following step

$$\begin{aligned}
\tilde{\mathbb{F}}_t^{(1)} &\equiv \tilde{\mathbb{F}}_t^{(1,1)} + \tilde{\mathbb{F}}_t^{(1,2)} + ... + \tilde{\mathbb{F}}_t^{(1,l)} \\
\tilde{\mathbb{F}}_t^{(2)} &\equiv \tilde{\mathbb{F}}_t^{(2,1)} + \tilde{\mathbb{F}}_t^{(2,2)} + ... + \tilde{\mathbb{F}}_t^{(2,l)} \\
&\vdots \\
\tilde{\mathbb{F}}_t^{(p)} &\equiv \tilde{\mathbb{F}}_t^{(p,1)} + \tilde{\mathbb{F}}_t^{(p,2)} + ... + \tilde{\mathbb{F}}_t^{(p,l)}
\end{aligned} \tag{2.16}$$

According to Golyandina et al. (2001), Hassani et al. (2009), the complete process of the reconstruction step is an optimal procedure. In the framework of these papers, $\|\boldsymbol{X}_{I_s} - \tilde{\boldsymbol{X}}_{I_s}\|_{\boldsymbol{F}}$ leads to the minimal $\|\cdot\|_{\boldsymbol{F}}$ amongst all other Hankel matrices with dimensions $L \times K$ and rank $r_{\boldsymbol{X}_{I_s}}$. In other words, the matrices $\tilde{\boldsymbol{X}}_{I_s}$'s are the closest in distance of all possible matrix computations with similar structure and lying in the same linear space to $\boldsymbol{X}_{I_s}$'s (matrices directly decomposed from the data matrix $\boldsymbol{X}$). The implication of this is that the SSA extracted components $\tilde{\mathbb{F}}_t^{(s)}$ achieve the most accurate approximation and the reconstruction step alters the data at minimal level attainable.

2.4.3 Obtaining peak frequencies cycle pairs at the Grouping Step

The characteristics of the underlying components in the frequency domain will be the main determinant in the grouping strategy of our SSA application. When the multivariate time series data is decomposed into their rank-1 components ($\boldsymbol{X}_j$), they will not be grouped based on any initial strategy. Skipping the grouping stage are reconstructed back to an approximated component of the original data set, we apply basic spectral analysis. The spectral analysis will conclude 2 important categorizations. Firstly, it will decide whether the component is within the business cycle frequency spectrum. Secondly, it will decide whether all the frequencies are in pairs such that every peak frequency has an equivalent to be matched and represent a sine/cosine pair.

For the analysis, the periodogram allows the identification of the peak frequencies. According to the lower and upper limits of a business cycle frequency, the peak frequencies are collected along the spectrum such that $\lambda_{LL} \leq \{\lambda_z^*\}_{z=min(z)}^{r_{BC}} \leq \lambda_{UL}$ where $min(z) \geq 1$ and $r_{BC} \leq l = pL$. This identification step deals with criteria-1 which allows the peak frequencies to be grouped so that $\{\lambda_z^*\}_{z=min(z)}^{r_{BC}} \in \Psi_{BC}$. Criteria-2 requires each peak frequency within Ψ_{BC} to be equal by at least another peak frequency in the group. This brings us to criteria-2 by matching the peak

frequencies in Ψ_{BC} so that $\lambda^*_{z^*_m} = \lambda^*_{z^*_n}$ where $min(z^*) \leq z^*_m, z^*_n \leq max(z^*)$ and z^* is an irregularly spaced sequence of indexes and a subset of z ($z^* \in z$) such that $min(z^*) \geq 1$ and $max(z^*) \leq r_{BC}$.

Since peaks are the dominant frequency along the full spectrum and therefore determine the frequency characteristics of the associated component, the frequency pair can represent a sinusoid pair such that it can be written in form of equation 2.4. As a result, the pair in fact corresponds to a pair of approximated components of the SSA; $\tilde{\mathbb{F}}^{(m)}_t, \tilde{\mathbb{F}}^{(n)}_t$. If a component has a peak frequency value not equal to any other peak frequency in Ψ_{BC}, then the associated component is excluded from the group in criteria-2 when constructing the common business cycle.

Below is the summary of the steps of the algorithm used in this study which selects components used for the construction of the business cycle:

1. Step-1: Carry out stage-1 of SSA; embedding and SVD
2. Step-2: Consider complete separability for grouping such that $I_i = \{i\}$; $\boldsymbol{X}_{I_1} = \boldsymbol{X}_1$, $\boldsymbol{X}_{I_2} = \boldsymbol{X}_2, ...$ and so on. Set $l = min\{L, K, r_x\} \leq pL$
3. Step-3: Carry out stage-2 of SSA; diagonalise and reconstruct pL number of components so that we compute $\tilde{\mathbb{F}}^{(j,s)}_t$ components, where $j = 1, 2, ..., pL$ and $s = 1, 2, ..., p$.[16]
4. Step-4: Identify peak frequencies by periodogram calculations such that they all fall within the business cycle bandwidth such that: $\frac{2\pi}{96} \leq \lambda^*_z \leq \frac{2\pi}{18}$. z is a subset of pL and includes indices of components with peak frequencies within the limits. $\{z_1, ..., z_l\}$ forms the sequence Ψ_{BC}. [17]
5. Step-5: Match peak frequencies in Ψ_{BC} and eliminate any single valued frequencies.
6. Step-6: Once peak frequencies (λ^*_z's) are paired, collect the corresponding indices to form the set of indices z^*'s contained in a new set; Ψ^*_{BC}, which is in turn a subset of pL.
7. Step-7: Aggregate components associated with indices in z^* and repeat steps $2-3$. This produces p number of cycle indicators which are linear combinations of sine/cosine pairs of the form: $\tilde{\mathbb{F}}^{(C,s)} = \sum_{i \in z^*} a_i \sin \lambda_i t + b_i \cos \lambda_i t$
8. Step-8: Select the $\tilde{\mathbb{F}}^{(C,s)}$ that could be aligned closely to the alternative business cycle indicator.

[16] pL is the upper limit of the number of decomposed components. In general application the value of l is set to one of the input parameters.

[17] This step is only repeated once since we are dealing with vertically stacked block matrices.

2.5 Simulated Data and Identification of a common cycle

Before applying the algorithm to a multivariate set of macroeconomic indicators, we would like to first carry out a simulated experiment and see whether a common cyclical component can be singled out for a bi-variate set. Henceforth, in this section we apply the algorithm set out in section 2.4.3 to simulated data in the form of equation 2.4. In this case, we have two time series data with number of observations for each being $n = 1000$. Both simulated series are of equal length and share a cyclical component within the business cycle frequency domain. It should be noted that the SVD step does not differentiate between common and idiosyncratic cycles of similar frequencies. As a result, the common component is integrated into one of the main groups that determine the evolution of both the time series; the trend, cycle or irregular group. The system that represents the two series can be put in perspective following the framework of the equation system of 2.4 as follows

$$\begin{aligned} y_{1t} &= \mu_{1t} + \epsilon_{1t} + x_{1t} + g_t \\ y_{2t} &= \mu_{2t} + \epsilon_{2t} + x_{2t} \\ \boldsymbol{\mu}_{t+1} &= \boldsymbol{\mu}_t + \boldsymbol{\eta}_t + \boldsymbol{e}_t \\ \boldsymbol{\eta}_{t+1} &= \boldsymbol{\eta}_t + \boldsymbol{\xi}_t \\ x_{t+1} &= \theta_1 x_{t-1} + \theta_2 x_{t-2} + u_t \\ g_{t+1} &= \theta_1 g_{t-1} + \theta_2 g_{t-2} + v_t \end{aligned} \tag{2.17}$$

where $\boldsymbol{\mu}_t = (\mu_{1t} \;\; \mu_{2t})'$, $\boldsymbol{\eta}_t = (\eta_{1t} \;\; \eta_{2t})'$, $\boldsymbol{e}_t = (e_{1t} \;\; e_{2t})' \sim N(\mathbf{0}_{2\times 1}, diag(0,0))$, $\boldsymbol{\xi}_t = (\xi_{1t} \;\; \xi_{2t})' \sim N(\mathbf{0}_{2\times 1}, diag(\sigma_1^2, \sigma_2^2))$, $x_{1t} = x_{t+0}$, $x_{2t} = x_{t-5}$, u_t and v_t are Gaussian error terms and the process x_t and g_t are both generated by an AR(2) with coefficients selected such that the AR(2) polynomial: $1 - \theta_1\varrho - \theta_2\varrho^2 = 0$ has complex roots $\varrho_1, \varrho_2 = -\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}$.[18] These roots $\varrho_1 = -\theta_1 + i$ and $\varrho_2 = -\theta_2 - i$ are complex conjugates of each other and have the straight forward representation of $\varrho_1 = \tau exp\{i\lambda_1\}$, $\varrho_2 = \tau exp\{-i\lambda_2\}$ where τ is the modulus and λ is the frequency. These expressions are known to have trigonometric forms such that; $exp\{i\lambda_1\} = \cos\lambda_1 + i\sin\lambda_1$ and $exp\{i\lambda_2\} = \cos\lambda_2 - i\sin\lambda_2$. Once it is established that both x_t and h_t are cycles in the equation system of 2.17, it can further be stated that x_t is the common cycle as it is present in both time series while g_t is the idiosyncratic cycle related to the first series y_{1t}. An alternative definition of x_t and g_t is that they are linear combinations of sine/cosine waves and can be traced back to the system of equation 2.4. In this framework, instead of treating x_{1t}, x_{2t} as lagged versions of x_t; they can be treated as phase shifted versions of a linear combination of sinusoids so that

[18] If $\theta_2 < 0$ and $\theta_1^2 < 4\phi_2$, the roots of the AR(2) polynomial are complex conjugates.

$$
\begin{aligned}
x_{1t} &= \sum_{i=1}^{m} \left[a_i^{(1)} \cos(\lambda_i t + \phi_1) + b_i^{(1)} \sin(\lambda_i t + \phi_1) \right] \\
x_{2t} &= \sum_{i=1}^{m} \left[a_i^{(2)} \cos(\lambda_i t + \phi_2) + b_i^{(2)} \sin(\lambda_i t + \phi_2) \right]
\end{aligned}
\tag{2.18}
$$

Considering the trigonometric definitions in equations 2.6, the mapping of coefficients $a_i^{(1)}$, $b_i^{(1)}$, $a_i^{(2)}$, $b_i^{(2)}$ to $R_{1i}sin(\lambda_i t + \phi_{1i})$, $R_{1i}cos(\lambda_i t + \phi_{1i})$, $R_{2i}sin(\lambda_i t + \phi_{2i})$, $R_{2i}cos(\lambda_i t + \phi_{2i})$ ensures the difference or equality in amplitude and phase among cyclical components. Since x_{1t} and x_{2t} are phase shifted versions of the same cycle (or the lagged version of an AR(2) process with complex roots), we can assume the amplitudes across linear combinations are identical so that; $R_{1i} = R_{2i}$. On the other hand, the phase between x_1 and x_2 are considered to be different so that: $\phi_1 \neq \phi_2$. However, this phase difference would be equal across j and represents the lag difference between $x_{1t}(x_{t+0})$ and $x_{2t}(x_{t-5})$. Depending on how different or similar the features between any two cycles are, these can be reflected by evaluation of coefficients. If all features are different, it can simply be treated that the two cycles are idiosyncratic. However in such a case, our study will be trivial. This is because a multivariate construct will not lead to any information gain as there is no commonality between series and that under these circumstances analysing each series individually under a univariate framework will be sufficient.

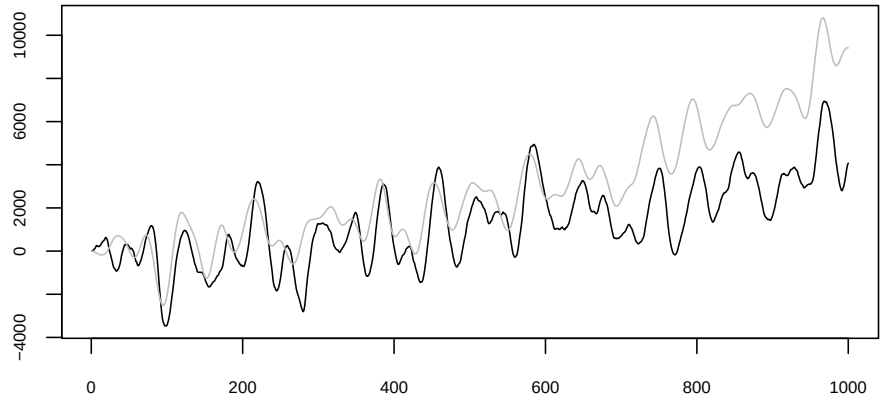
The integrated random walks μ_{1t} and μ_{2t} in equation 2.17 are generated by running random draws from ξ_1, $\xi_2 \sim N(0, 0.000001)$. Coefficients of the common AR(2) process are set to $\theta_1 = 1.88$, $\theta_2 = -0.90$ so that the roots ϱ_1, ϱ_2 are complex values: $-1.88 \mp i$. Errors of the AR(2) processes are $u, v \sim N(0, 1)$.[19] Figure 2.2 shows both series have the stochastic trends leading to an upward climb, with the uptick being prominent towards the end of the simulated samples. Additionally, there is an evident cyclical oscillation which seems to overlap in both series and underlies the common cyclical component that we aim to identify.

The identification of the common cyclical component in both series requires firstly the SVD decomposition procedures. However, as we have set the model accordingly, the common component should be found nested within the cyclical component and not in the trend or irregular components since the later two were generated independently. Furthermore, they lie outside the frequency spectrum and therefore will be excluded by our algorithm. Despite this fact, a common component may still be present in the trend component forming a vector $\mathbf{\Pi} = (1, -\pi)$ such that $\mathbf{\Pi}\boldsymbol{\mu}$ is a stationary process. However, these components by definition will not effect our results as they lie below the pre-determined frequency limit.

Once the cyclical components are extracted through the SSA steps, we move to step-4 which will isolate the cyclical components which will be considered to form the business cycle according to peak frequency values. Table 2.1 summarizes the results of this procedure. The first block of index sequences are the subscripts associated to the components $\widetilde{\mathbb{F}}_t^{(j,s)}$ (output of step-2) all of which have peak frequencies lying

[19] AR(2) process are separately achieved by running a linear filter on $N(0, 1)$ processes with parameters θ_1, θ_2.

Figure 2.2: Simulated bi-variate series with $\sigma_{\epsilon_1}, \sigma_{\epsilon_2} = 0.001$, $\theta_1 = 1.88$, $\theta_2 = -0.90$, $\sigma_u = 1$, $\sigma_v = 1$, $n = 1000$



Note: Data generated by the processes y_{1t} ——— and y_{2t} ——— (see equation 2.17). In SSA terminology these series refer to data inputs; $\mathbb{F}_t^{(1)}$ and $\mathbb{F}_t^{(2)}$ both of which are composed of individual trend, cycle and high frequency (noise) terms.

within the business cycle spectrum. Hence, these numbers can be directly traced to the subscripts for the components $\boldsymbol{X}_1$,...,$\boldsymbol{X}_l$ which are the decomposed matrices of the SVD-step-1 or identically the indices of $\boldsymbol{X}_{I_1}$,...,$\boldsymbol{X}_{I_l}$ at the initial grouping stage-step-2.[20] Results indicate that the cyclical components that fall within the frequency boundaries of the SSA filter, are mostly identified in the first 60 components. Hence, they constitute as the main determinants of the time series processes despite the fact that a visible trend is evident. As mentioned before, the indices z_i's form the initial set component indices of Ψ_{BC} which identify the components to be included in the business cycle indicator. However, the process includes further screening of the components according to whether the peak frequencies can be paired and eliminating odd pairs along the step. Hence, a final round of elimination is conducted by executing step-5 and 6 which its results are presented in block **II**. [21] The second block of index sequences are ordered by even pairs and by values of the corresponding peak frequencies ($\lambda^*_{z^*}$ with indices z^* in this sequence forming Ψ^*_{BC} and range between 85.3 and 18 time units in contrast to the predetermined period filter range of $\{96, 18\}$.

Figure 2.3 provides a visualization on how the evenly paired components found in Table 2.1 are factored into the bi-variate system according to the peak frequency ordering. The amount of deviation from a hypothetical bi-variate system of two identical series can be compared by the indicated middle line. Each column represents unit weight distributed between the estimated factor weights $\hat{a}^{(1)}$, $\hat{a}^{(2)}$ and between the preceding estimated factor weights $\hat{b}^{(1)}$, $\hat{b}^{(2)}$ corresponding to coefficients in equation 2.18. The factor weights with corresponding algorithm based subscripts are set to $\hat{a}^{(1)}_{z_i^*} = \sum_{j=1}^{L/2} u^2_{j,z_i^*}$, $\hat{a}^{(2)}_{z_i^*} = \sum_{j=\frac{L}{2}+1}^{L} u^2_{j,z_i^*}$, $\hat{b}^{(1)}_{z^*_{h_i}} = \sum_{j=1}^{L/2} u^2_{j,z^*_{h_i}}$,

[20] Complete separability at step-2; ($I_i = \{i\}$), ensures at the later stages of the algorithm, that peak frequencies are evaluated individually; $\lambda^*_{I_i} = \lambda^*_i$.

[21] If odd number of λ^*_i's have equal value, then the associated components σ_i's are evaluated with the component having the lowest one eliminated to form an even pair.

Table 2.1: Grouping components using peak frequencies; steps 4 to 6

I. Indices in the sequence $\{z\} = \{z_i = min(z), ..., z_j = max(z)\}$

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 29, 32, 33, 34, 35, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 55, 56, 58, 59, 60, 61, 66, 67, 70, 71, 78, 81, 86, 126

II. Component index sequence with corresponding peak frequencies $\lambda^*_{z^*}$

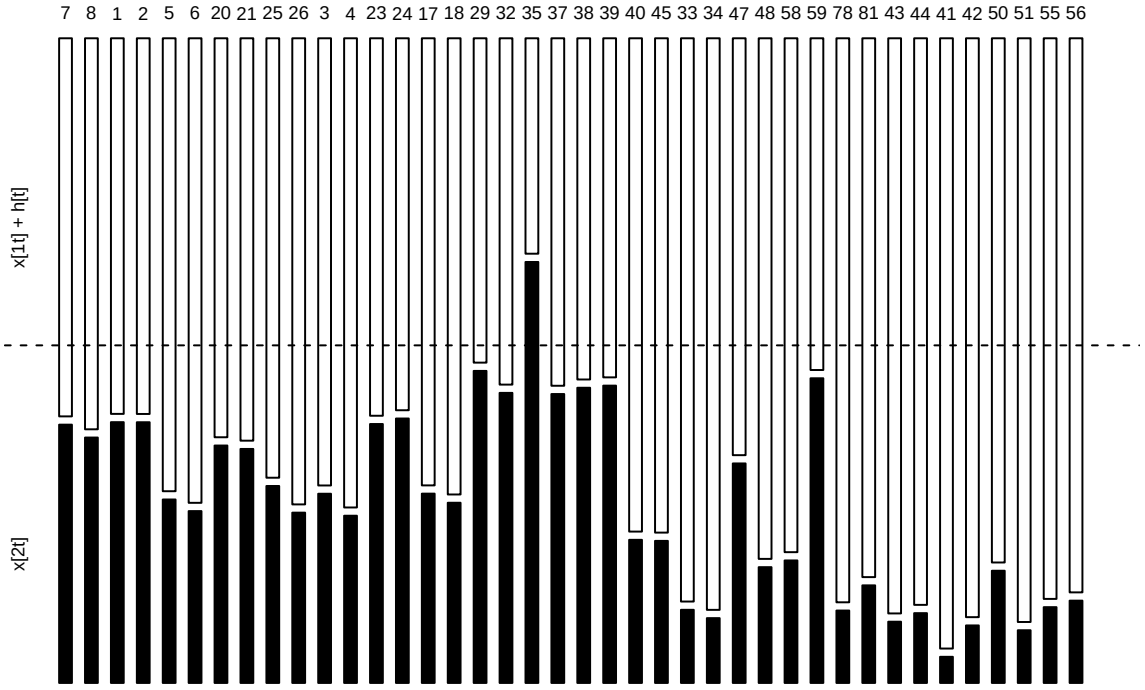
$\{z^*\}$	$\lambda^*_{z^*}$	$\{z^*\}$	$\lambda^*_{z^*}$	$\{z^*\}$	$\lambda^*_{z^*}$	$\{z^*\}$	$\lambda^*_{z^*}$
7	0.0736	3	0.1595	38	0.2209	78	0.2945
8	0.0736	4	0.1595	39	0.2209	81	0.2945
1	0.0859	23	0.1718	40	0.2332	43	0.3191
2	0.0859	24	0.1718	45	0.2332	44	0.3191
5	0.1104	17	0.1963	33	0.2454	41	0.3436
6	0.1104	18	0.1963	34	0.2454	42	0.3436
20	0.1227	29	0.2086	47	0.2577	50	0.3559
21	0.1227	32	0.2086	48	0.2577	51	0.3559
25	0.1350	35	0.2086	58	0.2700	55	0.3559
26	0.1350	37	0.2086	59	0.2700	56	0.3559

Note: Second columns is the frequency in radian degrees. To get the number of oscillations per unit of time, calculate: $\lambda_j/2\pi$.

