

SDS-16

A Robust-filtering Method for Noisy  
Non-Stationary Multivariate Time  
Series with Econometric Applications

Naoto Kunitomo and Seisho Sato

October 2020

Statistics & Data Science Series back numbers:  
<http://www.mims.meiji.ac.jp/publications/datascience.html>

# A Robust-filtering Method for Noisy Non-Stationary Multivariate Time Series with Econometric Applications \*

Naoto Kunitomo <sup>†</sup>

and

Seisho Sato <sup>‡</sup>

October 16, 2020  
(A Revised Version)

## Abstract

We investigate a new filtering method to estimate the hidden states of random variables for multiple non-stationary time series data. This helps in analyzing small sample non-stationary macro-economic time series in particular and it is based on the frequency domain application of the separating information maximum likelihood (SIML) method, developed by Kunitomo, Sato and Kurisu (2018), and Kunitomo, Awaya and Kurisu (2019), and Nishimura, Sato and Takahashi (2019). We solve the filtering problem of hidden random variables of trend-cycle, seasonal and measurement-errors components, and propose a method to handle macro-economic time series. We develop the asymptotic theory based on the frequency domain analysis for non-stationary time series. We illustrate applications, including some properties of the method of Müller and Watson (2018), and analyses of some macro-economic data in Japan.

---

\*This is a revised version of the paper (SDS-9, MIMS=RSP Statistics and Data Science Series, Meiji University) presented at the conference "Theory and Applications of Economic and Government Statistics 2019" held at University of Tokyo on January 31, 2019. We thank Akihiko Takahashi for giving a valuable suggestion on the SIML filtering method at the initial stage of this work.

<sup>†</sup>The corresponding author, School of Political Science and Economics, Meiji University, Kanda-Surugadai 1-1 (Sarugakucho 3rd Building C-106), Chiyoda-ku, Tokyo 101-8301, naoto.kunitomo@gmail.com

<sup>‡</sup>Graduate School of Economics, University of Tokyo, Bunkyo-ku, Hongo 7-3-1, Tokyo 113-0033, seisho@e.u-tokyo.ac.jp

## Key Words

Noisy Non-stationary time series, Errors-variables models, Measurement Error, trend-cycle, seasonality and noise, Robust-filtering, SIML, Real-Fourier-Inversion, Müller-Watson method, Macro-consumption in Japan

# 1 Introduction

There exists vast research on the use of statistical time series analysis for macro-economic time series. One important feature of macroeconomic time series, which is different from standard statistical time series analysis, is that the observed time series is an apparent mixture of non-stationary and stationary components. The second feature is that the measurement errors in economic time series play important roles, because macro-economic data are usually constructed from various sources, including sample surveys in major official statistics, while the statistical time series analysis often ignored measurement errors. Third, the sample size of macro-economic data is rather small, and we have about 120 time series observations for each series for quarterly data over 30 years. The quarterly GDP series, which is the most important data in the Japanese macro-economy, for instance, is regularly estimated and released by the cabinet office of Japan <sup>1</sup>. Fourth, to publish the seasonally adjusted data, the official agencies usually apply the X-12-ARIMA program, which uses the univariate reg-ARIMA model to remove the seasonality as the standard filtering procedure. As the sample size is small, it is important to use an appropriate statistical procedure to extract information on trend-cycle, seasonality and noise (or measurement error) components systematically from multiple time series data.

In this study, we investigate a new filtering procedure to estimate the hidden states of trend-cycle, which are non-stationary, and to handle multiple time series data, including small sample time series. Kunitomo and Sato (2017), Kunitomo, Sato and Kurisu (2018), and Kunitomo, Awaya and Kurisu (2019) have developed the separating information maximum likelihood (SIML) method for estimating the non-stationary errors-in-variables models. They have discussed the asymptotic and finite sample properties of the estimation of unknown parameters in the statistical models. We utilize their results to solve the filtering problem of hidden random variables and show that they lead to new a way to handle macro-economic time series.

Related literature on the non-stationary economic time series analysis are Engle and Granger (1987) and Johansen (1995), which examined multivariate non-stationary and stationary time series and developed the notion of co-integration without measurement errors. Our problem is related to their work, but it has different aspects, and our focus is on the non-stationary trend-cycle, seasonality and measurement error in the non-stationary errors-in-variable models and their frequency domain analysis. Some related econometric studies on time series in the

---

<sup>1</sup>Currently, GDP time series data in Japan are constructed and seasonally adjusted since 1994Q1, although some may think the historical data before 1994 was constructed and measured exactly in the same way. The measuring procedures of official GDP series has been changed several times including the base-year changes. See <https://www.esri.cao.go.jp/index-e.html> for the detailed explanation of how GDP in Japan is constructed.

frequency domain are Baxter and King (1999), Christiano and Fitzgerald (2003) and Müller and Watson (2018). See Yamada (2019) for a survey of related studies, including the well-known Hodrick-Prescot (HP) filter in econometrics.

In statistical multivariate analysis, some studies on the errors-in-variables models are Anderson (1984, 2003) and Fuller (1987); however, they considered multivariate cases of independent observations, and the underlying situation is different from ours.

For statistical filtering methods, Kitagawa (2010) discussed the standard statistical methods already known, including the Kalman-filtering and particle-filtering methods. Although many studies have examined statistical filtering theories, we must exercise caution in analyzing non-stationary multivariate economic time series. See Granger and Hatanaka (1964), Brillinger and Hatanaka (1969) on early studies, and Harvey and Trimbur (2008) on the relationship between HP filter and other methods, for instance. Here we should mention two issues. First, the existing methods often depend on the underlying distributions such as the Gaussian distributions for the Kalman-filtering, and second, they often depend on the dimension of state variables. There may be some difficulty in extending the existing methods to high-dimension cases, even when the dimension is about 10. On the other hand, we expect that our method has robustness properties when we handle small sample economic times series with non-stationary trend-cycle, and stationary seasonality and measurement errors because our method does not depend on the specific distribution as well as the dimension of the underlying random variables. See Kunitomo, Awaya and Kurisu (2017) for a comparison of small sample properties of the ML (maximum likelihood) and SIML methods for the non-stationary errors-in-variables models, and Nishimura, Sato and Takahashi (2019) for an application of financial data smoothing. The most important feature of the present procedure is that it may be applicable to small sample time series data with non-stationary trend-cycle, seasonal and noise components and it has a statistical foundation based on the (real-valued) spectral decomposition of stochastic processes by a (real-valued) Fourier transformation, as we shall explain in Sections 4 and 5.

In Section 2, we give some macro-economic data, which have motivated this present study. In Section 3, we define the non-stationary errors-in-variables models and the SIML method. In Section 4, we introduce the SIML filtering method. In Section 5, we give the statistical foundation of the method and in Section 6, we discuss the problem of choosing the number of orthogonal processes, and give some numerical examples based on simulation and data. Section 7 contains some applications including an interpretation on the Müller-Watson method in econometrics and gives two empirical applications of macro-consumption in Japan. Concluding remarks are given in Section 8. The Appendix contains some mathematical derivations of our results and figures.

## 2 Two Illustrative Examples

In the first illustrative example, we plot the graph of two macro-economic time series in Japan: quarterly (real) consumption and quarterly (real) GDP (1994Q1-2018Q2) as Figure 1 <sup>2</sup>. It looks like a simple example of linear regression in undergraduate textbooks. However, if we draw the time series sequences of these (original and seasonally unadjusted official) macro-data estimated by the Cabinet office of the Japanese Government as Figure 2, we find things to be a little more complex than in Figure 1. There are clear trend-cycle components, seasonal fluctuations, and noise components in two time series data. Although many economists usually use the seasonally adjusted (published) data, which were constructed by using the X-12-ARIMA program in different ministries within the government, the effects of filtering in the program are often unknown. The X-12-ARIMA program uses the univariate reg-ARIMA model, which is a mixture of univariate seasonal ARIMA and linear regression, and it decomposes univariate time series into the trend-cycle, seasonality and noise components as the standard procedure <sup>3</sup>. In contrast, the DECOMP program, which is explained by Kitagawa (2010), is a possible choice particularly in Japan, uses the univariate AR model and the Kalman filtering technique with AIC, which is based on the Gaussian likelihood. When each time series is handled using different filtering procedures (i.e., different ARIMA models or reg-ARIMA models for instance), it may cause a fundamental problem in their interpretation when the focus is on the relationships among different non-stationary time series.

Figure 3 gives three different macro-consumption data (2002 January - 2016 December), which are observed as monthly time series and widely used by economists in Japan to judge the current macro-business condition. The first series is *Kei-Chosa* (the data from monthly consumer-survey collected by Statistics Bureau, Ministry of Internal Affairs and Communications), the second is *Shougyo-Doutai-Statistics* (the data from monthly retail constructed by Ministry of Economy, Trade and Industry (METI)). and the third is *Dai-Sanji-Sangyo-Statistics* (the index data on commerce constructed by METI). We note that the data construction processes of these series based on sample surveys are complex in different ministries within the government, and each data reflects different aspects of macro-consumption. Although they show similar movements, we observe some differences in trend-cycle, seasonality and noises. Then it may be desirable to unify the monthly consumption series because we want to judge the business condition each month by just observing these data to evaluate the state of the Japanese macro-economy and forming macro-economic policy. Many economists in both governments and private sectors usually use seasonally adjusted data, which were constructed from the quarterly or monthly (original) time series via the univariate X-12-ARIMA seasonal adjustment program.