$\hat{b}^{(2)}_{z^*_{h_i}} = \sum_{j=\frac{L}{2}+1}^{L} u^2_{j,z^*_{h_i}}$ respectively and where the paired indices $\{z^*_i, z^*_{h_i}\} \in \Psi^*_{BC}$, $z^*_i < z^*_{h_i}$ such that $z^*_i \cap z^*_{h_i} = \emptyset$ are disjoint index sequences and $\hat{a}^{(1)}_{z^*_i} + \hat{a}^{(2)}_{z^*_i} = 1$, $\hat{b}^{(1)}_{z^*_{h_i}} + \hat{b}^{(2)}_{z^*_{h_i}} = 1$. Looking to the results, the sine/cosine pairs associated to the components 33, 34, 43, 44 41, 42, 50, 51, 55, 56 constitute features that are aligned heavily towards y_{1t}. The cycle component 41 exhibits the highest y_{1t} related features and records most of the unit weight (0.97) and is well beyond the 0.5 mark. In turn, the component's estimated weight for y_{2t} is left at only 0.03. The only weight estimate exceeding the half-line mark for y_{2t} is associated to component 35, which accounts for two thirds of the unit weight at 0.66. Nevertheless, factors for these components are present in both series and no single weight estimate has a completely zero value for any of the components in Ψ^*_{BC} which confirms the commonality. The reason why components with unbalanced allocation of weights exist is that, components defining g_t (the idiosyncratic cycle part of y_{1t} which we had nested) have cyclical properties in the frequency spectrum that are very much inline with the components forming the common cycle x_t. As such, the data matrix still passes theses components to the common cycle matrix ($\boldsymbol{X}_{\boldsymbol{\Psi}^*_{BC}}$) albeit designating asymmetric weights through the elements in $\boldsymbol{U}^{(1)}_{\Psi^*_{BC}}$ and $\boldsymbol{U}^{(2)}_{\Psi^*_{BC}}$. Because g_t is known to be idiosyncratic cycle part of y_{1t} and not related to y_{2t}, we are able to identify beforehand that high factor loads relate to y_{1t}. However, it could be the case that g_t may be a very similar cycle to x_t and hence reinforce the commonality in y_{1t}. Another case could be that the common component has a stronger presence in y_{1t} series being introduced in the form of $\beta_1 x_{2t}$ with $\beta_1 > 1$ while for the y_{2t} series, being introduced into the system by $\beta_2 x_{2t}$ with $0 < \beta_2 < 1$. The analysis on how factor

loading allocations change with respect to commonality structure of individual series is beyond the scope of the study. Therefore, we only take results from this part of the analysis mainly to highlight the significance of the $\boldsymbol{X}$ and its SVD elements in $\boldsymbol{U}$, matrix which acts as a weight distributing schema of the selected components. These weights will inevitably dictate the behavior of each filtered estimate that the SSA produces.

Figure 2.3: Estimated factor loadings across common component matrix ($\boldsymbol{X}_{\Psi^*_{BC}}$)

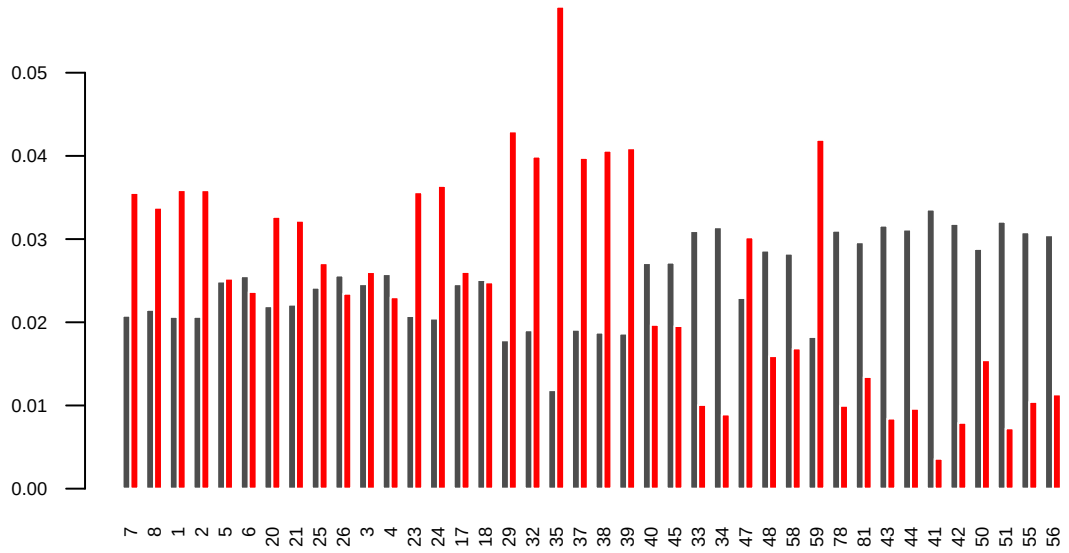


Note: Component loadings calculated from $\boldsymbol{X}^{(1)}_{\Psi^*_{BC}}$ related to coefficient estimates $\hat{a}^1_{z^*_i}$, $\hat{b}^1_{z^*_{h_i}}$ represented by □. Loadings of $\boldsymbol{X}^{(2)}_{\Psi^*_{BC}}$ related to coefficients estimates $\hat{a}^2_{z^*_i}$, $\hat{b}^2_{z^*_{h_i}}$ are represented by ■. See equation 2.4 for the representation and equation 2.6 for trigonometric identities.

To be able to understand how much these components are passed through into the evolution of the filter estimates, a check on the relative weights can provide some insight. Furthermore, it could give information on which of the filtered estimate from the common stacked matrices $X^{(1)}_{\Psi^*_{BC}}$ and $X^{(2)}_{\Psi^*_{BC}}$ provide a better fit to the common cycle. It should be mentioned that this is much easier done in a two series system, and it may result not as intuitive in case we deal with a larger multivariate system. In this regard, relative factor weights of the common cycle components for y_{1t} are each computed by $\bar{a}^1_{z^*_i} = \hat{a}^{(1)}_{z^*_i} / \sum_{z^*_i \in \Psi^*_{BC}} \hat{a}^{(1)}_{z^*_i}$, $\bar{b}^1_{z^*_{h_i}} = \hat{b}^{(1)}_{z^*_{h_i}} / \sum_{z^*_{h_i} \in \Psi^*_{BC}} \hat{b}^{(1)}_{z^*_{h_i}}$. Similarly, for the common cycle components, the relative weights are calculated by replacing the factor weights with $\hat{a}^{(2)}_{z^*_i}$, $\hat{b}^{(2)}_{z^*_{h_i}}$. The significance of the relative weights is that they will accommodate the reader on which filtered estimate reflects features that are more inline with common component cycles in the set of Ψ^*_{BC}. As the identification of which component reflected a higher degree of commonality was retrieved from the estimated factor weights in Figure 2.4, the relative weights will give an indication on which of the two filter estimates; $\tilde{\mathbb{F}}^{(1,C)}_t$, $\tilde{\mathbb{F}}^{(2,C)}_t$ would better reflect the underlying common cycle inherent in both time series; y_{1t}, y_{2t}. However, it should be highlighted once again that this is conditional on the fact that g_t is only present in y_{1t} or g_t

is not correlated with the common component x_t. Additionally we assume that x_t are equally present in y_{1t} andy_{2t} having coefficients $\beta_1 = \beta_2 = 1$. Otherwise, the filter estimate computed from $\boldsymbol{X}^{(1)}_{\Psi^*_{BC}}$ may have been a better fit as the weights will indicate a stronger presence of the common components in the filtered estimate $\tilde{\mathbb{F}}^{(1,C)}_t$ compared with $\tilde{\mathbb{F}}^{(2,C)}_t$. Therefore, the relative weights would play a vital role on the performance of the estimates capturing the underlying common structure. A main result from the figure would be that for the y_{1t} series, relative factor weights seem to be fairly consistent through out the block matrix $\boldsymbol{U}^{(1)}_{\Psi^*_{BC}}$. Moreover, for the relative factor weights; $\hat{a}^{(1)}_{z^*_i}$, $\hat{b}^{(1)}_{z^*_{h_i}}$, computed from $\boldsymbol{U}^{(1)}_{\Psi^*_{BC}}$, components that reflect idiosyncratic features have a slightly higher relative contribution compared to the more common component. This finding will reflect on to the estimate $\tilde{\mathbb{F}}^{(1,C)}_t$ such that similar to noise contamination problem for the signal, the tracking power of the estimate will be disrupted by the inclusion of relatively high factor loadings of components such as 41, 42, 51. For the relative factor loadings $\hat{a}^{(2)}_{z^*_i}$, $\hat{b}^{(2)}_{z^*_{h_i}}$ computed from $\boldsymbol{U}_{\boldsymbol{\Psi}^*_{BC}}$ on the other hand, calculated values vary significantly across the component set. What is more revealing is that, for components with factor loadings distributed highly uneven in Figure 2.3, have very low relative factor weights while components that showed large commonality (factor weights that had close to 0.5 loading distribution for both series), have the highest relative weights across the spectrum of components. For the multivariate SSA, this pattern of relative factor loading weights can be seen as an automatic correction between the individual series in the data matrix $\boldsymbol{X}^{(2)}$. What becomes a relatively high weighted component for one series, automatically results in a relatively low weighted component for the other.

Figure 2.4: Relative factor loadings across block common cycle matrices ($\boldsymbol{X}^{(1)}_{\Psi^*_{BC}}$, $\boldsymbol{X}^{(2)}_{\Psi^*_{BC}}$)



Note: Relative weights; $\bar{a}^{(1)}_{z^*_i}$, $\bar{b}^{(1)}_{z^*_{h_i}}$ are represented by ■. The relative weights; $\bar{a}^{(2)}_{z^*_i}$, $\bar{b}^{(2)}_{z^*_{h_i}}$ are represented by ■.

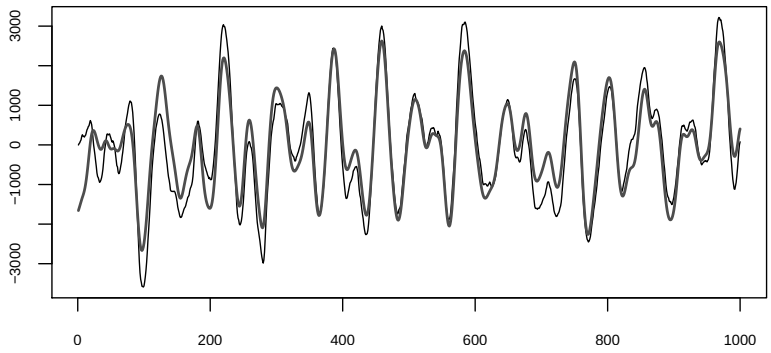
We have now all the required computations and the intuition to be able to execute the final reconstruction step of the SSA algorithm. Once the common component matrix XC is Haneklized and reconstructed according to the SSA steps outlined

in equations from 2.14 to 2.16, two separate cycle filters are produced; $\tilde{\mathbb{F}}_t^{(1,C)}$ and $\tilde{\mathbb{F}}_t^{(2,C)}$. Figure 2.5 depicts the results of both estimates with respect to the original cycles nested in the bi-variate equation system of 2.17. Because this is synthetic data, we are able to exploit the fact that we have built the data set prior to the experiment, therefore making it possible to separate the original cycle components of interest with their filtered estimates. Under this framework, the panel (a), we assume the full cycle component composed from both the common and idiosyncratic parts can be separated from the trend. The difference between the cycle estimate and the full cycle of y1t is fairly narrow through out the simulations. The SSA filter depicts the oscillations very accurately and identifies every turning point. Hence the clear disassociation of trend from cycle can be confirmed. However, the result should come with caution as $\tilde{\mathbb{F}}_t^{(2,C)}$ was designed to be the estimate of the common component x_{1t} while minimizing the influence of the idiosyncratic cycle ht on the full cycle in y_{1t}.

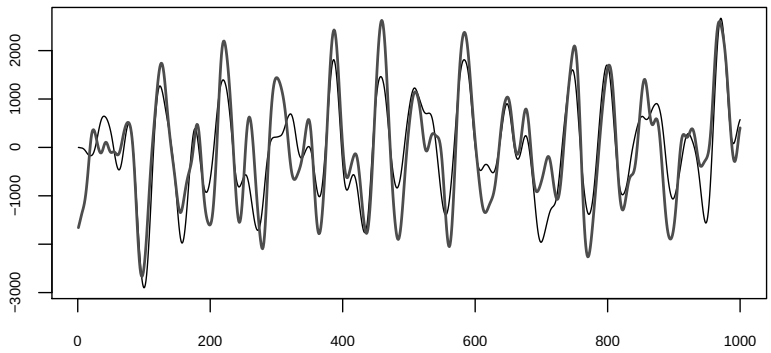
Therefore, we further progress our controlled exercise by extracting x_{1t} from h_t and see how $\tilde{\mathbb{F}}_t^{(1,C)}$ fairs in terms of representing the common cycle. It is evident at first sight that the overall estimation performance of the common cycle has declined with respect to what was presented in panel (a). This can be immediately linked to the influential presence of idiosyncratic cycles which in turn can be traced to the factor loading schema and relative weight across cycle components depicted in figures 2.3, 2.4. It would be misleading to criticize the SSA filter estimation when the presence of idiosyncratic cycle hampers $\tilde{\mathbb{F}}_t^{(1,C)}$ ability in identifying the common cycle. In reality, only two turning points are skipped by the estimate from the 39 turning points while there is considerable overlap of the the original cycle and the estimate. Moving on to the second filter estimate retrieved from y_{2t}, we are able to reach even more satisfying results. Panel (c) display the plot of $\tilde{\mathbb{F}}_t^{(2,C)}$ against the coinciding common cycle (at phase zero: $\phi_1 = 0$) and its shifted version ($x_{2t} = x_{t+5}$). The filter estimate obtained from the second time series not only demonstrates a highly close match to the oscillating common cycle patterns but furthermore it slightly leads the coinciding series at certain episodes of the simulation. In this respect, the filtered estimate can also be considered as a possible predictor of the common cycle. The fact that y_{2t} did not contain an idiosyncratic component made it clear-cut process to extract the estimate of the common component. Moreover, the SSA mechanism acknowledged y_{1t} related cycles by diminishing their relative importance when computing the evolution process of the estimate. The implication of this has given the access to produce a leading indicator of a common component for a multivariate series. This algorithmic framework can be of applied conveniently to a large group of macroeconomic indicators in order to come up with leading/coinciding estimate of a business cycle indicator that is in commercial use. The next section deals with this application for Euro-zone data to be able design a business cycle indicator for the EU region.

Figure 2.5: Fitting filter estimates to common and idiosyncratic cycles in a bi-variate SSA system

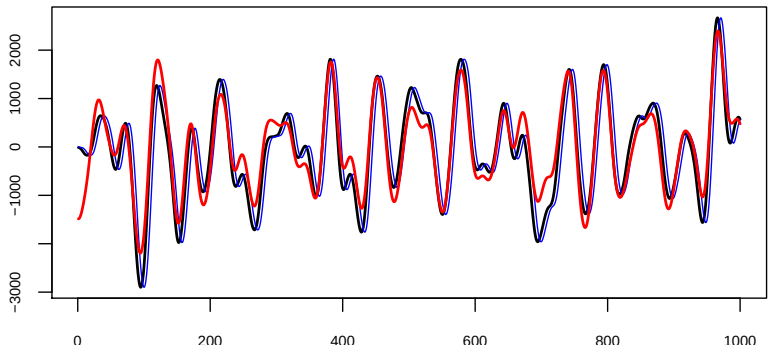
(a) $\mathbb{F}_t^{(1,c)}$ vs $\tilde{\mathbb{F}}_t^{(1,C)}$



(b) x_{1t}: $\left(\mathbb{F}_t^{(1,c_x)}\right)$ vs $\tilde{\mathbb{F}}_t^{(1,C)}$



(c) $\mathbb{F}_t^{(1,c_x)}$, $\mathbb{F}_t^{(2,c_x)}$ vs $\tilde{\mathbb{F}}_t^{(2,C)}$



Note: For panels (a) and (b): the grey line indicated by; — $\tilde{\mathbb{F}}_t^{(1,C)}$, is the SSA based business cycle estimate for $\mathbb{F}_t^{(1)}$. Black line in panel (a) indicates: — $\mathbb{F}_t^{(1,c)}$ which is the aggregated common and idiosyncratic cycles; $x_{1t} + g_t$. Black line in panel (b): indicates $\mathbb{F}_t^{(1,c_0)}$ is x_{1t} which is the common cycle for y1t. For panel (c): the red line indicated by; — $\tilde{\mathbb{F}}_t^{(2,C)}$, is the SSA based business cycle estimate for $\mathbb{F}_t^{(2)}$, black line represent: — $\tilde{\mathbb{F}}_t^{(1,c_0)}$, the blue line represents —$\mathbb{F}_t^{(2,c_0)}$ is the shifted common cycle. See equation 2.17 for the explicit forms.

2.6 Data and Empirical Application

Our study includes a multivariate data set for the EU region. We follow the work of Valle e Azevedo et al. (2006) in selecting the data set for our application. The data set includes the sentiment indicators; Construction Confidence (CNI), Industrial Confidence (ICI), Retail Confidence (RCI), Consumer Confidence (CCI), the unemployment rate (UNM), the interest rate spread (INT: 10-year maturity government bond yields minus 3-month Euro interbank offered rate), Industrial Production Index (IPI). We exclude Retail Sales Volume Index as it does not include data prior to the initial recording of the EuroCOIN indicator (01/1999).[22,23]

The EuroCOIN indicator is used to track the business cycle for the overall EU economy and identify its turning points. It is updated at monthly frequencies announced by the Center for Economic Policy Research (CEPR) website (www.cepr.org). Works of Altissimo et al. (2000) and Altissimo et al. (2010) show the essence of the indicator is designed. These papers rely on a large data set containing hundreds of macroeconomic series for major EU economies to construct an indicator that can estimate the medium terms component of GDP growth of the Euro region. The following sections will present the empirical results from the SSA methodology previously outlined and will conclude by comparing results with the EuroCOIN indicator series released by CEPR and the QoQ growth of Euro region GDP.

All our series are of equal length containing $N = 254$ monthly observations that spans the period starting from 01/1999 until 02/2020. We compare our constructed indicator from this multivariate set to the EuroCOIN indicator and the EU region GDP QoQ growth series with starting dates of 01/1999 eventhough there is data for constructed the EU cycle starting at an earlier date.[24] All our data have been standardized and have been adjusted to seasonality and calendar periods.

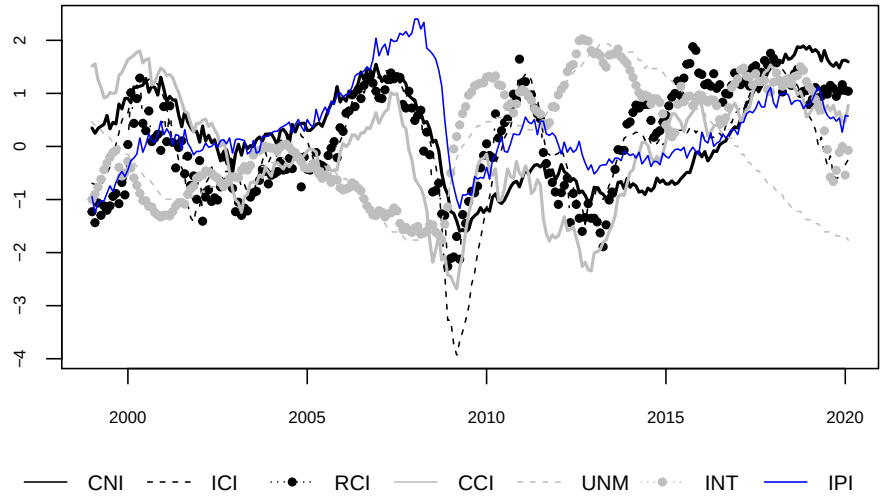
The macroeconomic variables in figure 2.6 contain cyclical components that should strongly relate to the business cycle of the related economic region. Whether their cyclical components coincide/lead or lag the business cycle and if they are pro or counter cyclical depends on the macroeconomic fundamentals they reflect. The IPI is the most commonly known variable used in empirical studies. It is an indicator usually considered to roughly coincides the business cycle. The confidence indicators of; Construction, Industrial, Retail, Consumer confidence indicators are in large forward looking and hence should be leading the business cycle. However, there are certain differences on these confidence indicators that should be made clear at the start of our analysis. Economic cycles are known to be strongly correlated with developments in the durable goods market. ICI and RCI closely follow developments in industrial and retail products which a large fraction is related to durable

[22]Missing data could have been retrieved by aggregating country level Retail Sales Volume and chained to where data period starts, however we exclude the series to not further complicate data collection process.