---

<sup>2</sup>In Japan, both the original quarterly series and the seasonally adjusted series of GDP and its major components are regularly published. It differs from macro-data in the U.S. in some aspect.

<sup>3</sup>See Findley et.al. (1998) for the detail of X-12-ARIMA program.

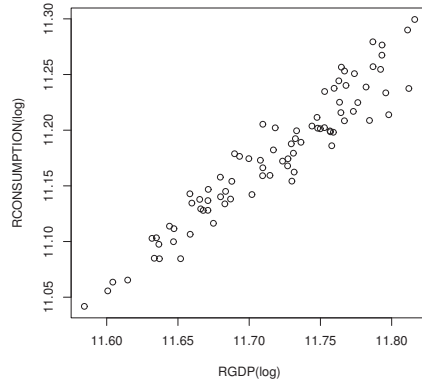


Figure 1: **GDP vs. Consumption.** (Data are the Quarterly real GDP and real-consumption between 1994Q1-2014Q3, which were published in 2015 by the Economic Social Research Institute (ESRI), Cabinet Office, Japan.)

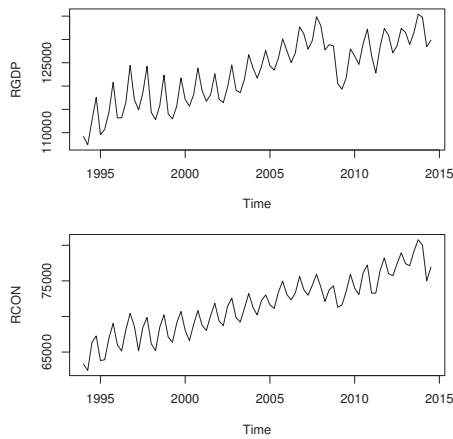


Figure 2: **GDP and Consumption.** (Data are the Quarterly real GDP and real-consumption between 1994Q1-2014Q3, which were published in 2015 by the Economic Social Research Institute (ESRI), Cabinet Office, Japan.)

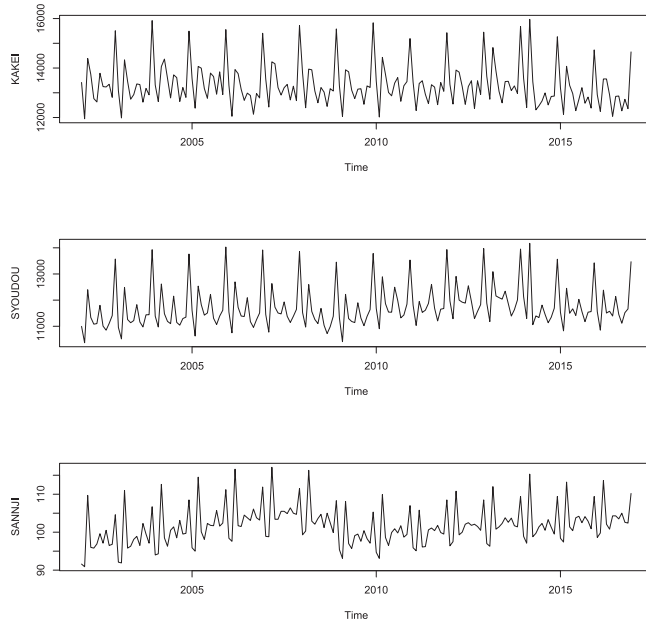


Figure 3: **Monthly Consumption Series.** (Data are the monthly consumptions between 2002M1-2016M12, which were published in 2017 by the Statistics Bureau and Ministry of Economy, Trade and Industry (METI), Japan.)

It is important to construct the monthly consumption index, which is consistent with the published quarterly macro-consumption data, which usually reported with substantial time lags. It was one of motivations to develop our filtering theory.

Some econometricians use the multivariate (parametric) time series models such as VAR (vector autoregressive process) for analyzing macro-economic data and investigating the relationships among them. They may use the seasonally adjusted (official) data, but we need caution to use such data because most published official data are already filtered by the X-12-ARIMA program. When the dimension is more than 2, some difficulties handling trend-cycle, seasonal, and measurement errors at the same time typically arise. A need to handle macro-economic data in simple non-parametric way motivated us to develop the multivariate non-stationary errors-in-variables models and the filtering method for the hidden state variables with measurement errors in Sections 3 and 4.



### 3 Non-Stationary Errors-in-Variables Model and SIML

#### 3.1 A Simple Non-Stationary Errors-in-Variables Model and the SIML Method

In this sub-section we first introduce a simple non-stationary errors-in-variables model with trend and noise components, and explain the SIML method. Then in the next subsection we shall investigate the general framework for non-stationary multivariate time series with trend-cycle, seasonal and measurement-error components.

Let  $y_{ji}$  be the  $i$ -th observation of the  $j$ -th time series at  $i$  for  $i = 1, \dots, n; j = 1, \dots, p$ . We set  $\mathbf{y}_i = (y_{1i}, \dots, y_{pi})'$  be a  $p \times 1$  vector and  $\mathbf{Y}_n = (\mathbf{y}'_i) (= (y_{ij}))$  be an  $n \times p$  matrix of observations and denote  $\mathbf{y}_0$  as the initial  $p \times 1$  vector and it is fixed. We consider the simple model that the underlying non-stationary trend is  $\mathbf{x}'_i (= (x_{1i}, \dots, x_{pi}), i = 1, \dots, n)$  and the noise component is  $\mathbf{v}'_i = (v_{1i}, \dots, v_{pi})$ , which is independent of  $\mathbf{x}_i$ . We write

$$(3.1) \quad \mathbf{y}_i = \mathbf{x}_i + \mathbf{v}_i \quad (i = 1, \dots, n).$$

When each pair of vectors  $\Delta \mathbf{x}_i$  and  $\mathbf{v}_i$  are independently, identically, and normally distributed (i.i.d.) as  $N_p(\mathbf{0}, \Sigma_x)$  and  $N_p(\mathbf{0}, \Sigma_v)$ , respectively, we have the observations of an  $n \times p$  matrix  $\mathbf{Y}_n = (\mathbf{y}'_i)$  and set the  $np \times 1$  random vector  $(\mathbf{y}'_1, \dots, \mathbf{y}'_n)'$ . Given the initial condition  $\mathbf{y}_0 (= \mathbf{x}_0)$ , we have

$$(3.2) \quad \text{vec}(\mathbf{Y}_n) \sim N_{n \times p} \left( \mathbf{1}_n \cdot \mathbf{y}'_0, \mathbf{I}_n \otimes \Sigma_v + \mathbf{C}_n \mathbf{C}'_n \otimes \Sigma_x \right),$$

where  $\mathbf{1}'_n = (1, \dots, 1)$  and

$$(3.3) \quad \mathbf{C}_n = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 1 & 1 & \dots & 0 \\ 1 & \dots & 1 & 1 & 0 \\ 1 & \dots & 1 & 1 & 1 \end{pmatrix}_{n \times n}.$$

We use the  $\mathbf{K}_n$ -transformation that from  $\mathbf{Y}_n$  to  $\mathbf{Z}_n (= (\mathbf{z}'_k))$  by

$$(3.4) \quad \mathbf{Z}_n = \mathbf{K}_n^* (\mathbf{Y}_n - \bar{\mathbf{Y}}_0), \quad \mathbf{K}_n^* = \mathbf{P}_n \mathbf{C}_n^{-1},$$

where  $\bar{\mathbf{Y}}_0 = \mathbf{1}_n \cdot \mathbf{y}'_0$ ,

$$(3.5) \quad \mathbf{C}_n^{-1} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & 0 & \dots \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}_{n \times n},$$

and

$$(3.6) \quad \mathbf{P}_n = (p_{jk}^{(n)}), p_{jk}^{(n)} = \sqrt{\frac{2}{n + \frac{1}{2}}} \cos \left[ \frac{2\pi}{2n + 1} \left(k - \frac{1}{2}\right) \left(j - \frac{1}{2}\right) \right].$$

Using the spectral decomposition  $\mathbf{C}_n^{-1} \mathbf{C}_n'^{-1} = \mathbf{P}_n \mathbf{D}_n \mathbf{P}_n$  and  $\mathbf{D}_n$  is a diagonal matrix with the  $k$ -th element  $d_k = 2[1 - \cos(\pi(\frac{2k-1}{2n+1}))]$  ( $k = 1, \dots, n$ ), we write

$$(3.7) \quad a_{kn}^* (= d_k) = 4 \sin^2 \left[ \frac{\pi}{2} \left( \frac{2k-1}{2n+1} \right) \right] \quad (k = 1, \dots, n).$$

Then the separating information maximum likelihood (SIML) estimator of  $\boldsymbol{\Sigma}_x$  in (3.2) can be defined by

$$(3.8) \quad \mathbf{G}_m = \hat{\boldsymbol{\Sigma}}_{x, SIML} = \frac{1}{m_n} \sum_{k=1}^{m_n} \mathbf{z}_k \mathbf{z}_k',$$

where we set  $m = m_n = [n^\alpha]$  ( $0 < \alpha < 1$ ).

We need to use  $m$  terms in  $n$  and the reason for this becomes clear from the frequency domain analysis, which will be explained in Section 5.

Kunitomo, Awaya and Kurisu (2019) discussed the estimation of the variance-covariance matrix  $\boldsymbol{\Sigma}_v$  when  $\mathbf{v}_i$  are i.i.d. vectors and some consistent estimators of  $\boldsymbol{\Sigma}_v$  were developed. As we shall see in Section 5, the SIML estimation method is quite robust, even when  $\mathbf{v}_i^{(x)}$  and  $\mathbf{v}_i$  are non-Gaussian stationary processes, and they are serially-correlated.

## 3.2 General Non-Stationary Errors-in-Variables Model

In order to investigate non-stationary trend-cycle component, and stationary seasonality and measurement error component, we consider the general non-stationary multivariate errors-in-variables model <sup>4</sup>

$$(3.9) \quad \mathbf{y}_i = \mathbf{x}_i + \mathbf{s}_i + \mathbf{v}_i \quad (i = 1, \dots, n).$$

We take a positive integer  $s$  ( $s > 1$ ),  $N$ , and  $n = sN$  for the resulting simplicity of exposition and arguments, We explain the general model in three steps.

(i) The trend-cycle factor  $\mathbf{x}_i$  ( $i = 1, \dots, n$ ) is a sequence of non-stationary I(1) process that satisfies

$$(3.10) \quad \Delta \mathbf{x}_i = (1 - \mathcal{L}) \mathbf{x}_i = \mathbf{v}_i^{(x)},$$

---

<sup>4</sup>It is possible to use the log-transformed data for multiplicative models. It is known that the standard model in X-12-ARIMA is a multiplicative one although it uses moving averages, for instance.

with the lag-operator  $\mathcal{L}\mathbf{x}_i = \mathbf{x}_{i-1}$ ,  $\Delta = 1 - \mathcal{L}$ ,

$$(3.11) \quad \mathbf{v}_i^{(x)} = \sum_{j=0}^{\infty} \mathbf{C}_j^{(x)} \mathbf{e}_{i-j}^{(x)},$$

and  $\mathbf{e}_i^{(x)}$  is a sequence of i.i.d. random vectors with  $\mathbf{E}(\mathbf{e}_i^{(x)}) = \mathbf{0}$  and  $\mathbf{E}(\mathbf{e}_i^{(x)} \mathbf{e}_i^{(x)'}) = \Sigma_e^{(x)}$  (positive-semi-definite). The  $p \times p$  coefficient matrices  $\mathbf{C}_j^{(x)}$  ( $= c_{kl}^{(x)}(j)$ ) are absolutely summable and  $\|\mathbf{C}_j^{(x)}\| = O(\rho^j)$ , where  $0 \leq \rho < 1$  and  $\|\mathbf{C}_j^{(x)}\| = \max_{k,l=1,\dots,p} |c_{kl}^{(x)}(j)|$ . The random vectors  $\mathbf{v}_i$  ( $i = 1, \dots, n$ ) are a sequence of stationary I(0) process with

$$(3.12) \quad \mathbf{v}_i = \sum_{j=0}^{\infty} \mathbf{C}_j^{(v)} \mathbf{e}_{i-j}^{(v)},$$

where the  $p \times p$  coefficient matrices  $\mathbf{C}_j^{(v)}$  are absolutely summable and  $\|\mathbf{C}_j^{(v)}\| = O(\rho^j)$ , where  $0 \leq \rho < 1$  and  $\mathbf{e}_i^{(v)}$  is a sequence of i.i.d. random vectors with  $\mathbf{E}(\mathbf{e}_i^{(v)}) = \mathbf{0}$ ,  $\mathbf{E}(\mathbf{e}_i^{(v)} \mathbf{e}_i^{(v)'}) = \Sigma_e^{(v)}$  (positive definite).

(ii) The seasonal factor  $\mathbf{s}_i$  ( $i = 1, \dots, n$ ) is a sequence of stationary process<sup>5</sup>, which satisfies

$$(3.13) \quad \mathbf{s}_i = \sum_{j=0}^{\infty} \mathbf{C}_{sj}^{(s)} \mathbf{e}_{i-sj}^{(s)},$$

where the lag-operator is defined by  $\mathcal{L}^s \mathbf{s}_i = \mathbf{s}_{i-s}$  ( $s \geq 2$ ), and  $\mathbf{e}_i^{(s)}$  is a a sequence of i.i.d. random vectors with  $\mathbf{E}(\mathbf{e}_i^{(s)}) = \mathbf{0}$  and  $\mathbf{E}(\mathbf{e}_i^{(s)} \mathbf{e}_i^{(s)'}) = \Sigma_e^{(s)}$  (a non-negative definite matrix). The  $p \times p$  coefficient matrices  $\mathbf{C}_j^{(s)}$  are absolutely summable and  $\|\mathbf{C}_j^{(s)}\| = O(\rho^j)$ , where  $0 \leq \rho < 1$ .

(iii) Let  $\mathbf{f}_{\Delta x}(\lambda)$ ,  $\mathbf{f}_v(\lambda)$  and  $\mathbf{f}_s(\lambda)$  be the spectral density<sup>6</sup> ( $p \times p$ ) matrices of  $\Delta \mathbf{x}_i$ ,  $\mathbf{v}_i$  and  $\mathbf{s}_i$  ( $i = 1, \dots, n$ ) defined by

$$(3.14) \quad \mathbf{f}_{\Delta x}(\lambda) = \left( \sum_{j=0}^{\infty} \mathbf{C}_j^{(x)} e^{2\pi i \lambda j} \right) \Sigma_e^{(x)} \left( \sum_{j=0}^{\infty} \mathbf{C}_j^{(x)'} e^{-2\pi i \lambda j} \right), \quad \left( -\frac{1}{2} \leq \lambda \leq \frac{1}{2} \right),$$

$$(3.15) \quad \mathbf{f}_v(\lambda) = \left( \sum_{j=0}^{\infty} \mathbf{C}_j^{(v)} e^{2\pi i \lambda j} \right) \Sigma_e^{(v)} \left( \sum_{j=0}^{\infty} \mathbf{C}_j^{(v)'} e^{-2\pi i \lambda j} \right), \quad \left( -\frac{1}{2} \leq \lambda \leq \frac{1}{2} \right),$$

---

<sup>5</sup>When we assume that  $\Delta \mathbf{s}_i$  is stationary, most arguments in the following sections would go through. We have omitted its details.

<sup>6</sup>Our notation of spectral density is slightly different from the standard notation used in Anderson (1971). Let  $\mu = 2\pi\lambda$  and  $f^A(\mu)$  ( $-\pi \leq \mu \leq \pi$ ) be the spectral density in Chapter 7 of Anderson (1971). Then,  $f(\lambda) = 2\pi f^A(\mu)$  ( $-\frac{1}{2} \leq \lambda \leq \frac{1}{2}$ ). The present defition of spectral density corresponds to  $2\pi f^A(-\mu)$  in some literature.

and

$$(3.16) \quad \mathbf{f}_s(\lambda) = \left( \sum_{j=0}^{\infty} \mathbf{C}_{sj}^{(s)} e^{2\pi i \lambda s j} \right) \boldsymbol{\Sigma}_e^{(s)} \left( \sum_{j=0}^{\infty} \mathbf{C}_{sj}^{(s)'} e^{-2\pi i \lambda s j} \right) \quad \left( -\frac{1}{2} \leq \lambda \leq \frac{1}{2} \right),$$

where we set  $\mathbf{C}_0^{(x)} = \mathbf{C}_0^{(v)} = \mathbf{C}_0^{(s)} = \mathbf{I}_p$  for normalization and  $i^2 = -1$  (see Chapter 7 of Anderson (1971)).

Then the  $p \times p$  spectral density matrix of the transformed vector process of difference series  $\Delta \mathbf{y}_i (= \mathbf{y}_i - \mathbf{y}_{i-1})$  can be represented as

$$(3.17) \quad \mathbf{f}_{\Delta y}(\lambda) = \mathbf{f}_{\Delta x}(\lambda) + (1 - e^{2\pi i \lambda}) [\mathbf{f}_s(\lambda) + f_v(\lambda)] (1 - e^{-2\pi i \lambda}) .$$

We denote the long-run variance-covariance matrices of trend-cycle and noise components for  $g, h = 1, \dots, p$  as

$$(3.18) \quad \boldsymbol{\Sigma}_x = \mathbf{f}_{\Delta x}(0) (= (\sigma_{gh}^{(x)})) , \quad \boldsymbol{\Sigma}_v = f_v(0) = (\sigma_{gh}^{(v)}) .$$

respectively.

One important often neglected issue is that when applying the differencing procedure to non-stationary time series and using the standard statistical method for multivariate stationary time series, there is no guarantee of keeping the relationships among the original time series as they were by using the transformations. Although Engle and Granger (1987) and Johansen (1995) noticed this problem, they did not consider the frequency domain aspect with seasonality and measurement errors. The SIML filtering approach may shed a new light on the relationships among the time domain and frequency decompositions of non-stationary multivariate time series.