[23]GDP data is of quarterly frequency and therefore excluded in the monthly frequency data set as it will contain missing data.

[24]Data source: https://ec.europa.eu/eurostat/databrowser. Secondary source: https://eurocoin.cepr.org/.

Figure 2.6: Macroeconomic Indicators for the construction of an EU region business cycle Index, (01/1999-02/2020)



Note: Each individual time series is standardized so that $\tilde{z}_t^{(j)} = \frac{f_t^{(j)} - \tilde{\mu}_{f^{(j)}}}{\tilde{\sigma}_{f^{(j)}}}$. The original data is seasonally/calendar adjusted with monthly frequencies. Data source: https://ec.europa.eu/eurostat/.

goods markets. Therefore it is plausible to say both these indicators are sensitive to future movements in the business cycle. On the other hand, CNI and CCI are expected to be less sensitive. CNI would be more correlated to cyclical patterns in housing and mortgage markets and CCI are closely linked with with non-durable goods market. Nevertheless, construction activity and consumer expenditures are important driving forces of economic growth. Hence they reflect vital information for business cycle analysis. INT is another leading indicator of the business cycle. It should be taken in to account that INT is a spread formed by the gap between short and long term interest rates both of which are governed by different fundamentals. Monetary authorities have stronger influence on the former while for the later, investor/market sentiment on economic outlook and inflation expectations are influential. Combining all these forces, the narrowing (below a certain positive level) of the spread or a negative value would indicate an expectation of recessionary buildup in the economy and a widening (above a certain positive level) would an outlook of rigorous economic activity. UNM is the only lagging indicator in our data and which is also the only counter-cyclical one to the business cycle. This means that the cyclical component extracted should rise (decline) even if the economy is expanding (contracting) since companies will not react immediately to economic conditions and hire (dismiss) workers simultaneously.

The remaining part of section will be devoted to presenting the results and analysis of the SSA based filtering procedures outlined in sections 2.4.1, 2.4.2 and 2.4.3 for the data set which we presented in Figure 2.6. We will evaluate the constructed business cycle indicator based on the SSA filter after which we will compare it with indicators commercially in use and with common cycles filtered from a more structural framework using Unobserved Component Models (UCM).

Given the explanation made in selecting the window length in section 2.4.1, we

start by setting model parameters $L = 120$, $K = N - L$. [25]

For a better understanding of the frequency domain characteristics dominating our data set, we also present the calculated smooth periodograms. After processing the data in Figure 2.6 through Steps $1 - 3$, we can compose the smoothed periodograms for the approximated filtered cycles. In time series practice, the general formula for the periodogram is

$$I^{(s)}(\lambda_j) = \frac{1}{N}\left(\sum_{1}^{N} \tilde{\mathbb{F}}t^{(C,s)} \cos \lambda_j t/N\right)^2 + \frac{1}{N}\left(\sum_{1}^{N} \tilde{\mathbb{F}}t^{(C,s)} \sin \lambda_j t/N\right)^2 \tag{2.19}$$

where $j = 1, 2, ..., (N-1)/2, N/2$, $\lambda_j = 2\pi j/N$. The visualization of periodograms in the form of equation 2.19 is usually noisy. Further, a large value of N will not lead to a smoother periodogram. To smooth the noise structure a moving average can reflect a clearer view of the underlying frequency characteristics of the data.[26] The smoothed (moving averaged) periodogram is then

$$\tilde{I}^{(s)}(\lambda_j) = \sum_{i=-m}^{m} k_i I^{(s)}(\lambda_j + i/N) \tag{2.20}$$

where k_i's are the smoothing weights of $I^{(s)}(\lambda_j)$ and m is the width of the smoothing range: $j - m, ..., j + m$.

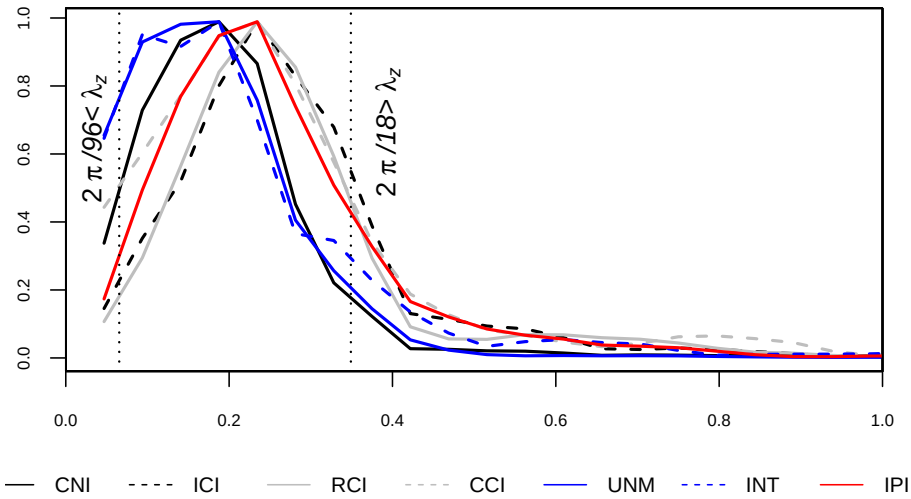
Figure 2.7 presents results of the smoothed periodograms in equation 2.20 associated to the SSA based filters. More specifically, the filters are the residuals once the first two leading components which are usually the trend components with the lowest frequency peaks are extracted from the data. The remaining spectral density is then formed by components along the business cycle frequency spectrum as well as with components that have higher frequencies than $2\pi/18$. The later components however will be constituting a relatively very low contribution to the variation of all the 7 individual series if business cycle features dominate the data. These components with peak frequencies beyond the upper limit of the business cycle can be classified as the irregular part of the individual series. The excluded two leading SSA based components can be identified as trend components both with frequencies at 0.0074 equivalent to 135 months (which is significantly below the lower limit of 0.065 or 96 months). Once these low frequency components are removed it can be clearly seen that the time series can be strongly characterized by components collected in Ψ_{BC} which is the business cycle frequency range.

Table 2.2 shows that amongst 120 components that were selected, only 15 peak frequencies and their respective indices correspond to the components of our interest. All these components lie within the first 18 components, with the exception of the

[25] As noted earlier; L, K, r are not parameters to be estimated but rather pre-determined input values to start the algorithm running.

[26] In empirical work this method is the Daniell kernel, see Shumway et al. (2000) for further details.

Figure 2.7: Smoothed periodograms for reconstructed series $\tilde{\mathbb{F}}_t^{(C,s)}$



Note: Sample periodograms have been smoothed according to equation 2.20 with $m = 5$. Grey dashed lines indicate the frequency spectrum allowed by Ψ_{BC}.

first two leading components being eliminated as a result of showing trend like features. [27] The summary results of Step-5 is also represented in the third column. The corresponding peak frequencies in the 2nd column of table 2.2 illustrates the ordering rule of SSA methodology. Notice that components 3 and 6,7 should be in consecutive order when in fact the ordering is interrupted by the consecutive pair 5,6. This is evidence that the SSA does not sort out components according to frequency levels but rather, the main determinant of ordering is with respect to the singular values. These singular values determine the weight each $\boldsymbol{X}_i$'s have in the original data matrix such that $\sigma_i / \sum_{min}^{max} \sigma_i$. The singular values can be retrieved simply by taking the square root of the eigenvalues found in the main diagonal of $\boldsymbol{Q}$; see equation 2.10). A further finding from columns 2 and 3 is that most components with peak frequencies within the business cycle frequency spectrum come in pairs of 2 or more as only component 15 has no pair in the frequency domain. As a result, following the rule in steps 5, we eliminate λ_{15}^*. We also eliminate λ_7^* and λ_{12}^* as a result of the definition in steps 6-7, because only two components should represent the sine/cosine pair. The reasoning behind eliminating λ_7 from its identical peak frequency components $\{3, 6\}$ and λ_{12} from its identical peak frequency components $\{10, 11\}$ is because both these values have the least significant component (lowest singular values) amongst their equal peak frequency peers; $\sigma_7 < \sigma_3, \sigma_6$ and $\sigma_{12} < \sigma_{10}, \sigma_{11}$. The paired components will be included in the summation of the harmonic parts of equation 2.3 once they are filtered as in equation 2.16 and allow us to construct the 7 business cycle indicators expressed in step-8. We are left with 6 unique sine/cosine pairs which we will use to construct the cycle indicator common to the 7 uni-variate series in our data set.

To complement the findings in Table 2.2, we present the scree-plot of the leading components. We stated that the data matrix is able to identify a strong trending tendency and hence the two highest logs of singular values are eliminated from

[27] Both leading component frequency peaks; λ_1^* and λ_2^* are calculated around 0.046 respectively and fall below the lower peak frequency limit $\lambda_L L$.

Table 2.2: Identification of Peak Frequencies and Corresponding Identities

index values:$\lambda_z^\top$	$2\pi j$	λ_{z^*}'s: peak frequency pairs
λ_3^*	0.0931	$\lambda_3^* = \lambda_6^*$
λ_4^*	0.1396	$\lambda_4^* = \lambda_5^*$
λ_5^*	0.1396	$\lambda_8^* = \lambda_9^*$
λ_6^*	0.0931	$\lambda_{10}^* = \lambda_{11}^*$
λ_7^*	0.0931	$\lambda_{13}^* = \lambda_{14}^*$
λ_8^*	0.1862	$\lambda_{17}^* = \lambda_{18}^*$
λ_9^*	0.1862	
λ_{10}^*	0.2793	
λ_{11}^*	0.2793	
λ_{12}^*	0.2793	
λ_{13}^*	0.3723	
λ_{14}^*	0.3723	
λ_{15}^*	0.2327	
λ_{17}^*	0.4189	
λ_{18}^*	0.4189	

Note: Second columns is the frequency in radian degrees. To get the number of oscillations per unit of time, calculate: $\lambda_j/2\pi$.

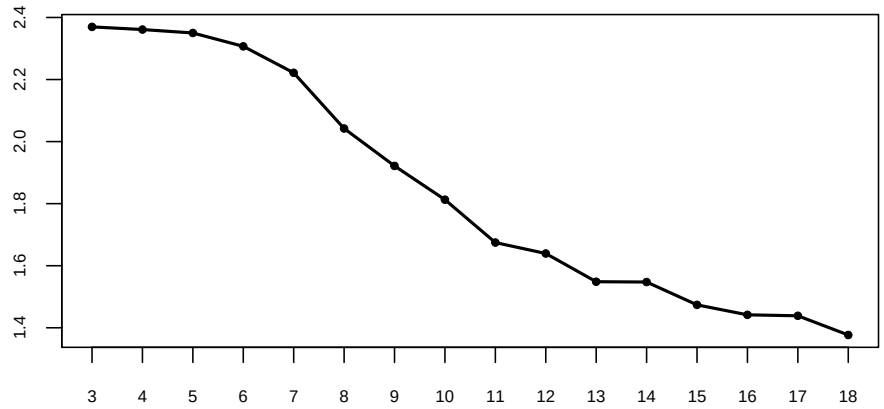
the plot. As previously mentioned, the indexing/ordering of components in the SSA framework is determined solely by the magnitude of σ_i's. In this context, the drawback of relying on visualization of eigen-triplet components may very well be misleading. One may consider the plateauing of the line between segments 3-7 may indicate two consecutive harmonic pairs 3,4 5,6 when our analysis in the frequency domain showed an interruption. Similarly, a sharp decline in the magnitude of the singular value when moving from 6 to 7 can wrongly be interpreted as a break in grouping when truth is the corresponding components have equal frequency peaks. Finally, the almost flat line from 16-17 again may very well be taken falsely as a strong sign of pairing when our previous findings reveal component 16 does not form a pair and 17 forms a pair with a component that has a lower singular-value.

It is important to asses some of our results for a filtering method regarding the extraction of our full component of interest in the context of signal-noise separability. As it was briefly discussed in section 2.4.1, a measure to check for the filter succeeded in separating the signal and error can be assessed by w-correlation of Hassani et al. (2009). The w-correlation between any two reconstructed group in the form of $\tilde{\mathbb{F}}_t^{(i)}$ is computed as

$$\rho_w\left(\tilde{\mathbb{F}}_t^{(i)}, \tilde{\mathbb{F}}_t^{(i')}\right) = \frac{\sum_{t=1}^{N} w_t \tilde{f}_t^{(i)} \tilde{f}_t^{(i')}}{\left(\sum_{t=1}^{N} w_t \tilde{f}_t^{(i)} \tilde{f}_t^{(i')}\right)\left(\sum_{t=1}^{N} w_t \tilde{f}_t^{(j)} \tilde{f}_t^{(i')}\right)} \tag{2.21}$$

where $i = i'$, $2 \leq i, i' \leq L$, $\tilde{f}_t^{(i)}, \tilde{f}_t^{(i')}$ are t^{th} observation estimates from reconstructed series $\tilde{\mathbb{F}}_t^{(i)}, \tilde{\mathbb{F}}_t^{(i')}$, w_t is an integer value which counts the number of times the t^{th} observations are repeated along the columns of reconstructed matrices

Figure 2.8: Contribution of elementary components in the frequency spectrum domain of Ψ_{BC}.



Note: The logarithms of singular values $\log(\sigma_1), ..., \log(\sigma_K)$ are plotted against their corresponding component indices $\{1, ..., K\}$ at an orderly fashion.

$\tilde{\boldsymbol{X}}_{I_i}, \tilde{\boldsymbol{X}}_{I_{i'}}$.[28]

Prior to moving on to results, we briefly discuss the reasoning behind why we also include the trend component in our separability analysis. While we do have a clear definition on the signal has been made in this study, no discussion on the error component has been put forward so far. The error component in SSA terminology does not necessarily require to be a Gaussian random variable in the classical sense. More over, for the filtering purpose, components with high frequencies that do not contribute to the information on the business cycle are referred as the noise. The $\boldsymbol{X}_i$ components converted to $\tilde{\mathbb{F}}_t^{(j,s)}$ time series dimensions that have peak frequencies above the cut off mark of λ_{UL} will be stacked into the error group, $I_{\boldsymbol{\epsilon}}$ to form matrix $\boldsymbol{X}_{I_{\boldsymbol{\epsilon}}}$. Each component within this group will have oscillating patterns that are beyond what can be considered as business cycle oscillations but rather will be classified as irregular patterns. The SSA based filter algorithm can separate the irregular components for p number of time series. The irregular components estimated as $\tilde{\mathbb{F}}_t^{(j,\epsilon)}$ represent error terms stated in the system of equations in 2.4 and 2.17. Eliminating the errors will facilitate to achieve smoothed common cycle estimates $\tilde{\mathbb{F}}_t^{(j,\Psi^*_{BC})}$. Similarly, the trend component does not contain any information content on the business cycle component, yet we can not treat the trend as an error as it is part of the signal. The two leading indicators which form the trend are components with the largest weights contributing to the evolution of the original time series. However, we are not interested on these terms and need to asses how well the filter is able to separate the cycle part of the signal from the trend part of the signal in our study. Table 2.3 provides findings there is a clear signal noise separation indicated by the negligible values for $\rho_w(I_1, \boldsymbol{\epsilon})$, $\rho_w(\Psi^*_{BC}, \boldsymbol{\epsilon})$. It is understandable to find a correlation value between two main parts of the signal ($\tilde{\mathbb{E}}(\mathbb{F}_{\approx})$=trend and cycle) to be considerably high at 0.18. This could most probably be attributed to the trend being of quadratic form or a polynomial function of higher order. These functional forms will be best representing the slow moving oscillations of the trend

[28]The more the observation lies in the mid sample range the higher the value of w_t will be. As the t^{th} moves towards the the initial or end sample range, w_t will decline.

component of the data.

Table 2.3: Filter based $\boldsymbol{w}$-Correlation matrix of reconstructed groups $\tilde{\mathbb{F}}_t^{(s)}$ components using grouped matrices $\boldsymbol{X}_{I_s}$, $(L = 120)$

	trend $(\tilde{\mathbb{F}}_t^{(I_1)})$	cycle $(\tilde{\mathbb{F}}_t^{(\Psi^*_{BC})})$	error $(\tilde{\mathbb{F}}_t^{(\Psi_{\boldsymbol{\epsilon}})})$
$\tilde{\mathbb{F}}_t^{(I_1)}$	1.000	0.182	0.0003
$\tilde{\mathbb{F}}_t^{(\Psi^*_{BC})}$		1.000	0.008

Note: The $\boldsymbol{w}$-correlation $(\rho_w(i,j))$ takes the value between $[0,1]$ and measures the orthogonality between any two SVD based matrices elementary; $\boldsymbol{X}_i, \boldsymbol{X}_j$ or grouped $\boldsymbol{X}_{I_y}, \boldsymbol{X}_{I_z}$. The computed ρ_w values serves as a filter based correlation for corresponding time series transformations of SVD based matrices. Trend, cycle and error components are grouped according to index sets; $I_1 = \{1,2\}$, $I_2 = \Psi^*_{BC}$, $I_3 = \Psi_{\boldsymbol{\epsilon}}$.

Table 2.4 gives a further detailed separability assessment related to the sine/cosine pairs that were selected to form the common business cycle with respect to the error term that was previously defined. Looking at the last column, reveals again that separability of the noise has even been achieved at cyclical pairs level. Additionally, correlations of each pair with respect to the error term all lie below the 0.18 correlation of $\{\tilde{\mathbb{F}}_t^{(I_1)},\ \tilde{\mathbb{F}}_t^{(\Psi^*_{BC})}\}$, with the highest level being recorded at 0.06 for $\rho_w\left(\tilde{\mathbb{F}}_t^{(\{17,18\})}, \tilde{\mathbb{F}}_t^{(\Psi_{\boldsymbol{\epsilon}})}\right)$. This is a further confirmation of separability having been achieved for our methodology. Another analysis can also be made regarding with-in correlation results for the sine/cosine pairs. There is a significant correlation between immediate preceding or succeeding pairs and negligible correlation between pairs beyond the first order. This pattern is inline with findings with the two main signal groups of trend and cycle in Table 2.3. A final note for the table findings is that they should note be taken in the view of assessing our grouping strategy in order to modify it with the new information we have. The aim of the filter based correlations would be to solely evaluate the performance of the model in terms of eliminating noise contamination of the filtered signal.

Table 2.4: Filter based $\boldsymbol{w}$-Correlation of paired components and the error term $\tilde{\mathbb{F}}_t^{(\Psi_{\boldsymbol{\epsilon}})}$, $(L = 120)$

	$\tilde{\mathbb{F}}_t^{(\{3,6\})}$	$\tilde{\mathbb{F}}_t^{(\{4,5\})}$	$\tilde{\mathbb{F}}_t^{(\{8,9\})}$	$\tilde{\mathbb{F}}_t^{(\{10,11\})}$	$\tilde{\mathbb{F}}_t^{(\{13,14\})}$	$\tilde{\mathbb{F}}_t^{(\{17,18\})}$	$\tilde{\mathbb{F}}_t^{(\Psi_{\boldsymbol{\epsilon}})}$
$\tilde{\mathbb{F}}_t^{(\{3,6\})}$	1.000	0.644	0.090	0.004	0.003	0.005	0.001
$\tilde{\mathbb{F}}_t^{(\{4,5\})}$		1.000	0.070	0.003	0.002	0.004	0.001
$\tilde{\mathbb{F}}_t^{(\{8,9\})}$			1.000	0.258	0.025	0.039	0.004
$\tilde{\mathbb{F}}_t^{(\{10,11\})}$				1.000	0.036	0.104	0.004
$\tilde{\mathbb{F}}_t^{(\{13,14\})}$					1.000	0.549	0.016
$\tilde{\mathbb{F}}_t^{(\{17,18\})}$						1.000	0.059
$\tilde{\mathbb{F}}_t^{(\Psi_{\boldsymbol{\epsilon}})}$							1.000

Note: For $\tilde{\mathbb{F}}_t^{(\{\cdot\})}$, the superscript $\{\cdot\}$ indicates the pairs in Ψ^*_{BC}. See Table 2.3 for further details.