## 4 The SIML Filtering Method

### 4.1 Basic Filtering

We introduce the general filtering procedure based on the  $\mathbf{K}_n$ -transformation in (3.4). When we interpret that the elements of the resulting  $n \times p$  random matrix  $\mathbf{Z}_n$  take real values in the frequency domain, it is easy to understand their roles. Since  $\mathbf{P}_n$  is a kind of real-valued discrete Fourier transformation, vectors  $\mathbf{z}_k$  ( $k = 1, \dots, n$ ) in  $\mathbf{Z}_n$  are asymptotically uncorrelated, as we shall discuss in Section 5.1. We investigate the partial inversion of the transformed orthogonal processes. Let an  $n \times p$  matrix

$$(4.19) \quad \hat{\mathbf{X}}_n(\mathbf{Q}) = \mathbf{C}_n \mathbf{P}_n \mathbf{Q}_n \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0)$$

and

$$(4.20) \quad \mathbf{Z}_n = \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0), \mathbf{Y}_n = \bar{\mathbf{Y}}_0 + \mathbf{X}_n^* + \mathbf{S}_n + \mathbf{V}_n,$$

where  $\mathbf{X}_n^* = (\mathbf{x}_t^*)'$ ,  $\mathbf{S}_n = (\mathbf{s}_t')$  and  $\mathbf{V}_n = (\mathbf{v}_t')$  are  $n \times p$  matrices,  $\mathbf{x}_t^* = \mathbf{x}_t - \mathbf{x}_0$  ( $t = 1, \dots, n$ ) and  $\mathbf{x}_0 = \mathbf{y}_0$  is the initial vector. (We use the subscript  $t$  instead of  $i$  in this subsection.)

The stochastic process  $\mathbf{Z}_n$  is the orthogonal decomposition of the original time series  $\mathbf{Y}_n$  in the frequency domain and  $\mathbf{Q}_n$  is an  $n \times n$  filtering matrix. Because  $\mathbf{Y}_n$  consist of non-stationary time series, we need a special form of transformation  $\mathbf{K}_n$  in (3.4). We give explicit forms of two examples, including the trend-cycle filtering and the band filtering procedures. Although there can be many possible filtering procedures within our general framework, it is useful to discuss some linear filtering procedures. Let an  $n \times n$  diagonal matrix

$$(4.21) \quad \mathbf{Q}_n = \sum_{t=1}^n w_{t,n} \mathbf{e}_t^{(n)} \mathbf{e}_t^{(n)'}$$

and  $\mathbf{e}_t^{(n)} = (0, \dots, 1, \dots)'$  ( $t = 1, \dots, n$ ) are the unit vectors and  $w_{t,n}$  ( $t = 1, \dots, n$ ) are some non-negative constants.

We start with the case when  $w_{t,n} = 1$  ( $t = 1, \dots, n$ ) and we have the identity matrix  $\mathbf{Q}_n = \mathbf{I}_n$ . Then we find  $\mathbf{C}_n \mathbf{P}_n \mathbf{Q}_n \mathbf{P}_n \mathbf{C}_n^{-1} = \mathbf{I}_n$ . There can be useful cases and we present two cases as the trend-cycle filtering and the band filtering by choosing  $w_{t,n} = 1$  or 0 for some  $t$ 's. Although the first example could be regarded a special case of the second one, the analysis of trend-cycle component in the first example has an important role in economic time series.

(i) **Trend-Cycle Filtering:** Let an  $m \times n$  ( $m < n$ ) choice matrix  $\mathbf{J}_m = (\mathbf{I}_m, \mathbf{O})$ , and let also  $n \times p$  matrix

$$(4.22) \quad \hat{\mathbf{X}}_n^{(m)} = \mathbf{C}_n \mathbf{P}_n \mathbf{J}_m' \mathbf{J}_m \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0)$$

and an  $n \times n$  matrix  $\mathbf{Q}_n = \mathbf{J}_m' \mathbf{J}_m$ .

We construct an estimator of  $n \times p$  hidden state matrix  $\mathbf{X}_n^*$  only in the lower frequency parts by using the inverse transformation of  $\mathbf{Z}_n$  and deleting the estimated seasonal and noise parts. We denote the hidden trend-cycle state based on  $m$  frequencies as

$$(4.23) \quad \mathbf{X}_n^{(m)} = \mathbf{C}_n \mathbf{P}_n \mathbf{J}_m' \mathbf{J}_m \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{X}_n^*.$$

This quantity is different from  $\mathbf{X}_n^*$  because  $\mathbf{x}_t$  ( $t = 1, \dots, n$ ) in (3.9) and (3.10) contains not only the trend-cycle component of  $\mathbf{y}_t$  ( $t = 1, \dots, n$ ), but also the noise component in the frequency domain, which is different from the measurement noise component  $\mathbf{v}_t$  ( $t = 1, \dots, n$ ) in (3.9)-(3.13). We try to estimate the trend-cycle component of  $\mathbf{x}_t$  by using (4.22) and recover the trend-cycle component of  $\mathbf{X}_n$  near

at the zero frequency because the effects of differenced measurement error noises  $(\mathbf{v}_t - \mathbf{v}_{t-1})$  are negligible around at zero frequency. This method differs from some existing procedures that consider the decomposition of time series only in the time domain. Our arguments can be justified by using the frequency decomposition of  $\mathbf{y}_t$  and  $\mathbf{r}_t^{(n)} = \Delta \mathbf{y}_t$  ( $= \mathbf{y}_t - \mathbf{y}_{t-1}$  and  $\mathbf{y}_0$  being fixed), and we shall discuss this issue in Section 5.2.

We partition  $\mathbf{P}_n$  into  $[m + (n - m)] \times [m + (n - m)]$  matrices as

$$\mathbf{P}_n = \begin{pmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{pmatrix}$$

and then

$$\mathbf{P}_n \mathbf{J}'_m \mathbf{J}_m \mathbf{P}_n = \begin{pmatrix} \mathbf{P}_{11} \\ \mathbf{P}_{21} \end{pmatrix} (\mathbf{P}_{11}, \mathbf{P}_{12}) = \mathbf{I}_n - \begin{pmatrix} \mathbf{P}_{12} \\ \mathbf{P}_{22} \end{pmatrix} (\mathbf{P}_{21}, \mathbf{P}_{22}) .$$

After straightforward calculations (see the Appendix for the derivation), the  $(j, j')$ -th element of  $\mathbf{A}_n = \mathbf{P}_n \mathbf{J}'_m \mathbf{J}_m \mathbf{P}_n$  ( $= (a_{j,j'}^{(n,m)})$ ) is given by

$$(4.24) \quad a_{j,j}^{(n,m)} = \frac{2m}{2n+1} + \frac{1}{2n+1} \left[ \frac{\sin \frac{2m\pi}{2n+1} (2j-1)}{\sin \frac{\pi}{2n+1} (2j-1)} \right] ,$$

$$a_{j,j'}^{(n,m)} = \frac{1}{2n+1} \left[ \frac{\sin \frac{2m\pi}{2n+1} (j+j'-1)}{\sin \frac{\pi}{2n+1} (j+j'-1)} + \frac{\sin \frac{2m\pi}{2n+1} (j-j')}{\sin \frac{\pi}{2n+1} (j-j')} \right] \quad (j \neq j') .$$

It is possible to evaluate MSE of statistical state vector estimation. Let  $\hat{\mathbf{X}}_n^{(m)} = (\hat{\mathbf{X}}_{ni}^{(m)})$  ( $\hat{\mathbf{X}}_{ni}^{(m)}$  is an  $n \times 1$  vector,  $i = 1, \dots, p$ ) and  $\mathbf{X}_n^{(m)} = (\mathbf{X}_{ni}^{(m)})$  ( $\mathbf{X}_{ni}^{(m)}$  is an  $n \times 1$  vector for  $i = 1, \dots, p$ ). By decomposing  $\mathbf{Y}_n - \bar{\mathbf{Y}}_0 = \mathbf{X}_n^* + \mathbf{S}_n + \mathbf{V}_n$  and  $\hat{\mathbf{X}}_n^{(m)} - \mathbf{X}_n^* = \mathbf{C}_n \mathbf{P}_n \mathbf{J}'_m \mathbf{J}_m \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{S}_n + \mathbf{V}_n)$ , we find

$$\mathbf{E}[(\hat{\mathbf{X}}_{ni}^{(m)} - \mathbf{X}_{ni}^{(m)})(\hat{\mathbf{X}}_{nj}^{(m)} - \mathbf{X}_{nj}^{(m)})'] = \mathbf{K}_n^{*-1} \mathbf{Q}_n \mathbf{K}_n^* \Gamma_n^{(s+v)}(i, j) \mathbf{K}_n^{*'} \mathbf{Q}_n \mathbf{K}_n^{*-1} ,$$

where  $\mathbf{K}_n^* = \mathbf{P}_n \mathbf{C}_n^{-1}$ ,  $\Gamma_n^{(s+v)}(i, j)$  is the  $n \times n$  variance-covariance matrix of  $\mathbf{S}_{ni} + \mathbf{V}_{ni}$  and  $\mathbf{S}_{nj} + \mathbf{V}_{nj}$  ( $n \times 1$  vectors for  $i, j = 1, \dots, p$ ), and,  $\mathbf{S}_{ni}$  and  $\mathbf{V}_{ni}$  ( $n \times 1$  vectors) are the  $i$ -th and  $j$ -th column vectors of  $\mathbf{S}_n$  and  $\mathbf{V}_n$ , respectively.

Since (25) does not depend on  $\mathbf{x}_t$ , (22) minimizes the MSE with respect to unknown state vector  $\mathbf{x}_t$  to estimate (23) and it is optimal in this sense.

(ii) **Band Filtering** : We consider a general filtering based on the  $\mathbf{K}_n$  transformation in (3.4) and use the inversion of some frequency parts of the random matrix  $\mathbf{Z}_n$ . The leading example is the analysis of seasonal frequencies in the discrete time series and we take  $s$  ( $> 1$ ) being a positive integer as the seasonal lag.

Let an  $m_2 \times [m_1 + m_2 + (n - m_1 - m_2)]$  choice matrix  $\mathbf{J}_{m_1, m_2} = (\mathbf{O}, \mathbf{I}_{m_2}, \mathbf{O})$  (we take  $m_1 + m_2 < n$ ), and let also  $n \times p$  matrix

$$(4.25) \quad \hat{\mathbf{X}}_n^{(m_1, m_2)} = \mathbf{C}_n \mathbf{P}_n \mathbf{J}'_{m_1, m_2} \mathbf{J}_{m_1, m_2} \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0)$$

and an  $n \times n$  matrix  $\mathbf{Q}_n = \mathbf{J}'_{m_1, m_2} \mathbf{J}_{m_1, m_2}$ .

When we have a particular seasonal frequency  $s$  ( $> 1$ ), for instance, we can take  $m_1 = [2n/s] - [m/2]$  and  $m_2 = m$ . We set  $s = 4$  for quarterly data and  $s = 12$  for monthly data. (We may have many seasonal frequencies in data analysis as pointed out by Granger and Hatanaka (1964) already.)

As in the trend-cycle filtering problem, (4.25) is the SIML-filtering value for

$$(4.26) \quad \mathbf{X}_n^{(m_1, m_2)} = \mathbf{C}_n \mathbf{P}_n \mathbf{J}'_{m_1, m_2} \mathbf{J}_{m_1, m_2} \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{X}_n^* + \mathbf{S}_n),$$

and it is an estimate of some frequency components of  $\mathbf{x}_t + \mathbf{s}_t$  ( $t = 1, \dots, n$ ) in (3.9), (3.10) and (3.13). The filtering problem becomes difficult because some frequency component of  $\mathbf{y}_t$  includes not only some component  $\mathbf{x}_t$  at the same frequency, but also some component of the measurement errors at the same frequency.

After straightforward calculations (see the Appendix) in this case, the  $(j, j')$ -th element of  $\mathbf{A}_n = \mathbf{P}_n \mathbf{J}'_{m_1, m_2} \mathbf{J}_{m_1, m_2} \mathbf{P}_n$  ( $= (a_{j, j'}^{(n, m_1, m_2)})$ ) is given by

$$(4.27) \quad a_{j, j}^{(n, m_1, m_2)} = \frac{2m_2}{2n+1} + \frac{1}{2n+1} \left[ \frac{\sin \frac{2(m_1+m_2)\pi}{2n+1} (2j-1) - \sin \frac{2(m_1)\pi}{2n+1} (2j-1)}{\sin \frac{\pi}{2n+1} (2j-1)} \right],$$

$$a_{j, j'}^{(n, m_1, m_2)} = \frac{1}{2n+1} \left[ \frac{\sin \frac{2(m_1+m_2)\pi}{2n+1} (j+j'-1) - \sin \frac{2(m_1)\pi}{2n+1} (j+j'-1)}{\sin \frac{\pi}{2n+1} (j+j'-1)} + \frac{\sin \frac{2(m_1+m_2)\pi}{2n+1} (j-j') - \sin \frac{2(m_1)\pi}{2n+1} (j-j')}{\sin \frac{\pi}{2n+1} (j-j')} \right] \quad (j \neq j').$$

When  $m_1 = 0$  and  $m_2 = m$ , the resulting formula reduces to the trend-cycle filtering case. There are existing filtering procedures having the frequency domain interpretation, but it seems that our procedure differs from some existing literature because of (4.24) and (4.27).

It is also possible to evaluate MSE of the state vector estimation in the same way as Case (i). Let  $\hat{\mathbf{X}}_n^{(m_1, m_2)}$  be an estimate of the hidden state as

$$\mathbf{X}_n^{(s, m_1, m_2)} = \mathbf{C}_n \mathbf{P}_n \mathbf{J}'_{m_1, m_2, n} \mathbf{J}_{m_1, m_2, n} \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{X}_n^* + \mathbf{S}_n).$$

By using the similar calculation as Case (i) and the notation of  $\Gamma_n^{(v)}(i, j)$  ( $i, j = 1, \dots, p$ ), we find

$$(4.28) \quad \mathbf{E}[(\hat{\mathbf{X}}_{ni}^{(m_1, m_2)} - \mathbf{X}_{ni}^{(s, m_1, m_2)})(\hat{\mathbf{X}}_{nj}^{(m_1, m_2)} - \mathbf{X}_{nj}^{(s, m_1, m_2)})']$$

$$= \mathbf{K}_n^{*-1} \mathbf{Q}_n \mathbf{K}_n^* \Gamma_n^{(v)}(i, j) \mathbf{K}_n^{*'} \mathbf{Q}_n \mathbf{K}_n^{*-1},$$

where  $\mathbf{\Gamma}_n^{(v)}(i, j)$  is the  $n \times n$  variance-covariance matrix of  $\mathbf{V}_{ni}$  and  $\mathbf{V}_{nj}$ . We use an  $(m_2 - m_1) \times n$  choice matrix  $\mathbf{J}_{m_1, m_2} = (\mathbf{O}, \mathbf{I}_{m_2 - m_1}, \mathbf{O})$  and  $\mathbf{Q}_n^* = \mathbf{I}_n - \mathbf{J}_{m_1, m_2}^T \mathbf{J}_{m_1, m_2}$ . If there are several seasonal frequencies, we need more complicated filtering procedures.

## 5 On a Statistical Foundation

### 5.1 Underlying Asymptotic Theory

At first glance, the SIML filtering procedure may be regarded as an *ad-hoc* statistical procedure without any mathematical foundation. However, it has a rather solid statistical foundation.

Let  $\theta_{jk} = \frac{2\pi}{2n+1}(j - \frac{1}{2})(k - \frac{1}{2})$ ,  $p_{jk}^{(n)} = \frac{1}{\sqrt{2n+1}}(e^{i\theta_{jk}} + e^{-i\theta_{jk}})$  and for  $\mathbf{Y}_n = (\mathbf{y}'_i)$  we write  $\mathbf{z}_k$  ( $k = 1, \dots, n$ ) as

$$(5.29) \quad \mathbf{z}_n(\lambda_k^{(n)}) = \sum_{j=1}^n p_{jk}^{(n)} \mathbf{r}_j, \quad \mathbf{r}_j = \mathbf{y}_j - \mathbf{y}_{j-1},$$

which is a (real-valued) Fourier-transformation and  $\mathbf{y}_0$  is fixed.

Then, we find that  $\mathbf{z}_n(\lambda_k^{(n)})$  ( $k = 1, \dots, n$ ) are the (real-valued) Fourier-transformation of data at the frequency  $\lambda_k^{(n)}$  ( $= (k - 1/2)/(2n + 1)$ ), which is a (real-part of) estimate of the orthogonal incremental process  $\mathbf{z}(\lambda)$  ( $0 \leq \lambda \leq 1/2$ ), which is continuous in the frequency domain.

We shall utilize an asymptotic theory for the stationary linear processes because the time series model defined by (3.9)-(3.13) can be regarded as a special case. Let

$$(5.30) \quad \mathbf{r}_i = \Delta \mathbf{y}_i = \boldsymbol{\mu} + \sum_{j=0}^{\infty} \mathbf{\Gamma}_j \mathbf{w}_{i-j},$$

where  $\boldsymbol{\mu}$  is a constant vector (we assume  $\boldsymbol{\mu} = \mathbf{0}$  for simplicity in this section),  $\mathbf{\Gamma}_j$  are  $p \times p$  matrices, and  $\mathbf{w}_i$  are a sequence of mutually independent random variables with  $\mathbf{E}[\mathbf{w}_i] = \mathbf{0}$ ,  $\mathbf{E}[\mathbf{w}_i \mathbf{w}_i'] = \boldsymbol{\Sigma}_u (> 0)$ . The errors-in-variables model in Section 3 implies that the  $p \times p$  matrices  $\mathbf{\Gamma}_j$  satisfy  $\sum_{h=0}^{\infty} \|\mathbf{\Gamma}_j\| < \infty$ .

We then summarize the useful result on  $\mathbf{z}(\lambda_k^{(n)})$ . Although it could be regarded as a direct extension of Theorem 8.4.3 of Anderson (1971) for discrete and (ergodic) stationary time series, we could not find the following representation. The proof is given in the Appendix.

**Proposition 1:** Let  $\mathbf{r}_j$  ( $j = 1, \dots, n$ ) be an ergodic stationary stochastic process given by (5.30) with  $\sum_{h=0}^{\infty} \|\mathbf{\Gamma}_j\| < \infty$  and the fourth order moments of each element of  $\mathbf{w}_i$  are finite.