Table 2.5 presents the signal-to-noise separability assessment at the most primitive signal level. As a result correlation results are for single component time series based on the reconstruction of rank-1 data matrices. However, we only include components with index values in Ψ^*_{BC}. Once more the final column highlights strong

separation of noise from individual components used to construct the business cycle indicator. The highest computed correlation is recorded for $\rho_w\left(\tilde{\mathbb{F}}_t^{(18)}, \tilde{\mathbb{F}}_t^{(\Psi_\epsilon)}\right)$ at 0.11. Nevertheless this value is still below the 0.12 correlation found between $\{\tilde{\mathbb{F}}_t^{(I_1)}, \tilde{\mathbb{F}}_t^{(\Psi^*_{BC})}\}$ and in Table 2.3. An interesting result retrieved from Table 2.5 is that components that form sine/cosine pairs do not translate to a higher correlation for the two paired series. For example, $\tilde{\mathbb{F}}_t^{(3)}$ and $\tilde{\mathbb{F}}_t^{(6)}$ have a correlation of 0.11 despite that they represent the sinusoidal pair $\tilde{\mathbb{F}}_t^{(3,6)}$. This value compares much lower than the correlations $\rho_w\left(\tilde{\mathbb{F}}_t^{(3)}, \tilde{\mathbb{F}}_t^{(4)}\right) = 0.53$ and $\rho_w\left(\tilde{\mathbb{F}}_t^{(5)}, \tilde{\mathbb{F}}_t^{(6)}\right) = 0.82$ even though $\{\tilde{\mathbb{F}}_t^{(3)}, \tilde{\mathbb{F}}_t^{(4)}\}$ and $\tilde{\mathbb{F}}_t^{(5)}, \tilde{\mathbb{F}}_t^{(6)}$ do not form oscillating pairs. The underlying cause for such an observation is straightforward; sine/cosine pairs are only phase shifted versions that translate to low correlation.

Table 2.5: Filter based $\boldsymbol{w}$-Correlation of single component series estimates $\tilde{\mathbb{F}}_t^{(\cdot)}$ and grouped error term $\tilde{\mathbb{F}}_t^{(\Psi_{\boldsymbol{\epsilon}})}$, $(L = 120)$

	$\tilde{\mathbb{F}}_t^{(3)}$	$\tilde{\mathbb{F}}_t^{(4)}$	$\tilde{\mathbb{F}}_t^{(5)}$	$\tilde{\mathbb{F}}_t^{(6)}$	$\tilde{\mathbb{F}}_t^{(8)}$	$\tilde{\mathbb{F}}_t^{(9)}$	$\tilde{\mathbb{F}}_t^{(10)}$	$\tilde{\mathbb{F}}_t^{(11)}$	$\tilde{\mathbb{F}}_t^{(13)}$	$\tilde{\mathbb{F}}_t^{(14)}$	$\tilde{\mathbb{F}}_t^{(17)}$	$\tilde{\mathbb{F}}_t^{(18)}$	$\tilde{\mathbb{F}}_t^{(\Psi_{\boldsymbol{\epsilon}})}$
$\tilde{\mathbb{F}}_t^{(3)}$	1.000	0.525	0.152	0.109	0.063	0.014	0.002	0.001	0.001	0.003	0.003	0.004	0.001
$\tilde{\mathbb{F}}_t^{(4)}$		1.000	0.553	0.206	0.072	0.022	0.004	0.004	0.001	0.003	0.004	0.003	0.001
$\tilde{\mathbb{F}}_t^{(5)}$			1.000	0.820	0.110	0.023	0.001	0.001	0.001	0.002	0.004	0.003	0.001
$\tilde{\mathbb{F}}_t^{(6)}$				1.000	0.134	0.032	0.004	0.004	0.001	0.003	0.005	0.003	0.001
$\tilde{\mathbb{F}}_t^{(8)}$					1.000	0.852	0.168	0.139	0.012	0.018	0.028	0.014	0.004
$\tilde{\mathbb{F}}_t^{(9)}$						1.000	0.381	0.348	0.023	0.046	0.065	0.031	0.005
$\tilde{\mathbb{F}}_t^{(10)}$							1.000	0.996	0.019	0.048	0.144	0.027	0.004
$\tilde{\mathbb{F}}_t^{(11)}$								1.000	0.024	0.054	0.146	0.030	0.004
$\tilde{\mathbb{F}}_t^{(13)}$									1.000	0.923	0.336	0.496	0.018
$\tilde{\mathbb{F}}_t^{(14)}$										1.000	0.623	0.590	0.012
$\tilde{\mathbb{F}}_t^{(17)}$											1.000	0.703	0.014
$\tilde{\mathbb{F}}_t^{(18)}$												1.000	0.111
$\tilde{\mathbb{F}}_t^{(\Psi_{\boldsymbol{\epsilon}})}$													1.000

Note: For $\tilde{\mathbb{F}}_t^{(\cdot)}$, the superscript $(\cdot)$ indicates the single component in Ψ_{BC}^*. See Table 2.3 for further details. Ordering of single components from left to right or top-down are from smallest to largest and does not follow paired orderings.

A remark before moving to the final constructed business cycle indicator, will be the analysis of factor vectors. The factorization of components according to the SVD step of equation 2.10, can be limited to a specific group of components (in our case the common cycle group) and to a specific uni-variate time series, in order to understand the similarities/differences of structural features (frequency). The ability of the SSA methodology to allow such an analysis will be exploited at this stage of the study. The analysis will give insight on the behaviour of the common component as it is passed across each individual time series. We would like to expect that the differences in the structural features of oscillations remain minimal and therefore only attributed to differences in amplitude and phases whilst its frequency remaining intact through out the multivariate set. We apply this factorization to the common cycle component matrix $\boldsymbol{X}_{\Psi^*_{BC}}$ to assess its changes across individual series, this could confirm the extent of commonality between the extracted common cyclical component for each individual series. The computation to factorize the common component matrix is done as follows

$$\underset{(1\times K)}{\boldsymbol{FV^{(i)}}} = \underset{(1\times D)}{\boldsymbol{U}^*_i} \times \underset{(D\times K)}{\boldsymbol{V}^{*\top}} \times \underset{(1\times 1)}{\vartheta_i} \tag{2.22}$$

where

$i = 1, L+1, 2L+1, 3L+1, ..., pL+1$, with p the number of time series in the data,

$D = dim[\Psi^*_{BC}]$, with $[\Psi^*_{BC}] = [z^*_1 \ ... \ z^*_l]$ is a vector of an index sequence that selects specific rows or columns of a matrix according to index numbers $z^*_1, ..., z^*_l$,

$\boldsymbol{U}^*_i = [u_{iz^*_1}, ..., u_{iz^*_l}] = (u_{ij})^{i=pL+1, j=z^*_l}_{i=1, j=z^*_1}$ is the vector of i'th row of matrix $\boldsymbol{U}$ which only includes its elements along columns $\{z^*_1, ..., z^*_l\} \in \Psi^*_{BC}$,

$\boldsymbol{V}^{*\top} = [\boldsymbol{V}_{z^*_1} \ ... \ \boldsymbol{V}_{z^*_l}]^\top = (v^*_{jk})^{j=z^*_l, k=K}_{j=z^*_1, k=1}$ is a $D \times K$ matrix that includes specific rows of matrix $\boldsymbol{V}^\top$ (of equation 2.10) according to index numbers $z^*_1, ..., z^*_l$,

$\vartheta_i = 1/\sqrt{\sum_{j\in\Psi^*_{BC}}(u^2_{ij})}$.

Equation 2.22 computes p number of $\boldsymbol{FV^{(p)}}$ series with length K, where each of them is formed by a linear combination of the row vectors in $V^{*\top}$. The row vectors $\boldsymbol{V}^{*\top}_j$ contains the factor vectors used in constructing the common business cycle indicator which remains the same for each series from 1 to p. $\boldsymbol{U}_i$'s on the other hand will be series specific and will contain the factor loadings which determine the row vector contribution on to $\boldsymbol{FV}$'s. Each series is then multiplied by a standardizing scalar ϑ_i. The resulting factor vector series can then be assessed in terms of their evolution path. Since we know that the matrix $\boldsymbol{V}^{*\top}$ is common for all factor vectors $\boldsymbol{FV(p)}$, the main determinant of deviation in their evolution path amongst each other will depend the factors loadings u_{ij}. [29]

[29] One comment should be made on the ordering of the sequences in Ψ^*_{BC} does not have any impact on the structure of $\boldsymbol{FV}$ as long as any change of the alignments are the same for $\boldsymbol{U}_i$ and $\boldsymbol{V}^{*\top}_j$.

Figure 2.9 provides results of the $\boldsymbol{FV^{(p)}}$ evolutions. We select $\boldsymbol{FV^{(7)}}$ which is the factor vector for the IP series as the baseline. Panels (a), (b), (c), (f) display high level of overlap which is especially evident at parts of the series. For the whole of the panels, segments of the vectors between 0-35 and 105-135 show strong overlap that also comes with a clear phase shift even if the overlap might be deemed low through out other vector segments. The panels also reveal an interesting information referred from macroeconomic theory that was perceived prior to the analysis. Specifically speaking, for calculation of $\boldsymbol{FV^{(5)}}$ and $\boldsymbol{FV^{(6)}}$ required a multiplication of -1, which reflects the common component being counter-cyclical for the macroeconomic indicators of unemployment (UNP) and interest rate differential (INF). Moreover, both these vectors are lagging macro indicators. It is also possible to assess this lag through rightward shifts of $\boldsymbol{FV^{(5)}}$ and $\boldsymbol{FV^{(6)}}$ with respect to the baseline $\boldsymbol{FV^{(7)}}$ in panels (e) and (f) in respective order.

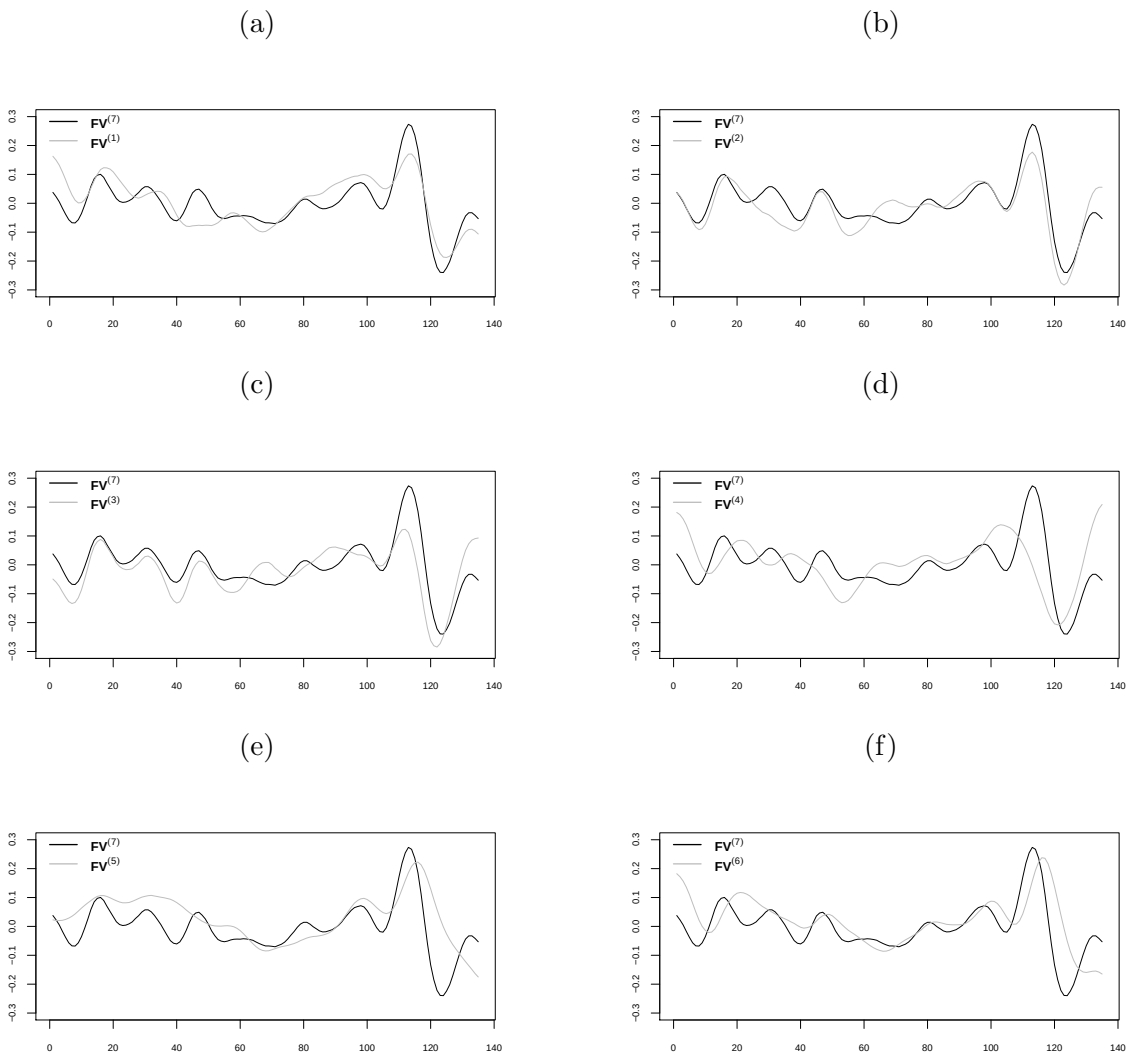
We conclude our comments on this part by stating that factorization of the data matrix reveal the overlap of the common component matrix block V despite factor loadings varying significantly across the multivariate series. The process allows for a final check of the grouping strategy prior to achieving the filter estimate of the buiness cycle indicator and our findings reassures to proceed to the estimation step of the algorithm.

To complement our factorization results, we move the focus on figure 2.10 where the allocations of factors between the number of time series in the data set is presented across the $\boldsymbol{\Psi}^*_{BC}$ indexed components. As was stated in section 1.3.1, the graph shows how loading weights are distributed when moving from one series to the other (the vertical columns) along the common components (horizontal movement). Hence, if all seven series were identical across the common components, then each factor loads $\hat{a}^p_{z^*_i}$, $\hat{b}^p_{z^*_{h_i}}$ would take the value of 1/7 for each series $1 < p < 7$ and across $\boldsymbol{U}3, \boldsymbol{U}4, ..., \boldsymbol{U}13$. It is evident that for each series the common cycle components are nested very diverse. Amongst all series, ICI is the only indicator where all components constitute more than the 1/7 share of total loading weights; $\hat{a}^{(2)}_{z^*_i}$, $\hat{b}^{(2)}_{z^*_{h_i}}$ $> 1/7$.[30] Furthermore, in 8 out of 12 common components, it leads in magnitude of having the highest factor loads. In terms of recording highest factor loadings, RCI, CCI, INT are the other three series which exhibit highest factor weights for the some of the component columns. These factor weights are more specifically being the following 4 weights; $\hat{a}^{(3)}_{13}$, $\hat{b}^{(3)}_{14}$, $\hat{a}^{(4)}_{17}$, $\hat{a}^{(6)}_{8}$.

From these results, it is possible to state that, the complexity of the factor loading matrix in visual form leads to a considerably different result compared to the plot in section 1.3.1. This complexity manifests itself by not being able to clearly separate which component has idiosyncratic or common tendencies from figure 2.10 as there are multiple factor weights along any column with values below 1/7 and above 1/7. In the simulation exercise, the sample size was large and in turn the window length L was wide enough to encompass a more densely populated set of components along the frequency spectrum of the business cycle. Hence we were able to draw a more detailed map of the frequency spectrum of the business cycle and

[30]Identical to the note in Figure 2.9, we denote superscript numbers; 1,2,3,4,5,6,7 to indicator tickers; CNI, ICI, RCI, CCI, UNM, INT, IPI in respective order.

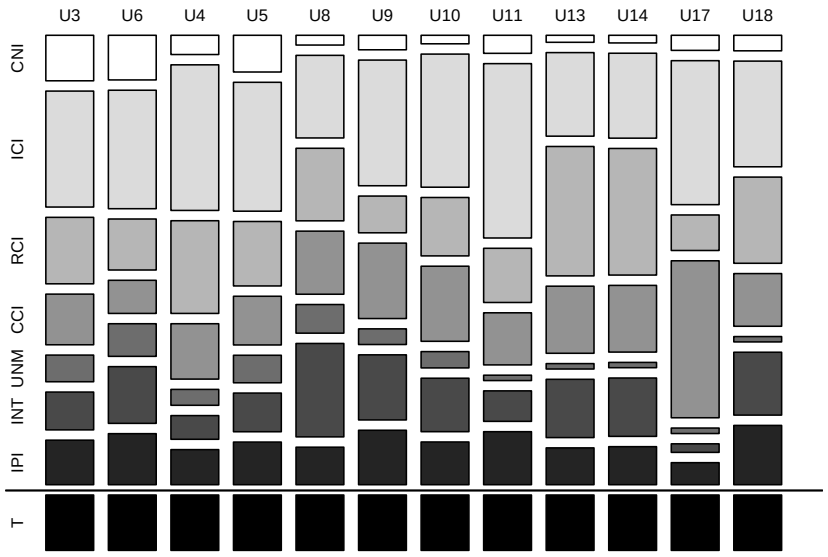
Figure 2.9: Factor Vectors of Business Cycle components ($L = 120$, $K = 135$)



Note: Superscripts 1, 2, 3, 4, 5, 6, 7 indicate K length factor vectors (FV) of the data series $\mathbb{F}_t^{(C,s)}$ with $s = 1, ..., 7$ associated to CNI, ICI, RCI, CCI, UNM, INT, IPI.

cover more components that fell between the upper and lower limits. In a 7 variable system with a much smaller sample size of $N = 254$ we are left with a decreased window length of $L = 120$, which is a fifth of the length for the simulation data. This leads us to have a lower resolution of the business cycle frequency spectrum. Therefore, we rely on the SSA method's capability to reflect the common component with limited number of components; the remaining 6 paired sinusoids. When the common cycle identified by the SSA is represented by large number of components, it would be meaningful to associate some of the components/pairs as being more related to features of some of the individual series while classifying others as common depending on the weight allocation between the series in the system. Seeing that we only have 6 pairs of sinusiods, would be only enough to account for a single common business cycle and certainly not adequate to divide them into idisoyncratic and common cycle classifications.

Figure 2.10: Mosaic plot of factor loadings for a Multivariate dataset



Note: Factor loadings for individual series representing coefficients $a_{z_i^*}^{(s)}$, $b_{z_j^*}^{(s)}$ where z_i^*, z_j^* are paired. Color scheme switches according to superscript $s = \{1, ..., 7\}$. Blocks below the solid black line are average factor loadings; $\sum_{s=1}^{7} a_{z_i^*}^{(s)} = \bar{a}_{z_i^*}$ and $\sum_{s=1}^{7} b_{z_i^*}^{(s)} = \bar{b}_{z_j^*}$ which are equal to 1/7 across z^*'s.