Let also  $\mathbf{z}_n(\lambda_k^{(n)}) = \sum_{j=1}^n p_{jk}^{(n)} \mathbf{r}_j$  and  $\mathbf{r}_j$  be an ergodic stationary sequence with  $\mathbf{E}[\mathbf{r}_j] = \mathbf{0}$ ,  $\Gamma(h) = \mathbf{E}(\mathbf{r}_j \mathbf{r}_{j-h}')$ , and the (symmetrized real-valued) spectral density matrix

$$(5.31) \quad \mathbf{f}_{SR}(\lambda) = \sum_{h=-\infty}^{\infty} \cos(2\pi h\lambda) \Gamma(h) \quad , \quad 0 \leq \lambda \leq \frac{1}{2}$$

is the positive definite and bounded (real-valued and symmetrized) spectral matrix. Assume that  $\lambda_k^{(n)} \rightarrow s$ ,  $\lambda_{k'}^{(n)} \rightarrow t$  as  $n \rightarrow \infty$  for  $0 < s < t < \frac{1}{2}$ . Then, as  $n \rightarrow \infty$

$$(5.32) \quad \begin{bmatrix} \mathbf{z}_n(\lambda_k^{(n)}) \\ \mathbf{z}_n(\lambda_{k'}^{(n)}) \end{bmatrix} \xrightarrow{w} N_{2p} \left[ \mathbf{0}, \begin{bmatrix} \mathbf{f}_{SR}(s) & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_{SR}(t) \end{bmatrix} \right] .$$

This proposition covers the general model with (3.9)-(3.13) with the moment conditions because  $\Delta \mathbf{y}_i$  are stationary. As the asymptotic variance-covariance matrix of the orthogonal random vectors  $\mathbf{z}_n(\lambda_k^{(n)})$  is the (symmetrized real) spectral density matrix, it can be estimated consistently. When we have noise terms as in Section 3, it is not possible to estimate the (long-run) variance-covariance matrix  $\Sigma_x$  of trend-cycle component simply by using the differenced time series  $\mathbf{r}_j (= \Delta \mathbf{y}_j) = \mathbf{y}_j - \mathbf{y}_{j-1}$ . It is because

$$(5.33) \quad \mathbf{E}[\mathbf{r}_j \mathbf{r}_j'] = \mathbf{E}[\Delta \mathbf{x}_j \Delta \mathbf{x}_j'] + \mathbf{E}[\Delta \mathbf{s}_j \Delta \mathbf{s}_j'] + \mathbf{E}[\Delta \mathbf{v}_j \Delta \mathbf{v}_j'] .$$

The SIML estimator of  $\mathbf{f}_{SR}(0)$  ( $= \mathbf{f}_{\Delta x}(0)$ )  $= \Sigma_x$  in the general case can be defined by

$$(5.34) \quad \mathbf{G}_m = \frac{1}{m_n} \sum_{k=1}^m (\mathbf{z}_n(\lambda_k^{(n)})) (\mathbf{z}_n(\lambda_k^{(n)}))' .$$

Then we summarize the basic property of the SIML estimation of  $\Sigma_x$  and the derivation is given in the Appendix.

**Proposition 2:** Assume that the fourth order moments of each element of  $\mathbf{v}_i^{(x)}$ ,  $\mathbf{v}_i$  and  $\mathbf{v}_i^{(s)}$  ( $i = 1, \dots, n$ ) in (3.9)-(3.13) are bounded. We set  $m_n = [n^\alpha]$  ( $0 < \alpha < 1$ ). Then, as  $n \rightarrow \infty$

$$(5.35) \quad \mathbf{G}_m \xrightarrow{p} \Sigma_x .$$

The above result has some implication on the use of the frequency domain analysis of (non-stationary) multiple time series. For  $0 \leq \mu \leq \frac{1}{2}$ , let the symmetrized (real-valued) spectral density matrix of  $\Delta \mathbf{y}_j$  ( $j = 1, \dots, n$ ) be  $f_{SR, \Delta y}(\mu) = (1/2)[f_{\Delta y}(\mu) +$

$\bar{f}_{\Delta y}(\mu)]$ , where  $\bar{f}(\cdot)$  is the complex conjugate of  $f$  and we have the initial condition  $\mathbf{y}_0$ .

Similarly, we find that  $f_{SR,\Delta x}(\lambda) = (1/2)[f_{\Delta x}(\lambda) + \bar{f}_{\Delta x}(\lambda)]$ ,

$f_{SR,\Delta s}(\lambda) = (1/2)(1 - e^{2\pi i\lambda})[f_s(\lambda) + \bar{f}_s(\lambda)](1 - e^{-2\pi i\lambda})$  and

$f_{SR,\Delta v}(\lambda) = (1/2)(1 - e^{2\pi i\lambda})[f_v(\lambda) + \bar{f}_v(\lambda)](1 - e^{-2\pi i\lambda})$ .

They are the consequence of (3.17), (5.29) and (5.31) for the symmetrized spectral (real-valued) density matrices.

From this interpretation, it may be easy to find that  $\mathbf{G}_m$  is a consistent estimator of the long-run variance-covariance matrix, which is the spectral density matrix at the zero-frequency  $f_{SR,\Delta x}(0)$  when we use the  $\mathbf{Z}_n$ -transformation of data.

## 5.2 Frequency Interpretation of SIML-filtering

In the traditional statistical time series analysis for a stationary discrete (vector) process with the (complex-valued) spectral distribution  $F$ , it has a representation with right-continuous (complex-valued) orthogonal increments in the frequency domain (see Doob (1953), Brockwell and Davis (1990) for the details). Chapter 7 of Anderson (1971) is informative because of its discussion on real-valued representations although it has only univariate cases. The real-valued multivariate orthogonal processes and the spectral density matrix play important roles in our formulation.

For  $\lambda_k^{(n)} = (k - 1/2)/(2n + 1)$  ( $k = 1, \dots, n$ ), we rewrite (5.29) as

$$(5.36) \quad \mathbf{z}_n(\lambda_k^{(n)}) = \sum_{j=1}^n \mathbf{r}_j^{(n)} \frac{2}{\sqrt{2n+1}} \cos[2\pi\lambda_k^{(n)}(j - \frac{1}{2})] \quad (k = 1, \dots, n).$$

Then, by using the inversion transformation with  $\mathbf{P}_n$ , we can confirm that

$$(5.37) \quad \mathbf{r}_s^{(n)} = \sum_{k=1}^n p_{sk} \mathbf{z}_n(\lambda_k^{(n)}) \quad (s = 1, \dots, n).$$

It is another representation of  $\mathbf{R}_n = (\mathbf{r}_i^{(n)'}) = \mathbf{C}_n^{-1} \hat{\mathbf{X}}_n(\mathbf{Q})$  in (4.19) when  $\mathbf{Q}_n = \mathbf{I}_n$ . For any  $s$  ( $s = 1, \dots, n$ ),  $\mathbf{r}_s^{(n)}$  can be recovered as the weighted sum of orthogonal processes  $\mathbf{z}_n(\lambda_k^{(n)})$  at frequency  $\lambda_k^{(n)}$  ( $k = 1, \dots, n$ ). We then, by using  $\mathbf{Y}_n = \mathbf{C}_n \mathbf{R}_n$ , recover the non-stationary process  $\mathbf{y}_t^{(n)}$  ( $t = 1, \dots, n$ ) given the initial condition  $\mathbf{y}_0$  as

$$(5.38) \quad \mathbf{y}_t^{(n)} = \mathbf{y}_0 + \sum_{s=1}^t \mathbf{r}_s^{(n)}.$$

Let

$$(5.39) \quad \alpha_n(\lambda_m^{(n)}, j - \frac{1}{2}) = \frac{1}{n} \sum_{k=1}^m [2 \cos 2\pi\lambda_k^{(n)}(j - \frac{1}{2})]$$

Then, when  $\lambda_m^{(n)} \rightarrow \lambda$  as  $n \rightarrow \infty$  ( $0 < \lambda < \frac{1}{2}$ ), using Lemma 5.1 of Kunitomo et al. (2018), we find

$$\alpha_n(\lambda_m^{(n)}, j - \frac{1}{2}) \rightarrow \alpha(\lambda, j - \frac{1}{2}) = \frac{2 \sin 2\pi\lambda(j - \frac{1}{2})}{\pi(j - \frac{1}{2})}.$$

If we set the uncorrelated stochastic process of uncorrelated increments with continuous parameter  $\lambda$  ( $0 \leq \lambda \leq \frac{1}{2}$ ) as  $A_n(\lambda) = \sum_{j=1}^n \alpha(\lambda, j - \frac{1}{2}) \mathbf{r}_j^{(n)}$ , then we find

$$(5.40) \quad \int_0^{\frac{1}{2}} \cos[2\pi\lambda(s - \frac{1}{2})] dA_n(\lambda) = \mathbf{r}_s^{(n)} \quad (s = 1, \dots, n).$$

This corresponds to the continuous representation of a discrete (real-valued) stationary time series in the frequency domain (see Chapter 7.4 of Anderson (1971)). If we write the limit of  $\mathbf{A}(\lambda) = \lim_{n \rightarrow \infty} \mathbf{A}_n(\lambda)$  (assuming it exists), the (real-valued) spectral distribution matrix  $F_{RS}$  for any  $0 \leq \lambda_1 < \lambda_2 \leq 1/2$  can be defined as

$$(5.41) \quad F_{RS}(\lambda_2 - \lambda_1) = \mathbf{E}[(\mathbf{A}(\lambda_2 - \lambda_1)\mathbf{A}(\lambda_2 - \lambda_1)')] = \int_{\lambda_1}^{\lambda_2} f_{RS}(\lambda) d\lambda$$

if  $F_{RS}$  is absolutely continuous and the matrix-valued density process  $f_{RS}(\lambda)$  ( $0 \leq \lambda_1 < \lambda_2 \leq 1/2$ ) exists.