In the synthetic data exercise, a clear definition of the idiosyncratic and common cycles were made. Furthermore, the components were separately introduced to form the cyclical part of one of the series prior to executing the SSA algorithm. However in economic data, such a prior separation is not possible and we are left to the SSA to identify according to the rules conditioned on the grouping strategy set by the researcher. Moreover, we were able to prove that despite presence of idiosyncratic cycle effects, the methodology was able to find amongst the filtered signals, a successful estimate to the common cycle. For the empirical economic application, we consider a 7-variable data set to be represented in the similar form to equation 2.4 but written in compactly such that

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{\mu}_t + (\hat{\boldsymbol{A}}_t + \hat{\boldsymbol{B}}_t)\boldsymbol{T}_{\lambda t} + \boldsymbol{\epsilon}_t \\ \boldsymbol{\mu}_{t+1} &= \boldsymbol{\mu}_t + \boldsymbol{\eta}_t + \boldsymbol{e}_t \\ \boldsymbol{\eta}_{t+1} &= \boldsymbol{\eta}_t + \boldsymbol{\xi}_t \end{aligned} \tag{2.23}$$

where $\boldsymbol{y}_t = (y_{1t}\, y_{2t} \ldots y_{7t})'$, $\boldsymbol{\mu}_t = (\mu_{1t}\, \mu_{2t} \ldots \mu_{7t})'$, $\boldsymbol{\eta}_t = (\eta_{1t} \ldots \eta_{7t})'$, $\boldsymbol{\epsilon}_t \sim N(\mathbf{0}_{[7]}, \boldsymbol{\Lambda}_{\boldsymbol{\epsilon}})$ with $\boldsymbol{\Lambda}_{\boldsymbol{\epsilon}} = diag(\sigma^2_{\epsilon_1}\, \sigma^2_{\epsilon_2} \ldots \sigma^2_{\epsilon_7})$, $\boldsymbol{\xi}_t \sim N(\mathbf{0}_{[7]}, \boldsymbol{\Lambda}_{\boldsymbol{\xi}})$ with $\boldsymbol{\Lambda}_{\boldsymbol{\xi}} = diag(\sigma^2_{\xi_1}\, \sigma^2_{\xi_2} \ldots \sigma^2_{\xi_7})$, $\boldsymbol{e}_t \sim N(\mathbf{0}_{[7]}, \boldsymbol{\Lambda}_{\boldsymbol{e}})$ with $\boldsymbol{\Lambda}_{\boldsymbol{e}} = \mathbf{0}_{[7\times 7]}$ and the system matrices

$$\hat{\boldsymbol{A}}_t = \begin{bmatrix} \hat{a}^{(1)}_{z^*_1} & \hat{a}^{(1)}_{z^*_2} & \cdots & \hat{a}^{(1)}_{z^*_l} \\ \vdots & \vdots & & \vdots \\ \hat{a}^{(7)}_{z^*_1} & \hat{a}^{(7)}_{z^*_2} & \cdots & \hat{a}^{(7)}_{z^*_l} \end{bmatrix}, \quad \hat{\boldsymbol{B}}_t = \begin{bmatrix} \hat{b}^{(1)}_{z^*_{h_1}} & \hat{b}^{(1)}_{z^*_{h_2}} & \cdots & \hat{b}^{(1)}_{z^*_{h_l}} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{b}^{(7)}_{z^*_{h_1}} & \hat{b}^{(7)}_{z^*_{h_2}} & \cdots & \hat{b}^{(7)}_{z^*_{h_l}} \end{bmatrix}$$

$$\boldsymbol{T}_{\lambda t} = \begin{bmatrix} \cos(\lambda^*_{z^*_1} t) & \sin(\lambda^*_{z^*_{h_1}} t) \\ \vdots & \vdots \\ \cos(\lambda^*_{z^*_l} t) & \sin(\lambda^*_{z^*_{h_l}} t) \end{bmatrix}$$

Figure 2.11 displays the tracking performance of the SSA filter based on our algorithm. The panels plots the SSA multivariate filter estimate selected from the ICI ($\mathbb{F}^{(2,\Psi^*_{BC})}$) series against two cycle indicators of economic activity for the Euro region; the EuroCOIN indicator in Altissimo et al. (2000) and Altissimo et al. (2010), and the 3-month growth indicator obtained by the EU wide real GDP series. These two series have been specifically selected to serve as benchmark cycles as they were also the choice in Valle e Azevedo et al. (2006) in assessing the performance of their parametric multivariate band-pass filter. Moreover, unlike in the controlled experiment of section 2.5, where we used the nested a common component as a benchmark to asses the SSA based filter estimates of our algorithm, in a real world application the benchmark cycle is impossible to identify a-prior. However, using an indicator of the common cycle provides a proxy to the common component prior to the analysis im our empirical part the study as each individual series in our data set is an economic activity indicator of the Euro-zone region. From our multivariate data set, the SSA algorithm allows the generation of 7 estimates of the common cycle. However, carefully plotting each filter estimate against the benchmark cycles and a through visual inspection leads us to consider the best fit of the common economic cycle to be represented by the filter estimate associate to ICI which we will denote as $\tilde{\mathbb{F}}^{(2,\Psi^*_{BC})}_t$.[31] We start by analysing panel (a) and see that $\tilde{\mathbb{F}}^{(2,\Psi^*_{BC})}_t$ has a high success rate in tracking the EU wide cycle. The panel depicts certain periods appearing to converge to a near perfect overlap (especially during the growth episodes of 2013 and 2015). Additionally, while synchronization features like phase between cycles cannot be directly retrieved from SSA analysis and one requires additional procedures for their computation, the figure illustrates that the estimate slightly lags the EU common cycle indicator. This lag is not persistent through out the sample period but in reality comes to present itself as a correction. This is because when looking at specific turning points where the recession episodes are ending, the lag in $\tilde{\mathbb{F}}^{(2,\Psi^*_{BC})}_t$ implies that those points are closer to the shaded edges compared to the EuroCOIN. We could take this conclusion as an improvement obtained by our proposed rule based algorithm. In panel (b), we also see a reasonably good fit in turning points of our estimate comparison to the EU growth cycle indicator. The

[31] We exclude the graphs depicting each $\tilde{\mathbb{F}}^{(p,\Psi^*_{BC})}_{\approx}$ for $p = 1, ..., 7$ business cycle estimate for the sake of brevity and for the reason that they do not add information to our analysis.

filter estimate provides a precise identification of the cyclical traces and almost can be assigned as the smoothed estimate of the data itself. Notice that the growth cycle naturally is a quarterly spaced data which is reflected by the step like movements of the data points. Therefore, one can use $\tilde{\mathbb{F}}_t^{(2,\Psi^*_{BC})}$ as a proxy to the unobserved monthly growth cycle indicator since growth cycles will require gap filling procedures to be considered as the underlying monthly cycle component.

Considering that the common EU cycle indicator was structured using a much larger information set (145 EU region economic data), our algorithm rule of the SSA based filter provides an alternative procedure to constructing and identifying region wide common cycle components because our estimate was computed only from a 7 series data set. Secondly, while a critique can be made regarding the fact that $\tilde{\mathbb{F}}_t^{(2,\Psi^*_{BC})}$ lags the common cycle indicator, in effect this increases slightly the accuracy for predicting the end time of certain EU recession episodes relevant to the sample period. Another outcome from the findings in figure 2.11 provides an additional purpose to our filter estimate. Our main purposes of our filter estimate was initially to extract the common business cycle component identified by our rule based algorithm. Seeing that we achieve a reasonably successful tracker of the positioning of the Euro region quarterly growth rates, we can adopt the algorithm rule for constructing latent monthly growth cycle indicators since these indicators require monthly transformation steps for smoothing the data.

To further inspect the achievement of the SSA filter, we develop two competing multivariate UCM models that are able to extract one common cycle. Since the SSA is a non-parametric model, these models will be viewed as an alternative in parametric form. We set the main structure of the UCM models in compact from in the following way;

$$\boldsymbol{y}_t = \boldsymbol{\mu}_t + \boldsymbol{\Theta}\boldsymbol{\varphi}_t + \boldsymbol{\Xi}\boldsymbol{\psi}_t + \boldsymbol{\epsilon}_t \tag{2.24}$$

where the 7×1 row vectors of $\boldsymbol{y}_t$, $\boldsymbol{\mu}_t$, $\boldsymbol{\epsilon}_t$ have identical forms to equation 2.23, with the idiosyncratic stochastic cycle vector $\boldsymbol{\varphi}_t = (\boldsymbol{\varphi_{1t}}\ \boldsymbol{\varphi_{2t}}\ \boldsymbol{\varphi_{3t}}\ \boldsymbol{\varphi_{4t}}\ \boldsymbol{\varphi_{5t}}\ \boldsymbol{\varphi_{6t}}\ \boldsymbol{\varphi_{7t}})'$ which can be expressed for each individual idiosyncratic cycle $\boldsymbol{\varphi}_{p,t+1}$ by the equation

$$\begin{bmatrix} \varphi_{p,t+1} \\ \varphi^*_{p,t+1} \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda & \sin\lambda \\ -\sin\lambda & \cos\lambda \end{bmatrix} \begin{bmatrix} \varphi_{p,t} \\ \varphi^*_{p,t} \end{bmatrix} + \begin{bmatrix} \kappa_{p,t} \\ \kappa^*_{p,t} \end{bmatrix} \tag{2.25}$$

.

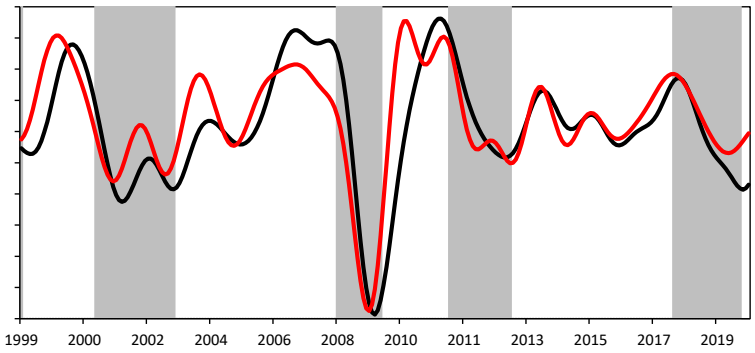
Equation 2.25 can be represented for the full data in more compact form represented by

$$\boldsymbol{\varphi_{t+1}} = \rho \left(\boldsymbol{I}_{[7]} \otimes \begin{bmatrix} \cos\lambda & \sin\lambda \\ -\sin\lambda & \cos\lambda \end{bmatrix} \right) \boldsymbol{\varphi_t} + \begin{bmatrix} \kappa_{p,t} \\ \kappa^*_{p,t} \end{bmatrix} \tag{2.26}$$

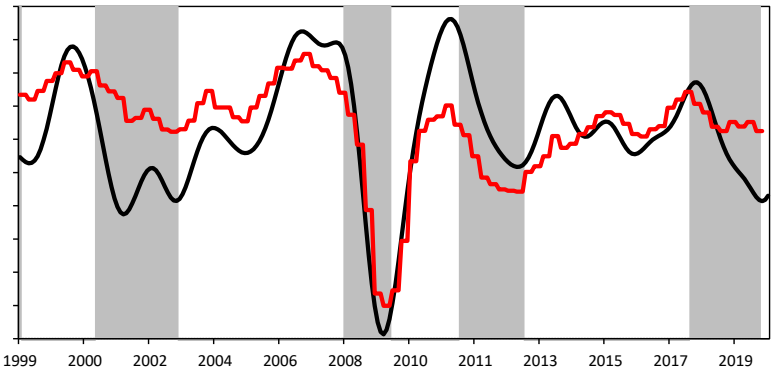
where $\boldsymbol{\kappa}_t = (\kappa_{1t}\ \kappa^*_{1t}\ \kappa_{2t}\ \kappa^*_{2t}\ \kappa_{3t}\ \kappa^*_{3t}\ \kappa_{4t}\ \kappa^*_{4t}\ \kappa_{5t}\ \kappa^*_{5t}\ \kappa_{6t}\ \kappa^*_{6t}\ \kappa_{7t}\ \kappa^*_{7t})' \sim N(\boldsymbol{0}_{[2p]}, \boldsymbol{\Lambda}_\kappa)$, and the operator $\otimes$ is the Kronecker product which multiplies each element of the first matrix with the complete second matrix.

Figure 2.11: Business cycle indicators vs the SSA filter estimate

(a) $\widetilde{\mathbb{F}}_t^{(2,\Psi^*_{BC})}$ vs EuroCOIN



(b) $\widetilde{\mathbb{F}}_t^{(2,\Psi^*_{BC})}$ vs Annualized GDP growth *qoq*



Note: Recession shades are based on the OECD recession indicator (see Appendix 2.7 for details). The recession indicator covers prior and the full sample period with only the final 3 data points missing; 11/2019-01/2020.

$\boldsymbol{\Lambda}_{\kappa} = Blockdiag\,[\boldsymbol{\Sigma}_{\kappa_1}\ \boldsymbol{\Sigma}_{\kappa_2}\ ...\ \boldsymbol{\Sigma}_{\kappa_7}]$, $\boldsymbol{\Sigma}_{\kappa_p} = \sigma^2_{\kappa_p}\boldsymbol{I}_{[2]}$ and $p = 1, 2, ..., 7$. The correlation structure between the idiosyncratic cycle disturbances for this model is zero which makes the cycle variance-covariance structure at its most simplest form to not further complicate computation.

The common stochastic cycle $\boldsymbol{\psi_t}$ follows the following representation

$$\begin{bmatrix}\psi_{t+1}\\ \psi^*_{t+1}\end{bmatrix} = \rho\begin{bmatrix}\cos\lambda & \sin\lambda\\ -\sin\lambda & \cos\lambda\end{bmatrix}\begin{bmatrix}\psi_t\\ \psi^*_t\end{bmatrix} + \begin{bmatrix}\kappa^{\dagger}_t\\ \kappa^{\dagger *}_t\end{bmatrix} \tag{2.27}$$

with $\rho \in (0,\ 1)$, $\boldsymbol{T}_{\psi}$ is the 2×2 trigonometric matrix also coined as rotation matrix, $\lambda \in [\pi/48,\ \pi/9]$, $(\kappa^{\dagger}_t\ \kappa^{\dagger *}_t)' \sim N(\boldsymbol{0}_{[2]}, \boldsymbol{\Lambda}_{\kappa^{\dagger}})$, $\boldsymbol{\Lambda}_{\kappa^{\dagger}} = \sigma^2_{\kappa^{\dagger}}\boldsymbol{I}_{[2]}$. The correlation structure of both the stochastic cycles aa and aa are determined by the following cross-covariance functions

$$\boldsymbol{\Gamma}_{\varphi}(k) = \frac{\rho^k\cos(k\lambda)}{1-\rho^2}\boldsymbol{\Lambda}_{\kappa}, \qquad \boldsymbol{\Gamma}_{\psi}(k) = \frac{\rho^k\cos(k\lambda)\sigma_{\kappa^{\dagger}}}{1-\rho^2}\boldsymbol{I}_{[2]} \tag{2.28}$$

To complete the UCM system, we define in equation 2.24 the idiosyncratic cycle coefficient matrix as $\boldsymbol{\Theta} = \boldsymbol{I}_{[7]} \otimes (1\ 0)$ and the common cycle coefficient matrix $\boldsymbol{\Xi}$ as follows

$$\boldsymbol{\Xi} = \begin{bmatrix}\gamma_1\cos(\lambda\zeta_1) & \gamma_1\sin(\lambda\zeta_1)\\ \gamma_2\cos(\lambda\zeta_2) & \gamma_2\sin(\lambda\zeta_2)\\ \vdots & \vdots\\ \gamma_6\cos(\lambda\zeta_6) & \gamma_6\sin(\lambda\zeta_6)\\ 1 & 0\end{bmatrix} \tag{2.29}$$

where $\{\gamma_1,\ \gamma_2,\ \gamma_3,\ \gamma_4,\ \gamma_5,\ \gamma_6, 1\} \subset \mathbb{R}$ and $\{\zeta_1,\ \zeta_2,\ \zeta_3,\ \zeta_4,\ \zeta_5,\ \zeta_6\} \in (-\pi/2,\ \pi/2)$. The above definitions and system matrices all form the UCM models in this study. Alternatively, a second UCM model is introduced where the only difference is that the common component is a second ordered cycle of Harvey and Trimbur (2003). In this case, the the common stochastic cycle specification of equation 2.27 is changed to

$$\begin{aligned}\boldsymbol{\psi}^{(m)}_{t+1} &= \rho\boldsymbol{T}_{\lambda}\boldsymbol{\psi}^{(m)}_t + \boldsymbol{\psi}^{(m-1)}_{t+1}\\ \boldsymbol{\psi}^{(0)}_{t+1} &= \boldsymbol{\kappa}^{\dagger}_t\end{aligned} \tag{2.30}$$

where $\boldsymbol{T}_{\psi}$ is the rotation matrix in equation 2.27, and with $\boldsymbol{\kappa}^{\dagger}_t = (\kappa^{\dagger}_t, \kappa^{\dagger *}_t)' \sim N(\boldsymbol{0}_{[4\times 4]}, \boldsymbol{\Lambda}^{(2)}_{\kappa^{\dagger}})$, with

$$\boldsymbol{\Lambda}^{(2)}_{\kappa^{\dagger}} = \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix} \otimes \sigma^2_{\kappa^{\dagger}}\boldsymbol{I}_{[2]}$$

Slight modifications are further required on the system matrices such as the rotation matrix and the auto-covariance function of the cycles in the following form

$$\boldsymbol{T}_{\psi^{(2)}} = \begin{bmatrix} \rho T_{\psi} & \boldsymbol{I}_{[2]} \\ \boldsymbol{0}_{[2\times 2]} & \rho T_{\psi} \end{bmatrix}, \boldsymbol{\Gamma}_{\psi^{(2)}}(k) = \rho^{k}\sigma^{2}_{\kappa^{\dagger}} d \begin{bmatrix} (1+\rho^{2})dT_{\psi} & \rho d\boldsymbol{T}'_{\psi} \\ \rho d\boldsymbol{T}_{\psi} & \boldsymbol{I}_{[2]} \end{bmatrix}, \; d = (1+\rho^{2})^{-1} \tag{2.31}$$

When the common cycle is in second order form, the model will be coined as UCM-2. Equation 2.24 also forms the measurement equation required for the state space form. Once the state update equation is also set using the system matrices in Appendix 2.34, we can carry out the Kalman filter recursions to extract the common cycle estimates from both models. Figure 2.12 presents the comparative analysis of the results for our structural models, UCM-1 and UCM-2 and the filter estimate of the SSA. In panel (a), the common cycle extracted by UCM-1 fairs similarly with its SSA counterpart. The main source of deviation would be linked to how trend component is identified which may explain the encirclement of complete cycle lengths observed in the estimation sample. However, trend specification is not our main concern as our focus is on capturing the common oscillating motion of the business cycle. However, the cycle estimate of UCM-1 underperforms in obtaining a smoothed estimate and therefore it is evident that the SSA procedure is superior in terms of noise reduction. We can refine our UCM-1 based estimate by specifying a second order stochastic cycle as a common component. Panel (b) displays the evolution path of increased smoothness of the cyclical component. The elimination of the irregular high frequencies contaminating the first oder cycle of $\widetilde{\psi}^{(1)}_{t|n}$ makes $\widetilde{\psi}^{(2)}_{t|n}$ to be an ideal substitute to $\widetilde{\psi}_{t|n}$ of UCM-1. Panel (c) allows for a more precise visual analysis on the similarities/discrepancies related to the positioning of oscillations and also on the synchronization of the competing cycle estimates. Similar to panel (a) the encirclement of complete peak and trough episodes can be traced back to the form of the trend defined in the model and therefore is discarded in our evaluation. Furthermore, with the exception of the first peak episode witnessed at the start of the millennia, there is no systematic lead of any form between the alternating indicators.

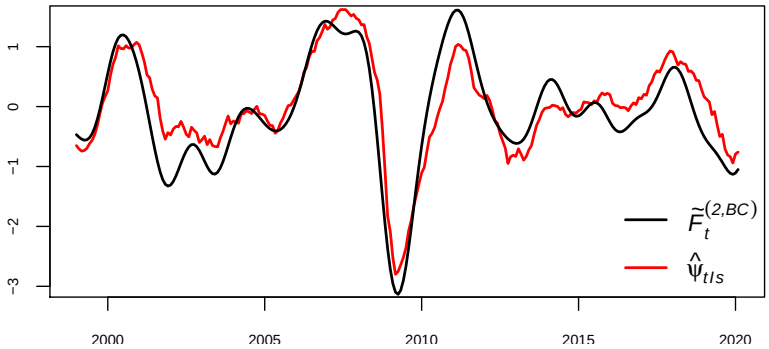
To verify some of the results obtained in figure 2.12 and to get further information on the identified common cycle, we investigate ML estimates of both UCM model parameters. We assume that the hyperparameter vector containing model parameters to take the same initial values to initialize the Kalman filter process. We first investigate the scale loadings associated to the common component. To normalize the loadings we set γ_{IP} to unit value which allows IP series to harbor the coinciding cycle benchmark with respect to all other coefficients to be estimated around the unit magnitude. The ML procedure reveals that the ICI series has the highest estimated value $\widehat{\gamma}_2$ at 1.9, meaning that it doubles the common cycle amplitude. This reinforces our findings in the mosaic graph of figure 2.10, where we found that the factor loadings estimated for ICI ($\{\hat{a}^{(2)}_{z^*_i}, \hat{b}^{(2)}_{z^*_{h_i}}\}_{z^*_i, z^*_{h_i} \in \Psi^*_{BC}}$) had the highest values across the common components that constructed the business cycle. RCI and IP follows ICI in terms of factor loading weights which is also evident in this table as $\widehat{\gamma}_3 \approx 1.4$ and $\gamma_{IP} = 1$. Moreover, common cycles of UNM and INT are found to be nested so that they are counter cyclical and also exhibit dampened common cycle

waves. This confirms our analysis in figure 2.9 where $\boldsymbol{FV^{(5)}}$ and $\boldsymbol{FV^{(6)}}$ were aligned with $\boldsymbol{FV^{(7)}}$ with a multiplicity of -1. Finally, we look into the phase shift between the common cycle evaluated by ζ_p which indicate the shift of the auto-correlation function of the common cycle of a series with respect to the baseline common cycle $\widetilde{\psi}_{t|n}$ of IP series. A right shift of the auto-correlation function will indicate a lead over $\widetilde{\psi}_{t|n}$. Hence, concentrating on positive values, it can be seen that ICI series possess a lead of 1.6 to 1.2 months over the coinciding cycle. Hence this final result establishes that the ICI (series $\mathbb{F}_t^{(2)}$ in the SSA specification) can be considered as a leading indicator of the common economic cycle.

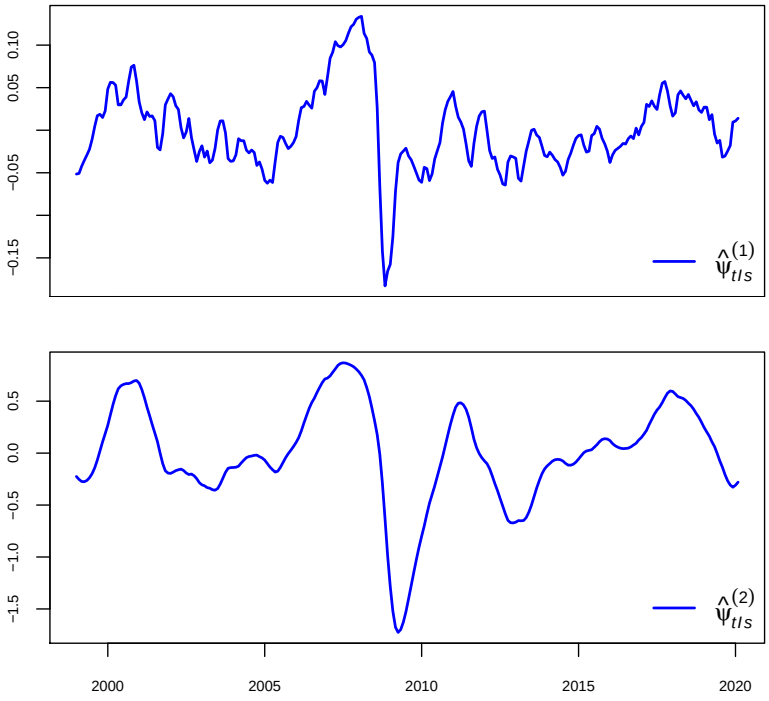
These results also convey that the non-parametric SSA based algorithm which we modified for the identification of common business cycle components can be an alternative to a parametric based multivariate cycle extraction procedure.

Figure 2.12: UCM based common cycle extraction

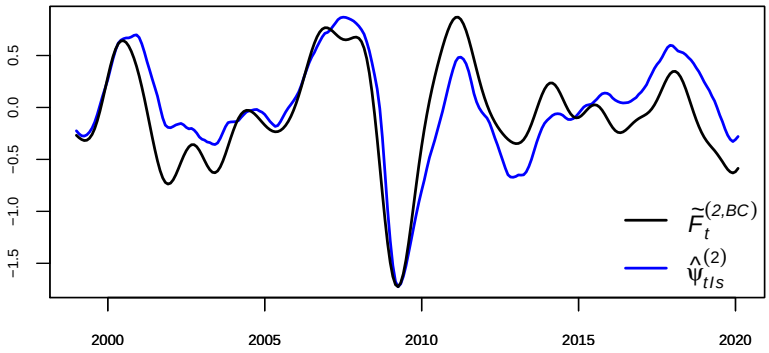
(a) UCM-1 common cycle estimate vs SSA: ($\hat{\psi}_{t|s}$ vs $\tilde{\mathbb{F}}_t^{(2,\boldsymbol{BC})}$)



(b) UCM-2 common ordered cycles ($\hat{\psi}_{t|s}^{(1)}$, $\hat{\psi}_{t|s}^{(2)}$)



(c) UCM-2 based 2nd order cycle vs SSA: ($\hat{\psi}_{t|s}^{(2)}$ vs $\tilde{\mathbb{F}}_t^{(2,\boldsymbol{BC})}$)



Note: The initial parameter vector required for running the Kalman filter in both models are identical. Red lines indicate UCM-1, blue indicate UCM-2 smoothed estimates.

Table 2.6: ML Estimates of the parameters in common cycle UCM

	UCM-1		UCM-2	
Parameter	Estimate	St.Error	Estimate	St.Error
$\sigma^2_{\epsilon_1}$	0.0076	0.0011	0.0075	0.0010
$\sigma^2_{\epsilon_2}$	0.0000	0.0001	0.0013	0.0012
$\sigma^2_{\epsilon_3}$	0.0345	0.0050	0.0325	0.0050
$\sigma^2_{\epsilon_4}$	0.0000	0.0003	0.0000	0.0001
$\sigma^2_{\epsilon_5}$	0.0004	0.0001	0.0005	0.0001
$\sigma^2_{\epsilon_6}$	0.0000	0.0000	0.0000	0.0000
$\sigma^2_{\epsilon_7}$	0.0089	0.0014	0.0087	0.0014
$\sigma^2_{\xi_1}$	0.0001	0.0001	0.0001	0.0000
$\sigma^2_{\xi_2}$	0.0000	0.0000	0.0000	0.0000
$\sigma^2_{\xi_3}$	0.0000	0.0000	0.0000	0.0000
$\sigma^2_{\xi_4}$	0.0000	0.0000	0.0000	0.0000
$\sigma^2_{\xi_5}$	0.0001	0.0001	0.0001	0.0000
$\sigma^2_{\xi_6}$	0.0000	0.0000	0.0001	0.0001
$\sigma^2_{\xi_7}$	0.0000	0.0000	0.0000	0.0000
$\sigma^2_{\kappa^\dagger}$	0.0040	0.0006	0.0006	0.0002
γ_1	0.6795	0.1126	0.6636	0.0915
γ_2	1.9598	0.3442	1.9974	0.2351
γ_3	1.3991	0.2804	1.4004	0.2157
γ_4	0.9862	0.2310	1.1026	0.2038
γ_5	-0.1826	0.0638	-0.3944	0.0692
γ_6	-0.3103	0.1707	-0.5642	0.1744
$\sigma^2_{\kappa_1}$	0.0009	0.0009	0.0013	0.0008
$\sigma^2_{\kappa_2}$	0.0074	0.0025	0.0097	0.0031
$\sigma^2_{\kappa_3}$	0.0170	0.0044	0.0229	0.0053
$\sigma^2_{\kappa_4}$	0.0236	0.0024	0.0242	0.0023
$\sigma^2_{\kappa_5}$	0.0009	0.0003	0.0006	0.0002
$\sigma^2_{\kappa_6}$	0.0166	0.0016	0.0159	0.0016
$\sigma^2_{\kappa_7}$	0.0034	0.0014	0.0038	0.0016
ρ	0.9735	0.0050	0.9269	0.0160
$p = 2\pi/\lambda$	54.9835	199.923	50.712	89.663
ζ_1	0.1943	49.991	0.1180	31.8514
ζ_2	1.5669	0.0206	1.2488	0.0240
ζ_3	1.5659	0.0382	1.0750	0.0313
ζ_4	1.5569	0.0990	1.0941	0.0965
ζ_5	-1.5682	0.0246	-1.1952	0.0168
ζ_6	-1.5676	0.0376	-1.2402	0.0328

Note: St.Error indicate the standard error, computed using the Delta method discussed in Appendix D.

2.7 Conclusion

In this paper we implement a pure data driven method in identifying common cycles for a multivariate data set and provide empirical analysis for constructing a Euro region wide economic cycle indicator. The selected methodology is a multivariate non-parametric filtering approach; the Vertical Multivariate Singular Spectrum Analysis (V-MSSA) of Hassani and Mahmoudvand (2013) and Golyandina et al. (2013). The methodology allows decompose a time series into many sub components without depending on a structural model. By exploiting this ability of the SSA, the paper aims to first extract cyclical components based on frequency characteristics and then follow by choosing only common cyclical component pairs with-in the business cycle frequency spectrum according to an algorithm rule. These components will then be aggregated for constructing an EU region wide Business cycle indicator.

The independence of the SSA methodology to any economic model and its data intense approach gives it high level of flexibility and a wide area of applications. Furthermore the number of applications on business cycles is limited and therefore a literature gap is prevalent. In terms of application, our paper shares a parallel aim of papers Altissimo et al. (2000), (2006) and de Carvalho and Rua (2017). These papers try to construct regional or national level business cycle indicators when there are competing time series variables in the data set that could play the role of a business cycle tracker.

The paper outlines each steps of the algorithm that will eventually identify the SSA filter to act as a band-pass filter. Modifying the algorithm to incorporate the frequency characteristics that maps the business cycle spectrum and the SSA based definition of sine/cosine pairs, we are able to define the business cycle (the common cycle component) as a linear combination of sine/cosine waves that have peak frequencies that fall within the business cycle spectrum. By specifying the business cycle definition, the paper carries out a controlled experiment where common and idiosyncratic cycles are nested into the individual series of a multivariate data set. The simulation results confirm that the algorithm can identify and construct a common cycle that fall within the business cycle frequency and that the filtered estimate is successful in tracking the original common component nested in each series even in the presence of idiosyncratic similar cycles. The study follows with an empirical analysis in-line to the framework set in Valle e Azevedo et al. (2006) with the aim to identify a Euro region business cycle indicator. The SSA based filter estimate is compared with Euro region economic activity indicators; the EuroCoin and the quarterly GDP growth rate of the EU area. Our results presents evidence of a successful candidate for tracing the cyclical position of the EU economy from a much smaller data set. Moreover, the constructed indicator also could serve as an unobserved proxy for a monthly growth cycle. A further analysis is also conducted to reveal whether the SSA based approach can be considered as an alternative to parametric filtering methods. In this framework, the study presents results from the empirical analysis of common cycle extraction using two multivariate UCM alternatives. Comparing results with the SSA based filter, we are able to reinforce our initial findings and use the parametric based results as a robustness check.

References

Allen, M. R. and Smith, L. A. (1997). Optimal filtering in singular spectrum analysis. *Physics letters A*, 234(6):419–428.

Altissimo, F., Cristadoro, R., Forni, M., Lippi, M., and Veronese, G. (2010). New eurocoin: Tracking economic growth in real time. *The review of economics and statistics*, 92(4):1024–1034.

Altissimo, F., Marchetti, D. J., and Oneto, G. P. (2000). The italian business cycle: coincident and leading indicators and some stylized facts. *Giornale degli economisti e Annali di economia*, pages 147–220.

Atkeson, A., Ohanian, L. E., et al. (2001). Are phillips curves useful for forecasting inflation? *Federal Reserve bank of Minneapolis quarterly review*, 25(1):2–11.

Baxter, M. and King, R. G. (1999). Measuring business cycles: approximate band-pass filters for economic time series. *Review of economics and statistics*, 81(4):575–593.

Beneki, C., Eeckels, B., and Leon, C. (2012). Signal extraction and forecasting of the uk tourism income time series: A singular spectrum analysis approach. *Journal of Forecasting*, 31(5):391–400.

Broomhead, D., King, G., and Pike, E. (1987). Singular spectrum analysis with application to dynamical systems. *Noise and Fractal, IOP*.

Broomhead, D. S. and King, G. P. (1986). Extracting qualitative dynamics from experimental data. *Physica D: Nonlinear Phenomena*, 20(2-3):217–236.

Christiano, L. J. and Fitzgerald, T. J. (2003). The band pass filter. *international economic review*, 44(2):435–465.

de Carvalho, M. and Martos, G. (2020). Brexit: Tracking and disentangling the sentiment towards leaving the eu. *International Journal of Forecasting*, 14(1):1128–1137.

de Carvalho, M., Rodrigues, P. C., and Rua, A. (2012). Tracking the us business cycle with a singular spectrum analysis. *Economics Letters*, 114(1):32–35.

de Carvalho, M. and Rua, A. (2017). Real-time nowcasting the us output gap: Singular spectrum analysis at work. *International Journal of Forecasting*, 33(1):185–198.

Durbin, J. and Koopman, S. J. (2012). *Time series analysis by state space methods*. Oxford university press.

Ghil, M., Allen, M., Dettinger, M., Ide, K., Kondrashov, D., Mann, M., Robertson, A. W., Saunders, A., Tian, Y., Varadi, F., et al. (2002). Advanced spectral methods for climatic time series. *Reviews of geophysics*, 40(1):3–1.

Ghodsi, M., Hassani, H., Rahmani, D., and Silva, E. S. (2018). Vector and recurrent singular spectrum analysis: which is better at forecasting? *Journal of Applied Statistics*, 45(10):1872–1899.

Golyandina, N. (2010). On the choice of parameters in singular spectrum analysis and related subspace-based methods. *arXiv preprint arXiv:1005.4374*.

Golyandina, N. (2020). Particularities and commonalities of singular spectrum analysis as a method of time series analysis and signal processing. *Wiley Interdisciplinary Reviews: Computational Statistics*, 12(4):e1487.

Golyandina, N., Korobeynikov, A., Shlemov, A., and Usevich, K. (2013). Multivariate and 2d extensions of singular spectrum analysis with the rssa package. *arXiv preprint arXiv:1309.5050*.

Golyandina, N., Korobeynikov, A., and Zhigljavsky, A. (2018). *Singular spectrum analysis with R*. Springer.

Golyandina, N., Nekrutkin, V., and Zhigljavsky, A. A. (2001). *Analysis of time series structure: SSA and related techniques*. CRC press.

Golyandina, N. and Shlemov, A. (2013). Variations of singular spectrum analysis for separability improvement: non-orthogonal decompositions of time series. *arXiv preprint arXiv:1308.4022*.

Golyandina, N. and Shlemov, A. (2015). Variations of singular spectrum analysis for separability improvement: non-orthogonal decompositions of time series. *Statistics and its Inference*, 8(03):277–294.

Gonzalez, R. B., Lima, J., Marinho, L., et al. (2015). Business and financial cycles: an estimation of cycles' length focusing on macroprudential policy. *Central Bank of Brazil Working Papers*, (385).

Groth, A., Ghil, M., Hallegatte, S., and Dumas, P. (2011). Identification and reconstruction of oscillatory modes in us business cycles using multivariate singular spectrum analysis. In *Workshop on Frequency Domain Research in Macroeconomics and Finance*. Bank of Finland, Helsinki.

Groth, A., Ghil, M., Hallegatte, S., and Dumas, P. (2015). The role of oscillatory modes in us business cycles. *OECD Journal: Journal of Business Cycle Measurement and Analysis*, 2015(1):63–81.

Hamilton, J. D. (2018). Why you should never use the hodrick-prescott filter. *Review of Economics and Statistics*, 100(5):831–843.

Harvey, A. C. and Jaeger, A. (1993). Detrending, stylized facts and the business cycle. *Journal of applied econometrics*, 8(3):231–247.

Harvey, A. C. and Trimbur, T. M. (2003). General model-based filters for extracting cycles and trends in economic time series. *Review of Economics and Statistics*, 85(2):244–255.

Hassani, H., Dionisio, A., and Ghodsi, M. (2010a). The effect of noise reduction in measuring the linear and nonlinear dependency of financial markets. *Nonlinear Analysis: Real World Applications*, 11(1):492–502.

Hassani, H., Heravi, S., and Zhigljavsky, A. (2009). Forecasting european industrial production with singular spectrum analysis. *International journal of forecasting*, 25(1):103–118.

Hassani, H. and Mahmoudvand, R. (2013). Multivariate singular spectrum analysis: A general view and new vector forecasting approach. *International Journal of Energy and Statistics*, 1(01):55–83.

Hassani, H. and Mahmoudvand, R. (2018). *Singular Spectrum Analysis: Using R*. Springer.

Hassani, H., Mahmoudvand, R., and Zokaei, M. (2011a). Separability and window length in singular spectrum analysis. *Comptes rendus mathematique*, 349(17-18):987–990.

Hassani, H., Mahmoudvand, R., Zokaei, M., and Ghodsi, M. (2012). On the separability between signal and noise in singular spectrum analysis. *Fluctuation and Noise Letters*, 11(02):1250014.

Hassani, H., Rua, A., Silva, E. S., and Thomakos, D. (2019). Monthly forecasting of gdp with mixed-frequency multivariate singular spectrum analysis. *International Journal of Forecasting*, 35(4):1263–1272.

Hassani, H., Soofi, A., and Avazalipour, M. S. (2011b). Forecasting gdp with aggregated and sectoral data. *Fluctuation and Noise Letters*, 10(03):249–265.

Hassani, H., Soofi, A. S., and Zhigljavsky, A. (2013). Predicting inflation dynamics with singular spectrum analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176(3):743–760.

Hassani, H., Soofi, A. S., and Zhigljavsky, A. A. (2010b). Predicting daily exchange rate with singular spectrum analysis. *Nonlinear Analysis: Real World Applications*, 11(3):2023–2034.

Hassani, H. and Thomakos, D. (2010). A review on singular spectrum analysis for economic and financial time series. *Statistics and its Interface*, 3(3):377–397.

Hassani, H., Webster, A., Silva, E. S., and Heravi, S. (2015). Forecasting us tourist arrivals using optimal singular spectrum analysis. *Tourism Management*, 46:322–335.

Hodrick, R. J. and Prescott, E. C. (1997). Postwar us business cycles: an empirical investigation. *Journal of Money, credit, and Banking*, pages 1–16.

Hyndman, R. J. and Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.

Hyndman, R. J., Khandakar, Y., et al. (2007). *Automatic time series for forecasting: the forecast package for R*. Number 6/07. Monash University, Department of Econometrics and Business Statistics ....

Ivanov, L. T. (2015). Is" the ideal filter" really ideal: the usage of frequency filtering and spurious cycles. *South-Eastern Europe Journal of Economics*, 3(1).

Kitagawa, G. (2020). *Introduction to Time Series Modeling with Applications in R*. CRC press.

Leon, C. (2015). Decomposition of the european gdp based on singular spectrum analysis. *MPRA Paper No. 65812*.

Mahmoudvand, R., Najari, N., and Zokaei, M. (2013). On the optimal parameters for reconstruction and forecasting in singular spectrum analysis. *Communications in Statistics-Simulation and Computation*, 42(4):860–870.

Marques, C., Ferreira, J., Rocha, A., Castanheira, J., Melo-Gonçalves, P., Vaz, N., and Dias, J. (2006). Singular spectrum analysis and forecasting of hydrological time series. *Physics and Chemistry of the Earth, Parts A/B/C*, 31(18):1172–1179.

Mohr, M. F. (2005). A trend-cycle (-season) filter. *ECB Working Paper*, page 499.

Nelson, C. R. and Kang, H. (1981). Spurious periodicity in inappropriately detrended time series. *Econometrica: Journal of the Econometric Society*, pages 741–751.

Palus, M. and Novotna, D. (2004). Enhanced monte carlo singular system analysis and detection of period 7.8 years oscillatory modes in the monthly nao index and temperature records. *Nonlinear Processes in Geophysics, European Geosciences Union*, 11(5/6):721–729.

Patterson, K., Hassani, H., Heravi, S., and Zhigljavsky, A. (2011). Multivariate singular spectrum analysis for forecasting revisions to real-time data. *Journal of Applied Statistics*, 38(10):2183–2211.

Pelagatti, M. M. (2006). Estimation and optimal filtering for a common stochastic cycle shifted in continuous time. *in Società Italiana di Statistica: Atti della XLIII Riunione Scientifica*.

Pelagatti, M. M. (2019). *Time series modelling with unobserved components*. Chapman and Hall/CRC.

Rodriguez-Aragon, L. J. and Zhigljavsky, A. (2010). Singular spectrum analysis for image processing. *Statistics and Its Interface*, 3(3):419–426.

Saayman, A. and de Klerk, J. (2019). Forecasting tourist arrivals using multivariate singular spectrum analysis. *Tourism Economics*, 25(3):330–354.

Schoellhamer, D. H. (2001). Singular spectrum analysis for time series with missing data. *Geophysical research letters*, 28(16):3187–3190.

Sella, L., Vivaldo, G., Groth, A., and Ghil, M. (2016). Economic cycles and their synchronization: A comparison of cyclic modes in three european countries. *Journal of Business Cycle Research*, 12(1):25–48.

Shumway, R. H., Stoffer, D. S., and Stoffer, D. S. (2000). *Time series analysis and its applications*, volume 3. Springer.

Silva, E. S. and Hassani, H. (2015). On the use of singular spectrum analysis for forecasting us trade before, during and after the 2008 recession. *International Economics*, 141:34–49.

Silva, E. S., Hassani, H., and Heravi, S. (2018a). Modeling european industrial production with multivariate singular spectrum analysis: A cross-industry analysis. *Journal of Forecasting*, 37(3):371–384.

Silva, E. S., Hassani, H., and Otero, J. (2018b). Forecasting inflation under varying frequencies. *Electronic Journal of Applied Statistical Analysis*, 11(1):307–339.

Skare, M. and Buterin, V. (2015). Modelling and forecasting unemployment nonlinear dynamics using spectral analysis. *Engineering Economics*, 26(4):373–383.

Skare, M. and Porada-Rochoń, M. (2019). Tracking financial cycles in ten transitional economies 2005–2018 using singular spectrum analysis (ssa) techniques. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 14(1):7–29.

Škare, M. and Porada-Rochoń, M. (2020). The synchronisation between financial and business cycles: a cross spectral analysis view. *Technological and Economic Development of Economy*, pages 1–13.

Stock, J. H. and Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2):147–162.

Thomakos, D., Hassani, H., Patterson, K., et al. (2013). *Optimal Linear Filtering, Smoothing and Trend Extraction for the m-th Differences of a Unit Root Process: A Singular Spectrum Analysis Approach*. University of Reading, Department of Economics.

Thomakos, D. D. (2008). Optimal linear filtering, smoothing and trend extraction of m-period differences of processes with a unit root. *Smoothing and Trend Extraction of M-Period Differences of Processes with a Unit Root (November 5, 2008)*.

Thomakos, D. D. and Hassani, H. (2020). Using singular spectrum analysis for inference on seasonal time series with seasonal unit roots. *International Journal of Computational Economics and Econometrics*, 10(2):149–182.

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

Vautard, R. and Ghil, M. (1989). Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Physica D-Nonlinear Phenomena*, 35:395–424.