From (4.22), we set  $\hat{\mathbf{R}}_n(m) = (\hat{\mathbf{r}}_i^{(m,n)})' = \mathbf{C}_n^{-1} \hat{\mathbf{X}}_n(m)$  and  $\hat{\mathbf{r}}_i^{(m,n)}$  are  $p \times 1$  vectors for  $i = 1, \dots, n$ . If we write

$$(5.42) \quad \hat{\mathbf{r}}_s^{(m,n)} = \sum_{k=1}^m p_{sk} \mathbf{z}_n(\lambda_k^{(n)}) \quad (s = 1, \dots, m; 0 < m < n),$$

it is the trend-cycle estimate of the SIML-filtering value for  $\mathbf{r}_s^{(m,n)}$ , which is the corresponding element of  $\mathbf{R}_n(m)$  ( $= \mathbf{C}_n \mathbf{X}_n(m)$ ). It corresponds to

$$(5.43) \quad \mathbf{r}_s^{(m,n)} = \sum_{k=1}^m p_{sk} \mathbf{z}_n^*(\lambda_k^{(n)}) \quad (s = 1, \dots, m; 0 < m < n),$$

where  $\mathbf{z}_n^*(\lambda_k^{(n)})$  are constructed from the  $n \times p$  hidden states matrix  $\mathbf{X}_n$  instead of the observed  $n \times p$  matrix data  $\mathbf{Y}_n$ . Hence (5.42) is the same as the element of  $\mathbf{C}_n^{-1} \hat{\mathbf{X}}_n(m)$  in (4.22), and for  $\lambda_m^{(n)} = m/(2n)$  in the frequency domain it is a discrete version of

$$(5.44) \quad \hat{\mathbf{r}}_s^{(n)}(\lambda_m^{(n)}) = \int_0^{\lambda_m^{(n)}} \cos[2\pi\lambda(s - \frac{1}{2})] dA_n(\lambda).$$

Then, (5.43) is the same as the element of  $\mathbf{C}_n^{-1} \mathbf{X}_n^{(m)}$  in (4.23), and it has the corresponding (continuous) version in the frequency domain.

Similarly,  $\hat{\mathbf{r}}_s^{(m_1, m_2, n)} = \sum_{k=m_1+1}^{m_1+m_2} p_{sk} \mathbf{z}_n(\lambda_k^{(n)}) = \hat{\mathbf{r}}_s^{(m_2, n)} - \hat{\mathbf{r}}_s^{(m_1, n)}$  ( $s = 1, \dots, m$ ;  $0 < m_1 < m_2 < n$ ) can be regarded as a discrete version of

$$(5.45) \quad \hat{\mathbf{r}}_s^{(n)}(\lambda_{m_1}^{(n)}, \lambda_{m_2}^{(n)}) = \int_{\lambda_{m_1}^{(n)}}^{\lambda_{m_2}^{(n)}} \cos[2\pi\lambda(s - \frac{1}{2})] dA_n(\lambda).$$

From our interpretation of the SIML filtering we find an interesting representation of (discrete time and real-valued) stationary processes and orthogonal incremental stochastic processes. They are closely related to our method of data analysis for non-stationary vector time series.

## 6 Some Simulation and the Choice of Frequencies

### 6.1 A Guide of Choosing Frequencies

When we are interested in filtering a non-stationary time series with trend-cycle, seasonality and measurement errors, we need to choose the parameter  $m$ . We give a guide to set  $m$  in the trend-cycle filtering case for practical purpose. From the discussion of Section 5.2, the orthogonal process  $\mathbf{z}_n(\lambda_k^{(n)})$  corresponds to the frequency  $\lambda_k^{(n)} = (k - \frac{1}{2})/(2n + 1)$  ( $k = 1, \dots, n$ ). When we are interested in the trend-cycle component, we may only use  $\lambda_k^{(n)} = (k - \frac{1}{2})/(2n + 1)$  ( $k = 1, \dots, m$ ) and then the maximum frequency is approximately

$$(6.46) \quad \lambda_{max}^{(n)} = \frac{m}{2n}.$$

For instance, when we have monthly data over 20 years as an example, we have  $n = 240$  and  $s = 12$ . Since we have seasonal frequencies, we want to find trend-cycle components as business cycles more than 1.5 year, say. Then an appropriate maximum frequency would be  $\lambda_{max}^{(n)} = 1.5/24$  and then we could take  $m^* = 480 \times (1.5/24) = 30$ .

As another example if we have quarterly data over 30 years, we have  $n = 120$  and  $s = 4$ . Since we have seasonal frequency, we want to find trend-cycle components as business cycles more than 1.5 year, say. Then an appropriate maximum frequency would be  $\lambda_{max}^{(n)} = 1.5/8$  and then we could take  $m^* = 240 \times (1.5/8) = 45$ .

If we were interested in the trend-cycle component of the non-stationary time series, these choices might be reasonable candidates.

As we shall see in the next subsection, this point could be checked by simulation for prediction.

### 6.2 Some Simulation

When we have estimates of the state variables  $\mathbf{x}_i^{(m)}$  ( $i = 1, \dots, n$ ), the estimates of error components are  $\hat{\mathbf{v}}_i^{(m)} = \mathbf{y}_i - \hat{\mathbf{x}}_i^{(m)}$  ( $i = 1, \dots, n$ ). Then, an estimated MSE

of the one-step ahead prediction errors based on the SIML-smoothing or filtering is given by

$$(6.47) \quad \text{PMSE}_n(h) = \mathbf{E}[(\mathbf{y}_{n+h} - \mathbf{y}_n)(\mathbf{y}_{n+h} - \mathbf{y}_n)' | \mathcal{F}_n] ,$$

where  $\mathcal{F}_n$  is the  $\sigma$ -field (information) available at  $n$ .

One may try to minimize the estimated h-step prediction MSE by choosing an appropriate  $m$ . It may be reasonable to take  $h = 2, 3$  as short-run prediction and  $h = 4, 8$  for long-run prediction from our limited experiments.

We only show a result on the simple trend plus noise model when  $p = 1$ ,  $x_i = x_{i-1} + v_i^{(x)}$  and  $y_i = x_i + v_i$  ( $i = 1, \dots, n$ ) with  $v_i^{(x)} \sim N(0, \sigma_x^2)$  and  $v_i \sim N(0, \sigma_v^2)$ . The criterion function is the prediction MSE given by

$$\text{PMSE}_n^*(h) = \frac{1}{h} \sum_{i=n-h+1}^n (y_i - \hat{x}_i^{(m)})^2 ,$$

where  $\hat{x}_i^{(m)}$  is the estimate of  $x_i^{(m)}$ .

We present the minimum  $m$  as  $m^*$  based on the trend-cycle filtering as Tables 6.1-6.3 by taking  $h = 2, 4, 6, 8$  and  $n = 80, 120, 200, 400$ . In our simulations, as  $n$  increases, we have larger choice of  $m^*$  while  $m^*$  decreases as  $h$  increases. When we have a long horizon with  $h$ , it may be natural to use a small number of lower frequencies, and for a wide range of  $\sigma_x$  and  $\sigma_v$ .

**Table 6.1 :** PMSE and Choice of  $m$   
( $\sigma_x = 0.3$ ,  $\sigma_v = 0.05$ )

n	80	120	200	400
h=2	12	19	32	65
h=3	8	12	21	42
h=4	6	9	15	32
h=5	5	7	13	26
h=6	4	6	10	21
h=7	4	5	9	18
h=8	3	4	8	16

**Table 6.2 :** PMSE and Choice of  $m$   
 $(\sigma_x = 0.3, \sigma_v = 0.4)$

n	80	120	200	400
h=2	13	19	33	66
h=3	8	13	21	43
h=4	6	9	16	32
h=5	5	8	13	26
h=6	4	6	11	21
h=7	4	5	9	18
h=8	3	5	8	16

**Table 6.3 :** PMSE and Choice of  $m$   
 $(\sigma_x = 0.3, \sigma_v = 1.0)$

n	80	120	200	400
h=2	13	20	33	67
h=3	9	13	22	44
h=4	6	10	16	33
h=5	5	8	13	26
h=6	4	7	11	22
h=7	4	6	9	19
h=8	4	5	8	16

### 6.3 A Comparison of the SIML Filtering and HP Filtering

In order to characterize some property of the SIML filtering, we give an illustrative example. For this purpose, we take a monthly original consumption data in Section 2 (Kakei-Chosa series in Figure 3) and show the result of the SIML filtering and the HP filtering with two parameter values ( $\lambda$ ) as Figure 4. (We used *hpfilter* procedure in mFilter R-package for HP filter and took  $m = 12$  for the SIML filter, which corresponds to frequencies over 2 year cycles since  $\lambda_{max}^{(n)} = 1/24$ .) An important point is that when we have strong seasonality in data, the SIML often gives reasonable estimates while the usefulness of HP filter depends on the the particular parameter value chosen for  $\lambda$ .

Further comparison would be interesting, but it is beyond the scope of this paper.