2.8 APPENDIX

2.8.1 Diagonalization of Grouped Matrices

In section 2.4.2, a brief explanation of the diagonalization procedure was made. In the appendix, we demonstrate the application of the Hankelization rule once the grouping of the SVD based matrices have been carried out. The SVD does not maintain the Hankel structure of the vertically stacked sub matrices in $\boldsymbol{X}$. Hence the matrices $\boldsymbol{X}_1$, $\boldsymbol{X}_2$,..., $\boldsymbol{X}_j$ need to be transformed to Hankel structures. Only once the matrices are diagonalized, we are able to use the inverse maping operator $\boldsymbol{J}_{MD}^{-1}$ so that the mapping of the matrices to time series domain can be achieved. To make matters clear, amongst any group matrix $\boldsymbol{X}_{I_s}$, let us take the first group matrix $\boldsymbol{X}_{I_1}$ formed by the leading 5 components such that; $I_1 = \{1, 2, 3, 4, 5\}$:

$$\boldsymbol{X}_{I_1} = \left[\begin{array}{cccccc}
x_{11}^{I_1} & x_{12}^{I_1} & x_{13}^{I_1} & \cdots & x_{1,K-1}^{I_1} & x_{1K}^{I_1} \\
x_{21}^{I_1} & x_{22}^{I_1} & x_{23}^{I_1} & \cdots & x_{2,K-1}^{I_1} & x_{2,2K}^{I_1} \\
x_{31}^{I_1} & x_{32}^{I_1} & x_{33}^{I_1} & \cdots & x_{3,K-1}^{I_1} & x_{3,2K}^{I_1} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
x_{L-1,1}^{I_1} & x_{L-1,2}^{I_1} & x_{L-1,3}^{I_1} & \cdots & x_{L-1,K-1}^{I_1} & x_{L-1,K}^{I_1} \\
x_{L1}^{I_1} & x_{L2}^{I_1} & x_{L3}^{I_1} & \cdots & x_{L,K-1}^{I_1} & x_{LK}^{I_1} \\
\hdashline
x_{L+1,1}^{I_1} & x_{L+1,2}^{I_1} & x_{L+1,3}^{I_1} & \cdots & x_{L+1,K-1}^{I_1} & x_{L+1,K}^{I_1} \\
x_{L+2,1}^{I_1} & x_{L+2,2}^{I_1} & x_{L+2,3}^{I_1} & \cdots & x_{2,K-1}^{I_1} & x_{L+2,K}^{I_1} \\
x_{L+3,1}^{I_1} & x_{L+3,2}^{I_1} & x_{L+3,3}^{I_1} & \cdots & x_{L+3,K-1}^{I_1} & x_{L+3,K}^{I_1} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
x_{L-1,1}^{I_1} & x_{L-1,2}^{I_1} & x_{L-1,3}^{I_1} & \cdots & x_{L-1,K-1}^{I_1} & x_{L-1,K}^{I_1} \\
x_{2L,1}^{I_1} & x_{2L,2}^{I_1} & x_{2L,3}^{I_1} & \cdots & x_{2L,K-1}^{I_1} & x_{2LK}^{I_1} \\
\hdashline
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\hdashline
x_{(p-1)L+1,1}^{I_1} & x_{(p-1)L+1,2}^{I_1} & x_{(p-1)L+1,3}^{I_1} & \cdots & x_{(p-1)L+1,K-1}^{I_1} & x_{(p-1)L+1,K}^{I_1} \\
x_{(p-1)L+2,1}^{I_1} & x_{(p-1)L+2,2}^{I_1} & x_{(p-1)L+2,3}^{I_1} & \cdots & x_{(p-1)L+2,K-1}^{I_1} & x_{(p-1)L+2,K}^{I_1} \\
x_{(p-1)L+3,1}^{I_1} & x_{(p-1)L+3,2}^{I_1} & x_{(p-1)L+3,3}^{I_1} & \cdots & x_{(p-1)L+3,K-1}^{I_1} & x_{(p-1)L+3,K}^{I_1} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
x_{L-1,1}^{I_1} & x_{L-1,2}^{I_1} & x_{L-1,3}^{I_1} & \cdots & x_{L-1,K-1}^{I_1} & x_{L-1,K}^{I_1} \\
x_{pL,1}^{I_1} & x_{pL,2}^{I_1} & x_{pL,3}^{I_1} & \cdots & x_{pL,K-1}^{I_1} & x_{pLK}^{I_1}
\end{array}\right] \tag{2.32}$$

For the first block in the above matrix, $L+K-2$ number of sequences $\{x_{21}^{I_1}, x_{12}^{I_1}\}$, $\{x_{31}^{I_1}, x_{22}^{I_1}, x_{13}^{I_1}\}$,...,$\{x_{L,K-1}^{I_1}, x_{L-1,K}^{I_1}\}$ are separately averaged along there respective anti-diagonals. Similarly in the second block, $L+K-2$ sequences $\{x_{L+2,1}^{I_1}, x_{L+1,2}^{I_1}\}$, $\{x_{L+3,1}^{I_1}, x_{L+2,2}^{I_1}, x_{L+1,3}^{I_1}\}$,...,$\{x_{2L,K-1}^{I_1}, x_{L-1,K}^{I_1}\}$. The averaging of these $L+K-2$ sequences are made at every block in $\boldsymbol{X}_{I_1}$ so that the vertically stacked matrices are diagonalized so that we have

$$\tilde{\boldsymbol{X}}_{I_1} = \boldsymbol{J}_v\left(\boldsymbol{X}_{I_1}\right) = \left[\begin{array}{cccccc}
\tilde{x}_{k=0}^{I_1(1)} & \tilde{x}_{k=1}^{I_1(1)} & \tilde{x}_{k=2}^{I_1(1)} & \cdots & \tilde{x}_{k=K-2}^{I_1(1)} & \tilde{x}_{k=K-1}^{I_1(1)} \\
\tilde{x}_{k=1}^{I_1(1)} & \tilde{x}_{k=2}^{I_1(1)} & \tilde{x}_{k=3}^{I_1(1)} & \cdots & \tilde{x}_{k=K-1}^{I_1(1)} & \tilde{x}_{k=K}^{I_1} \\
\tilde{x}_{k=2}^{I_1(1)} & \tilde{x}_{k=3}^{I_1(1)} & \tilde{x}_{k=4}^{I_1(1)} & \cdots & \tilde{x}_{k=K}^{I_1(1)} & \tilde{x}_{k=K+1}^{I_1(1)} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\tilde{x}_{k=L-2}^{I_1(1)} & \tilde{x}_{k=L-1}^{I_1(1)} & \tilde{x}_{k=L}^{I_1(1)} & \cdots & \tilde{x}_{k=L+K-4}^{I_1(1)} & x_{k=L+K-3}^{I_1} \\
\tilde{x}_{k=L-1}^{I_1(1)} & \tilde{x}_{k=L}^{I_1(1)} & \tilde{x}_{k=L+1}^{I_1(1)} & \cdots & \tilde{x}_{k=L+K-3}^{I_1(1)} & \tilde{x}_{k=L+K-2}^{I_1(1)} \\
\hline
\tilde{x}_{k=0}^{I_1(2)} & \tilde{x}_{k=1}^{I_1(2)} & \tilde{x}_{k=2}^{I_1(2)} & \cdots & \tilde{x}_{k=K-2}^{I_1(2)} & \tilde{x}_{k=K-1}^{I_1(2)} \\
\tilde{x}_{k=1}^{I_1(2)} & \tilde{x}_{k=2}^{I_1(2)} & \tilde{x}_{k=3}^{I_1(2)} & \cdots & \tilde{x}_{k=K-1}^{I_1(2)} & \tilde{x}_{k=K}^{I_1(2)} \\
\tilde{x}_{k=2}^{I_1(2)} & \tilde{x}_{k=3}^{I_1(2)} & \tilde{x}_{k=4}^{I_1(2)} & \cdots & \tilde{x}_{k=K}^{I_1(2)} & \tilde{x}_{k=K+1}^{I_1(2)} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\tilde{x}_{k=L-2}^{I_1(2)} & \tilde{x}_{k=L-1}^{I_1(2)} & \tilde{x}_{k=L}^{I_1(2)} & \cdots & \tilde{x}_{k=L+K-4}^{I_1(2)} & x_{k=L+K-3}^{I_1(2)} \\
\tilde{x}_{k=L-1}^{I_1(2)} & \tilde{x}_{k=L}^{I_1(2)} & \tilde{x}_{k=L+1}^{I_1(2)} & \cdots & \tilde{x}_{k=L+K-3}^{I_1(2)} & \tilde{x}_{k=L+K-2}^{I_1(2)} \\
\hline
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\hline
\tilde{x}_{k=0}^{I_1(p)} & \tilde{x}_{k=1}^{I_1(p)} & \tilde{x}_{k=2}^{I_1(p)} & \cdots & \tilde{x}_{k=K-2}^{I_1(p)} & \tilde{x}_{k=K-1}^{I_1(p)} \\
\tilde{x}_{k=1}^{I_1(p)} & \tilde{x}_{k=2}^{I_1(p)} & \tilde{x}_{k=3}^{I_1(p)} & \cdots & \tilde{x}_{k=K-1}^{I_1(p)} & \tilde{x}_{k=K}^{I_1(p)} \\
\tilde{x}_{k=2}^{I_1(p)} & \tilde{x}_{k=3}^{I_1(p)} & \tilde{x}_{k=4}^{I_1(p)} & \cdots & \tilde{x}_{k=K}^{I_1(p)} & \tilde{x}_{k=K+1}^{I_1(p)} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\tilde{x}_{k=L-2}^{I_1(p)} & \tilde{x}_{k=L-1}^{I_1(p)} & \tilde{x}_{k=L}^{I_1(p)} & \cdots & \tilde{x}_{k=L+K-4}^{I_1(p)} & x_{k=L+K-3}^{I_1(p)} \\
\tilde{x}_{k=L-1}^{I_1(p)} & \tilde{x}_{k=L}^{I_1(p)} & \tilde{x}_{k=L+1}^{I_1(p)} & \cdots & \tilde{x}_{k=L+K-3}^{I_1(p)} & \tilde{x}_{k=L+K-2}^{I_1(p)}
\end{array}\right] \tag{2.33}$$

where $\boldsymbol{J}_v$ is an orthogonal projection operator that diagonalizes the matrix. The elements along the skewed diagonals are computed simply by the following rule for the

$$\tilde{x}_{k=0}^{I_1(1)} = x_{11}^{I_1}$$
$$\tilde{x}_{k=1(1)}^{I_1} = \frac{x_{21}^{I_1} + x_{12}^{I_1}}{2}$$
$$\tilde{x}_{k=2}^{I_1} = \frac{x_{31}^{I_1(1)} + x_{22}^{I_1} + x_{13}^{I_1}}{3}$$
$$\vdots \quad = \quad \vdots$$
$$\tilde{x}_{k=L+K-3}^{I_1(1)} = \frac{x_{L,K-1}^{I_1} + x_{L-1,K}^{I_1}}{2}$$
$$\tilde{x}_{k=L+K-2}^{I_1(1)} = x_{LK}^{I_1}$$

The above rule is done also for the remaining $p-1$ sequences $\{x_{L+i,j}^{I_1}\}_{i,j=1}^{L,K}$,..., $\{x_{(p-1)L+i,j}^{I_1}\}_{i,j=1}^{L,K}$.

According to Golyandina et al. (2001), Hassani et al. (2009), the complete process of the reconstruction step is an optimal procedure. In the framework of these papers, any possible group matrix once diagonalized will follow the general formula that $\|\boldsymbol{X}_{I_s} - \tilde{\boldsymbol{X}}_{I_s}\|_{\boldsymbol{F}}$ leads to the minimal $\|\cdot\|_{\boldsymbol{F}}$ amongst all other Hankel matrices with dimensions $pL \times K$ and rank $r_{\boldsymbol{X}_{I_s}}$. In other words, the matrices $\tilde{\boldsymbol{X}}_{I_s}$'s are the closest in distance of all possible matrix computations with similar structure and lying in the same linear space to $\boldsymbol{X}_{I_s}$'s (matrices directly decomposed from the data matrix $\boldsymbol{X}$). The implication of this is that the SSA extracted components $\tilde{\mathbb{F}}_t^{(s)}$ achieve the most accurate approximation and the reconstruction step alters the data at minimal level attainable.

2.8.2 Euro Region Recession Timings

The full list of OECD based recession chronology of EU area business cycles can be obtained from OECD Composite Leading Indicators website: http://www.oecd.org/std/leading-indicators/. The indicator series can also be retrieved from FRED, Federal Reserve Bank of St. Louis website with ticker code EUROREC. The recession cycle period will then be identified by the succeeding month of the peak date plus the number succeeding and uninterrupted months the binary valued indicator is equal to unity.

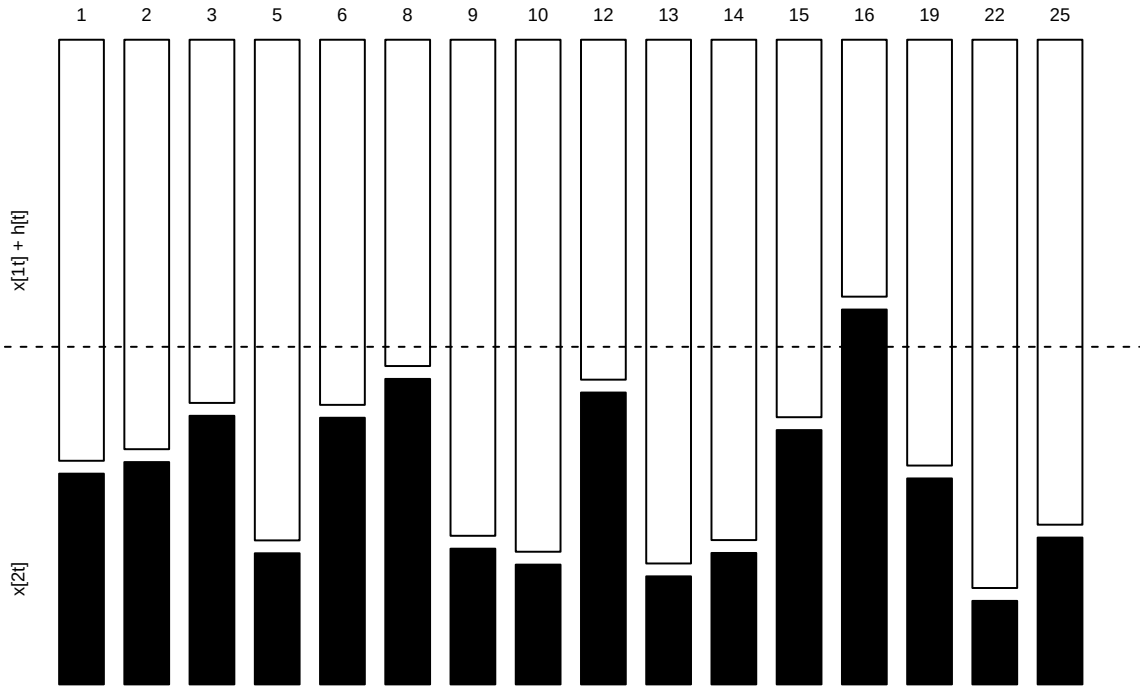
Table 2.7: Reference chronology of EU Recession Cycles

Turning Points;		
Peak	**Trough**	**Period**
08/1998	01/1999	16
02/2001	06/2003	25
02/2008	06/2009	16
05/2011	03/2013	22
11/2017	-	24

Note: Periods column are in terms of months. The final 3 points of the recession indicator are missing however the indicator cover the rest of the data period and beyond.

2.8.3 The effect of sample size and window length on filter resolution

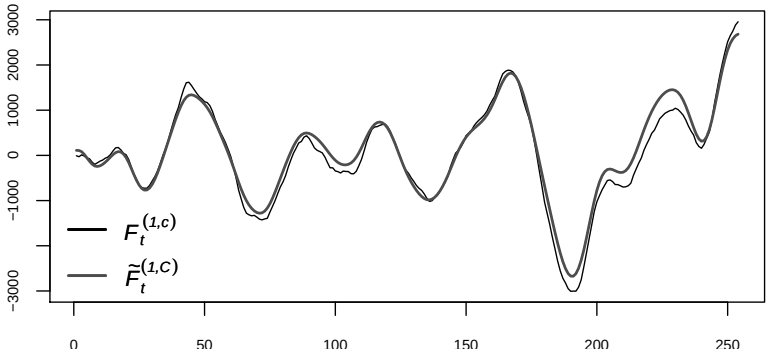
Figure 2.13: Estimated factor loadings across common component matrix ($\boldsymbol{X}_{\Psi^*_{BC}}$, $L = 120$)



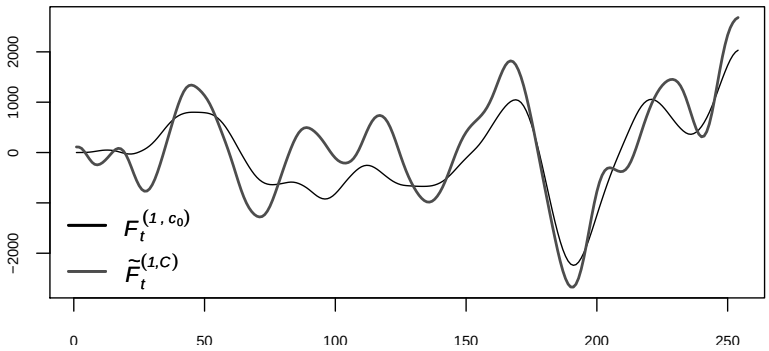
Note: Component loadings calculated from $\boldsymbol{X}^{(1)}_{\Psi^*_{BC}}$ related to coefficient estimates $\hat{a}^1_{z^*_i}$, $\hat{b}^1_{z^*_{h_i}}$ represented by □. Loadings of $\boldsymbol{X}^{(2)}_{\Psi^*_{BC}}$ related to coefficients estimates $\hat{a}^2_{z^*_i}$, $\hat{b}^2_{z^*_{h_i}}$ are represented by ■. See equation 2.4 for the representation and equation 2.6 for trigonometric identities.

Figure 2.14: Sensitivity of SSA-based filter, $L = 120$

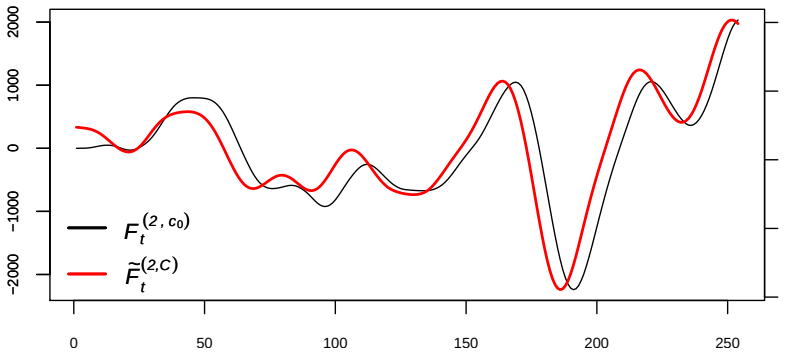
(a) $\mathbb{F}_t^{(1,c)}$ vs $\tilde{\mathbb{F}}_t^{(1,C)}$



(b) x_{1t}: $\left(\mathbb{F}_t^{(1,c_x)}\right)$ vs $\tilde{\mathbb{F}}_t^{(1,C)}$



(c) $\mathbb{F}_t^{(1,c_x)}$, $\mathbb{F}_t^{(2,c_x)}$ vs $\tilde{\mathbb{F}}_t^{(2,C)}$



Note: For panels (a) and (b): the grey line indicated by; — $\tilde{\mathbb{F}}_t^{(1,C)}$, is the SSA based business cycle estimate for $\mathbb{F}_t^{(1)}$. Black line in panel (a) indicates: — $\mathbb{F}_t^{(1,c)}$ which is the aggregated common and idiosyncratic cycles; $x_{1t} + g_t$. Black line in panel (b): indicates $\mathbb{F}_t^{(1,c_0)}$ is x_{1t} which is the common cycle for y1t. For panel (c): the red line indicated by; — $\tilde{\mathbb{F}}_t^{(2,C)}$, is the SSA based business cycle estimate for $\mathbb{F}_t^{(2)}$, black line represent: — $\tilde{\mathbb{F}}_t^{(1,c_0)}$, the blue line represents —$\mathbb{F}_t^{(2,c_0)}$ is the shifted common cycle. See equation 2.17 for the explicit forms.

2.8.4 State Space representation for UCM-1 and UCM-2

To obtain the smoothed filter estimates of the common cycle extracted from UCM-1 and UCM-2, we require the Kalman filter system matrices. These matrices form the State Space From (SSF) of structural equation system. The full details of the general expressions presented in this appendix can be found in Durbin and Koopman (2012), Pelagatti (2019) and Kitagawa (2020). We present only the necessary equations which will enable to compute the Kalman filter smoothed estimates of the states in the state vector $\boldsymbol{\alpha}_t$. Consider a 7 series data in stacked form $\boldsymbol{Y}_N = (\boldsymbol{y}_1', ..., \boldsymbol{y}_N')$ where $\boldsymbol{y}_1, ..., \boldsymbol{y}_N$ are each 7×1 time observation vectors. For any t, the time observation vectors can be estimated by the following UCM-1 in SSF form

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{Z}_t\boldsymbol{\alpha}_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim N(\boldsymbol{0}_{[7]}, \boldsymbol{H}_t) \\ \boldsymbol{\alpha}_{t+1} &= \boldsymbol{T}_t + \boldsymbol{R}_t\boldsymbol{\nu}_t \quad \boldsymbol{\nu}_t \sim N(\boldsymbol{0}_{[30]}, \boldsymbol{Q}_t) \end{aligned} \tag{2.34}$$

Equation 2.34 is composed of the measurement equation and the state update equation with the necessary non-unique system matrices. The system matrices build specifically for UCM-1 will be presented in compact form shortly. Before expressing each one, we write the set of Kalman filter recursions that will be required to obtain the Kalman smoothing equation of the state vector. These filter recursions are obtained based on satisfying conditions for the commonly known first two lemmas in multivariate normal regression and using linear expectation operators. The basis of the filters is by defining the one step ahead predictor of the state estimate $\hat{\boldsymbol{\alpha}}_{t|t-1} = \mathbb{E}(\boldsymbol{\alpha}_t|\boldsymbol{Y}_{t-1})$ and its variance $\boldsymbol{P}_{t|t-1} = \mathbb{V}(\boldsymbol{\alpha}_t|\boldsymbol{Y}_{t-1})$. Both these objects are in effect conditional mean and variances. The elements for the Kalman filter are then determined accordingly

$$\begin{aligned} \boldsymbol{\upsilon}_t &= \boldsymbol{y}_t - \mathbb{E}(\boldsymbol{y}_t|\boldsymbol{Y}_{t-1}) \\ &= \boldsymbol{y}_t - \boldsymbol{Z}_t\hat{\boldsymbol{\alpha}}_{t|t-1}, && (\textit{1-step ahead prediction error}) \\ \boldsymbol{F}_t &= \mathbb{V}(\boldsymbol{\upsilon}_t|\boldsymbol{Y}_{t-1}) \\ &= \boldsymbol{Z_t}\boldsymbol{P_{t|t-1}}\boldsymbol{Z_t'} + \boldsymbol{H_t}, && (\textit{1-step ahead prediction error variance}) \\ \boldsymbol{K}_t &= \boldsymbol{T_t}\boldsymbol{P_{t|t-1}}\boldsymbol{Z_t'}\boldsymbol{F_t^{-1}}, && (\textit{Kalman gain}) \\ \hat{\boldsymbol{\alpha}}_{t|t} &= \boldsymbol{T_t}\hat{\boldsymbol{\alpha}}_{t|t-1} + \boldsymbol{K_t}\boldsymbol{\upsilon_t}, && (\textit{filtered estimator of } \boldsymbol{\alpha_t}) \\ \boldsymbol{P}_{t|t} &= \boldsymbol{P}_{t|t-1} - \boldsymbol{P_{t|t-1}}\boldsymbol{Z_t'}\boldsymbol{F_t^{-1}}\boldsymbol{Z_t}\boldsymbol{P_{t|t-1}}, && (\textit{variance-covariance matrix of } \boldsymbol{\alpha_t}) \end{aligned} \tag{2.35}$$

where the matrix defined as the Kalman gain is the improvement at time t made on the 1-step ahead prediction once the observation set $\boldsymbol{y}_t$ is observed. Recursions in equation 2.35 allows for the computation of the prediction steps

$$\begin{aligned} \hat{\boldsymbol{\alpha}}_{t+1|t} &= \boldsymbol{T_t}\hat{\boldsymbol{\alpha}}_{t+1|t} + \boldsymbol{K_t}\boldsymbol{\upsilon_t} \\ \boldsymbol{P}_{t+1|t} &= \boldsymbol{T_t}\boldsymbol{P}_{t|t}\boldsymbol{T_t'} + \boldsymbol{R_t}\boldsymbol{Q_t}\boldsymbol{R_t'} - \boldsymbol{K_t}\boldsymbol{F_t}\boldsymbol{K_t'} \end{aligned} \tag{2.36}$$

The prediction step equations 2.36, also known as forward recursions are necessary for obtaining the backward recursions of the the Kalman filter smoothed

estimates of the state vector $\hat{\boldsymbol{\alpha}}_{\boldsymbol{t|N}} = \mathbb{E}(\boldsymbol{\alpha}_t|\boldsymbol{Y}_N)$, where the state vector is conditioned on the full data sample such that $N > t$. The backward recursive elements to compute the smoothed estimate of the signal ($\hat{\boldsymbol{\alpha}}_{\boldsymbol{t|N}}$) will be

$$\begin{aligned} \boldsymbol{r}_{t-1} &= \boldsymbol{Z_t F_t^{-1} v_t} + (\boldsymbol{T_t - K_t Z_t})' \boldsymbol{r_t} \\ \boldsymbol{r}_N &= 0 \\ \hat{\boldsymbol{\alpha}}_{\boldsymbol{t|N}} &= \hat{\boldsymbol{\alpha}}_{\boldsymbol{t|t-1}} + \boldsymbol{P_{t|t-1} r}_{t-1} \end{aligned} \tag{2.37}$$

The final line of equation 2.37 allows for the calculation of the smoothed estimate of the Kalman filter based structural econometric model that has been applied.

The initialization of the Kalman filter is based on certain distributional assumptions going forward. Mainly, the initial state vector distribution and the joint distribution of the model. These assumptions are a further additional requirements needed to be able to run the equations in 2.35. The distributional assumptions are presented in general Kalman filter terminology as follows

$$\begin{aligned} &\boldsymbol{\alpha_1} \sim N(\hat{\boldsymbol{\alpha}}_{\boldsymbol{1|0}}, \boldsymbol{P_{1|0}}), \quad \textit{(initial state distribution)} \\ &\textbf{model implied distributions:} \\ &\boldsymbol{y_t}|\boldsymbol{Y_{t-1}} \sim N(\boldsymbol{Z_t}\hat{\boldsymbol{\alpha}}_{\boldsymbol{t|t-1}}, \boldsymbol{F_t}), \quad \textit{(measurement equation implied)} \\ &\boldsymbol{\alpha_t}|\boldsymbol{Y_{t-1}} \sim N(\hat{\boldsymbol{\alpha}}_{\boldsymbol{t|t-1}}, \boldsymbol{P_{t|t-1}}), \quad \textit{(state update equation implied)} \end{aligned} \tag{2.38}$$

where $\hat{\boldsymbol{\alpha}}_{\boldsymbol{1|0}} = \mathbf{0}_{\boldsymbol{[30]}}$ is a column vector of zeros indicating every state vector initial value at $t = 1$ is set to 0. Similarly, $\boldsymbol{P_{1|0}}$ is a symmetric matrix of dimensions 30×30 and is the initial variance covariance matrix of the one step ahead prediction of $\boldsymbol{\alpha_t}$. The matrix will be explicitly presented together with the complete set of system matrices shortly. A final note on the model implied distribution assumption, will be that, it should essentially be seen as a multivariate joint distribution which can be written more compactly in the form of

$$\begin{bmatrix} \boldsymbol{y_t} \\ \boldsymbol{\alpha_t} \end{bmatrix} |\boldsymbol{Y_{t-1}} \sim N\left(\begin{bmatrix} \hat{\boldsymbol{y}}_{\boldsymbol{tt-1}} \\ \hat{\boldsymbol{\alpha}}_{\boldsymbol{t|t-1}} \end{bmatrix}, \begin{bmatrix} \boldsymbol{F_t} & \boldsymbol{Z_t P_{t|t-1}} \\ \boldsymbol{P_{t|t-1} Z_t'} & \boldsymbol{P_{t|t-1}} \end{bmatrix} \right). \tag{2.39}$$

The system matrices for the UCM-1 model associated to the SSF of equation 2.34 is presented as follows

System matrices for UCM-1

We first collect the state variables that form the state vector and the update equation

$\boldsymbol{\alpha_t} = (\boldsymbol{\alpha_t^{(\mu,\eta)}}\ \boldsymbol{\alpha_t^{\psi}}\ \boldsymbol{\alpha_t^{\varphi}})'$

$\boldsymbol{\alpha^{(\mu,\eta)}} = (\mu_{1t}\ \eta_{1t}\ \mu_{2t}\ \eta_{2t}\ \mu_{3t}\ \eta_{3t}\ \mu_{4t}\ \eta_{4t}\ \mu_{5t}\ \eta_{5t}\ \mu_{6t}\ \eta_{6t}\ \mu_{7t}\ \eta_{7t})'$

$\boldsymbol{\alpha^{\psi}} = (\psi_t\ \psi_t^*)'$

$$\boldsymbol{\alpha^{\varphi}} = (\varphi_{1t}\ \varphi^{*}_{1t}\ \varphi_{2t}\ \varphi^{*}_{2t}\ \varphi_{3t}\ \varphi^{*}_{3t}\ \varphi_{4t}\ \varphi^{*}_{4t}\ \varphi_{5t}\ \varphi^{*}_{5t}\ \varphi_{6t}\ \varphi^{*}_{6t}\ \varphi_{7t}\ \psi^{*}_{7t})'$$

The state transition matrix and its block matrix components are defined in the following forms

$$\boldsymbol{T_t} = \begin{bmatrix} \boldsymbol{T_t^{(\mu,\eta)}} & \boldsymbol{0_{[14\times16]}} \\ \boldsymbol{0_{[16\times14]}} & \boldsymbol{I_{[8]}} \otimes \boldsymbol{T_t^{\psi}} \end{bmatrix}$$

$$\boldsymbol{T_t^{(\mu,\eta)}} = \boldsymbol{I_{[7]}} \otimes \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$\boldsymbol{T_t^{\psi}}$ is the rotation matrix presented in equation 2.27.

The vector containing the error terms of the system of transition equation is stacked vertically so that

$$\boldsymbol{\nu_t} = (\boldsymbol{\nu_t^{\xi}}\ \boldsymbol{\nu_t^{\kappa^{\dagger}}}\ \boldsymbol{\nu_t^{\kappa}})'$$

$$\boldsymbol{\nu_t^{\xi}} = (\xi_{1t}\ \xi_{2t}\ \xi_{3t}\ \xi_{4t}\ \xi_{5t}\ \xi_{6t}\ \xi_{7t})'$$

$$\boldsymbol{\nu_t^{\kappa^{\dagger}}} = (\kappa^{\dagger}\ \kappa^{\dagger *})'$$

$$\boldsymbol{\nu_t^{\kappa}} = (\kappa_{1t}\ \kappa^{*}_{1t}\ \kappa_{2t}\ \kappa^{*}_{2t}\ \kappa_{3t}\ \kappa^{*}_{3t}\ \kappa_{4t}\ \kappa^{*}_{4t}\ \kappa_{5t}\ \kappa^{*}_{5t}\ \kappa_{6t}\ \kappa^{*}_{6t}\ \kappa_{7t}\ \kappa^{*}_{7t})'$$

The $\boldsymbol{R_t}$ matrix designates the state errors in the vector $\boldsymbol{\nu_t}$ to the correct state update equations. The matrix results in $\boldsymbol{\nu_t}$ to be parsimonious, such that no single error term within the vector is degenerate. The $\boldsymbol{R_t}$ structure for the model is

$$\boldsymbol{R_t} = \left[\boldsymbol{I_{[7]}} \otimes \begin{pmatrix} 1 & 0 \end{pmatrix}'\right] \oplus \boldsymbol{I_{[16]}}$$

The state error variance-covariance matrix which is simply $\boldsymbol{Q_t} = \mathbb{E}(\boldsymbol{\nu_t}\boldsymbol{\nu_t'})$ has the following representational form

$$\boldsymbol{Q_t} = Blockdiag[\boldsymbol{\Lambda^{\xi}}\ \boldsymbol{\Lambda^{\kappa^{\dagger}}}\ \boldsymbol{\Lambda^{\kappa}}]$$

$$\boldsymbol{\Lambda^{\xi}} = diag(\sigma^2_{\xi_1}\ \sigma^2_{\xi_2}\ \sigma^2_{\xi_3}\ \sigma^2_{\xi_4}\ \sigma^2_{\xi_5}\ \sigma^2_{\xi_6}\ \sigma^2_{\xi_7}\ )$$

$$\boldsymbol{\Lambda^{\kappa^{\dagger}}} = \sigma^2_{\kappa^{\dagger}}\boldsymbol{I_{[2]}}$$

$$\boldsymbol{\Lambda^{\kappa}} = diag(\sigma^2_{\kappa_1}\ \sigma^2_{\kappa_1}\ \sigma^2_{\kappa_2}\ \sigma^2_{\kappa_2}\ \sigma^2_{\kappa_3}\ \sigma^2_{\kappa_3}\ \sigma^2_{\kappa_4}\ \sigma^2_{\kappa_4}\ \sigma^2_{\kappa_5}\ \sigma^2_{\kappa_5}\ \sigma^2_{\kappa_6}\ \sigma^2_{\kappa_6}\ \sigma^2_{\kappa_7}\ \sigma^2_{\kappa_7})$$

Elements related to the measurement equation design matrix is presented in the following manner

$$\boldsymbol{Z_t} = \left[\boldsymbol{I_{[7]}} \otimes \begin{pmatrix} 1 & 0 \end{pmatrix}\ \ \boldsymbol{\Xi}\ \ \boldsymbol{I_{[7]}} \otimes \begin{pmatrix} 1 & 0 \end{pmatrix}\right]$$

where $\boldsymbol{\Xi}$ is the coefficient matrix in equation 2.29 The observational error in the measurement equation (top line in 2.34) was assumed to be jointly normally distributed with zero unconditional means and a covariance structure being a diagonal matrix: $\mathbb{E}(bm\epsilon_t bm\epsilon_t') = \boldsymbol{H_t}$.

The initialization of the Kalman filter for the state vector estimate and covariance matrix of the prediction error; $\hat{\boldsymbol{\alpha}}_{1|0}, \boldsymbol{P}_{1|0}$ is

$$\hat{\boldsymbol{\alpha}}_{\mathbf{1|0}} = \mathbf{0}'_{\mathbf{[30]}}$$

$$\boldsymbol{P}_{1|0} = Blockdiag\left[\tau \boldsymbol{I}_{\mathbf{[14]}} \quad \boldsymbol{\Gamma_\psi(0)} \quad \boldsymbol{\Gamma_\varphi(0)}\right]$$

where $\tau \longrightarrow \infty$ is constant set for the diffuse priors of non-stationary states, $\boldsymbol{\Gamma_\psi(k)}$ and $\boldsymbol{\Gamma_\phi(k)}$ are expressions obtained from 2.28 with $k = 0$ and $\boldsymbol{\theta}_0$ indicating the initial values for the hyper parameter vector $\boldsymbol{\theta}$ $=(\boldsymbol{\sigma_\epsilon} \quad \sigma_{\kappa^\dagger} \quad \boldsymbol{\gamma} \quad \boldsymbol{\sigma_\kappa})'$ with $\boldsymbol{\sigma_\epsilon}$ $=(\sigma_{\epsilon_1} \ \sigma_{\epsilon_2} \ \sigma_{\epsilon_3} \ \sigma_{\epsilon_4} \ \sigma_{\epsilon_5} \ \sigma_{\epsilon_6} \ \sigma_{\epsilon_7})'$, $\boldsymbol{\sigma_\kappa}$ $=(\sigma_{\kappa_1} \ \sigma_{\kappa_2} \ \sigma_{\kappa_3} \ \sigma_{\kappa_4} \ \sigma_{\kappa_5} \ \sigma_{\kappa_6} \ \sigma_{\kappa_7})'$

System matrices for UCM-2

$$\boldsymbol{\alpha_t} = (\boldsymbol{\alpha_t^{(\mu,\eta)}} \ \boldsymbol{\alpha_t^{\varphi}} \ \boldsymbol{\alpha_t^{\psi(2)}})'$$

$$\boldsymbol{\alpha^{(\mu,\eta)}} = (\mu_{1t} \ \eta_{1t} \ \mu_{2t} \ \eta_{2t} \ \mu_{3t} \ \eta_{3t} \ \mu_{4t} \ \eta_{4t} \ \mu_{5t} \ \eta_{5t} \ \mu_{6t} \ \eta_{6t} \ \mu_{7t} \ \eta_{7t})'$$

$$\boldsymbol{\alpha^{\varphi}} = (\varphi_{1t} \ \varphi_{1t}^* \ \varphi_{2t} \ \varphi_{2t}^* \ \varphi_{3t} \ \varphi_{3t}^* \ \varphi_{4t} \ \varphi_{4t}^* \ \varphi_{5t} \ \varphi_{5t}^* \ \varphi_{6t} \ \varphi_{6t}^* \ \varphi_{7t} \ \varphi_{7t}^*)'$$

$$\boldsymbol{\alpha^{\psi}} = (\psi_t^{(2)} \ \psi_t^{(2)*} \ \psi_t^{(1)} \ \psi_t^{(1)*})'$$

$$\boldsymbol{T_t} = \begin{bmatrix} \boldsymbol{T_t^{(\mu,\,\eta)}} & \mathbf{0_{[14\times18]}} \\ \mathbf{0_{[18\times14]}} & \boldsymbol{T_t^{(\phi,\psi^{[2]})}} \end{bmatrix}$$

$$\boldsymbol{T_t^{(\phi,\psi^{(2)})}} = Blockdiag\left[\boldsymbol{I_{[7]}} \otimes \boldsymbol{T_t^{\psi}} \quad \boldsymbol{T_t^{\psi(2)}}\right], \quad \boldsymbol{T_t^{\psi(2)}} = \begin{bmatrix} \rho \boldsymbol{T_t^{\psi}} & \boldsymbol{I_{[2]}} \\ \mathbf{0_{[2]}} & \rho \boldsymbol{T_t^{\psi}} \end{bmatrix}$$

$$\boldsymbol{\nu_t} = (\boldsymbol{\nu_t^{\xi}} \ \boldsymbol{\nu_t^{\kappa}} \ \boldsymbol{\nu_t^{\kappa^\dagger}})'$$

$$\boldsymbol{\nu_t^{\xi}} = (\xi_{1t} \ \xi_{2t} \ \xi_{3t} \ \xi_{4t} \ \xi_{5t} \ \xi_{6t} \ \xi_{7t})'$$

$$\boldsymbol{\nu_t^{\kappa}} = (\kappa_{1t} \ \kappa_{1t}^* \ \kappa_{2t} \ \kappa_{2t}^* \ \kappa_{3t} \ \kappa_{3t}^* \ \kappa_{4t} \ \kappa_{4t}^* \ \kappa_{5t} \ \kappa_{5t}^* \ \kappa_{6t} \ \kappa_{6t}^* \ \kappa_{7t} \ \kappa_{7t}^*)'$$

$$\boldsymbol{\nu_t^{\kappa^\dagger}} = (\kappa^\dagger \ \kappa^{\dagger *})'$$

$$\boldsymbol{Z_t} = \left[\boldsymbol{I_{[7]}} \otimes \begin{pmatrix} 1 & 0 \end{pmatrix} \quad \boldsymbol{I_{[7]}} \otimes \begin{pmatrix} 1 & 0 \end{pmatrix} \quad \boldsymbol{\Xi} \quad \mathbf{0_{[7\times2]}}\right]$$

$$\boldsymbol{Q_t} = Blockdiag[\boldsymbol{\Lambda^{\xi}}\ \boldsymbol{\Lambda^{\kappa}}\ \boldsymbol{\Lambda^{\kappa^{\dagger}}}]$$

$$\boldsymbol{H_t} = \boldsymbol{\Lambda^{\epsilon}}$$

$$\boldsymbol{R_t} = Blockdiag\left[\boldsymbol{I_{[7]}} \otimes \begin{pmatrix} 0 & 1 \end{pmatrix}' \quad \boldsymbol{I_{[14]}}\right] \oplus \boldsymbol{R_t^{\psi^{(2)}}}, \quad \boldsymbol{R_t^{\psi^{(2)}}} = \begin{bmatrix} \mathbf{0_{[2\times 2]}} \\ \boldsymbol{I_{[2]}} \end{bmatrix}$$

$$\boldsymbol{\hat{\alpha}_{1|0}} = \mathbf{0'_{[32]}}$$

$$\boldsymbol{P}_{1|0} = Blockdiag\left[\tau \boldsymbol{I_{[14]}} \quad \boldsymbol{\Gamma_{\varphi}(0)}\right] \oplus \boldsymbol{\Gamma_{\psi^{(2)}}(0)}$$

where $\boldsymbol{\Gamma_{\psi^{(2)}}(k)}$ is obtained from equation 2.31 with $k = 0$. Hyper parameter vector ($\boldsymbol{\theta}$) and its initial vector ($\boldsymbol{\theta_0}$) are identical to UCM-1.