## 7 Applications

### 7.1 An interpretation of Müller and Watson (2018)

Recently, Müller and Watson (2018) have proposed the *so-called* long-run co-variability of macro-economic time series. They have investigated many non-stationary time

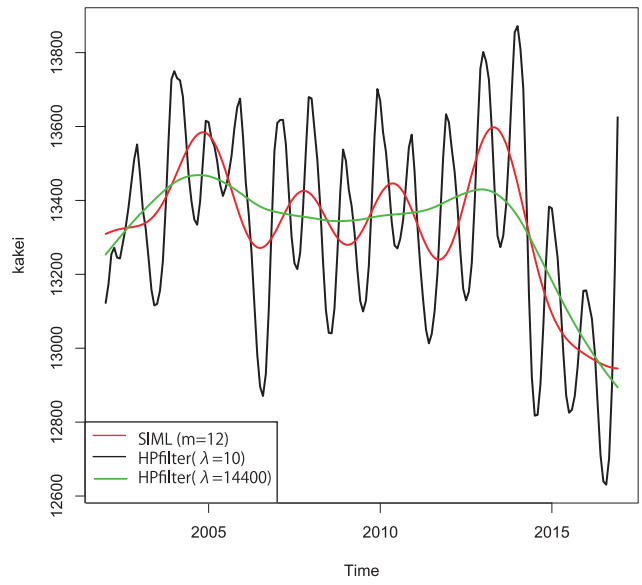


Figure 4: **An Example of SIML and HP Filters.** (An illustrated comparison of the SIML and HP Filters by using consumption data.)

series and obtained some empirical findings. In this sub-section, an interpretation of their method will be given as measuring the relationships among long-run trend-cycle in our framework when  $p = 2$ . Moreover, we obtain an important consequence from Proposition 2 in Section 5.1.

In this sub-section, we consider the case of  $p = 2$  in (3.9)-(3.13). Let  $2 \times 2$  matrices  $\Sigma_x = (\sigma_{ij}^{(x)})$ , we then define the (long-run) regression coefficient  $\beta = [\sigma_{22}^{(x)}]^{-1} \sigma_{21}^{(x)}$  under the assumption that  $\sigma_{22}^{(x)} > 0$ . Let also  $\mathbf{G}_m(0) = (g_{ij}^{(m)})$ , and an  $n \times 2$  matrix

$$(7.48) \quad (\mathbf{w}_1, \mathbf{w}_2) = \mathbf{C}_n^{-1}(\mathbf{Y}_n - \mathbf{Y}_0) .$$

For estimating  $\beta$ , we define

$$(7.49) \quad \hat{\beta} = [g_{22}^{(m)}]^{-1} g_{21}^{(m)} = [\mathbf{w}_2' \mathbf{P}_n \mathbf{J}_m \mathbf{J}_m' \mathbf{P}_n \mathbf{w}_2]^{-1} [\mathbf{w}_2' \mathbf{P}_n \mathbf{J}_m \mathbf{J}_m' \mathbf{P}_n \mathbf{w}_1] .$$

This quantity can be interpreted as the least squares slope of the transformed vector from  $\mathbf{y}_{1n}$  on the transformed vector from  $\mathbf{y}_{2n}$  for an  $n \times 2$  matrix  $\mathbf{Y}_n = (\mathbf{y}_{1n}, \mathbf{y}_{2n})$ , which is essentially the same as the one proposed by Müller and Watson (2018)<sup>7</sup>. Then, from Proposition 2 in Section 5 and its proof in the Appendix we immediately obtain the following result.

**Proposition 3** : In (3.9)-(3.13) with  $p = 2$ , we assume that the fourth order moments of each element of  $\mathbf{v}_i^{(x)}$ ,  $\mathbf{v}_i$  and  $\mathbf{v}_i^{(s)}$  ( $i = 1, \dots, n$ ) are bounded, and  $\Sigma_x$  is positive definite.

- (i) We fix an  $m$ , then  $\hat{\beta}$  is not consistent when  $n \rightarrow \infty$ .
- (ii) Set  $m_n = [n^\alpha]$  and  $0 < \alpha < 1$ , then as  $n \rightarrow \infty$

$$(7.50) \quad \hat{\beta} - \beta \xrightarrow{p} \mathbf{0} .$$

The second part of this proposition is an extended version of the first part of Theorem 4.1 of Kunitomo and Sato (2017).

### A Simulation Example

To illustrate our arguments in Proposition 3, we performed a set of Monte Carlo experiments under the simple situation with  $p = 2$  (the replication was 10,00 in each simulation). The model is given by

$$(7.51) \quad \mathbf{x}_i = \begin{pmatrix} x_{1,i} \\ x_{2,i} \end{pmatrix} = \begin{pmatrix} x_{1,i-1} \\ x_{2,i-1} \end{pmatrix} + \begin{pmatrix} v_{1,i}^{(x)} \\ v_{2,i}^{(x)} \end{pmatrix} ,$$

---

<sup>7</sup>In their notation,  $m$  corresponds to  $q$ , which is fixed. They did use the (differenced) stationary data, and thus, we could interpret that they calculated the linear regression from the filtered data  $\hat{\mathbf{X}}_n^* = \mathbf{P}_n' \mathbf{J}_m' \mathbf{J}_m \mathbf{P}_n \mathbf{C}_n^{-1}(\mathbf{Y}_n - \mathbf{Y}_0)$  as a modification of (4.19) in our notation.



$$(7.52) \quad \mathbf{y}_i = \begin{pmatrix} y_{1,i} \\ y_{2,i} \end{pmatrix} = \begin{pmatrix} x_{1,i} + \beta x_{2,i} \\ x_{2,i} \end{pmatrix} + \begin{pmatrix} v_{1,i} \\ v_{2,i} \end{pmatrix},$$

where we have generated the normal errors  $\mathbf{v}_i^{(x)'} = (v_{1,i}^{(x)}, v_{2,i}^{(x)})$ ,  $\mathbf{v}_i' = (v_{1,i}, v_{2,i})$ , ( $i = 1, \dots, n$ ) with zero means, variances  $\sigma_j^{(x)} = 0.75$  and  $\sigma_j^{(v)} = 1.5$  ( $j = 1, 2$ ), zero covariances and an initial value  $\mathbf{x}_0$ .

This is a two dimensional I(1)-process, which is not co-integrated, but close to a cointegrated-process. We present the finite sample properties of the (naive) least squares (LS) estimator from the original data, the Müller-Watson (MW) estimator, and the SIML (SIML) estimator from the transformed data.

The simulation results in Tables 7.1-7.3 are consistent with our arguments in Proposition 3. In the tables, Direct LS stands for the standard regression on  $\mathbf{y}_{1n}$  on  $\mathbf{y}_{2n}$ , while MW and SIML stand for Muller-Watson and SIML, respectively. When we have two-dimensional time series and they are not cointegrated, the standard least square method is badly biased. The procedure proposed by Muller and Watson (2018) often gives reasonable results, but its variance does not decrease as the sample size increases. The SIML estimation method has reasonable finite sample properties as well as reasonable asymptotic properties and it is applicable to more general cases.

**Table 7.1 :** Simulations (Direct LS)

n	200	1000	30000
true	1.2000	1.2000	1.2000
mean	2.2016	2.2287	2.0918
SD	0.0930	0.2183	1.4901

**Table 7.2 :** Simulations (MW)

n	200	1000	30000
true	1.2000	1.2000	1.2000
mean	1.1418	1.2129	1.1875
SD	0.5279	0.5301	0.5337

**Table 7.3 :** Simulations (SIML)

n	200	1000	30000
true	1.2000	1.2000	1.2000
mean	1.0828	1.1807	1.2020
SD	0.3278	0.1904	0.0766

## 7.2 Some Macro-economic data and Constructing a Monthly Consumption Index

As illustrations for the empirical analysis, we have used our filtering method to analyze Japanese quarterly (real) consumption-GDP data as the first example, and three sets of monthly consumption data as the second example, which have been discussed in Section 2. Appendix B contains all figures in this subsection.

As the first example, the ratio of real GDP and real consumption (quarterly original series) and their time series plots are given in Figures 1 and 2. They show non-stationarity in their trend-cycle and seasonality as typical macro-economic variables. We then calculate the transformed data using  $\mathbf{P}_n$  (a kind of real Fourier transformation) as Figure 5 from the original consumption data. In this case, the transformed series gives a wild up and down fluctuation, and trend-cycle and seasonality are sucked. To find seasonal components from the original series, we calculated the realized  $z_k$  ( $k = 1, \dots, n$ ) (Figure 6) and the empirical cumulative distribution of  $z_k^2$  ( $k = 1, \dots, n$ ) (Figure 7), which roughly correspond to the normalized and real-valued sample spectral distribution function. Each component of  $\mathbf{z}_k$  ( $k = 1, \dots, n$ ) is an orthogonal decomposition in the frequency domain. Because we have quarterly macro-data, we have a large up and down around ( $\lambda_k^{(n)} =$ ) 0.25, which corresponds to the seasonal frequency at  $s = 4$ . The empirical spectral distribution has an abrupt change at this frequency. From these figures, we can judge that the real Fourier transformation based on  $\mathbf{K}_n$  does give useful information.

As it has been a practice in time series data analysis to use seasonal differencing  $\Delta_s \mathbf{y}_i = \mathbf{y}_i - \mathbf{y}_{i-s}$  ( $s = 4$ ) in the Box-Jenkins method, we calculated the real Fourier transformation based on  $\mathbf{P}_n$  (Figure 8) after seasonal differencing. Although the contribution of the resulting orthogonal process around the seasonal frequency could be significant, there are some rather wild fluctuations on at many other frequencies by using  $\mathbf{P}_n$ -transformation. Because we have some difficulty in interpreting the resulting time series, it may not be possible to justify the seasonal differencing procedure, and we recommend not to use this representation. In the following analysis, we simply use the differencing and then use the frequency domain analysis.

In Figure 9, we have analyzed real Quarterly-GDP. We show one example with  $m = \lfloor n^{.99} \rfloor$  and the deleted seasonal frequency are around 48-52 (48/196-52/196 in  $[0, 1/2]$ ) and we delete extremely high-frequency part. We deleted only five transformed data around the seasonal frequency and the main intention was to investigate the effect of seasonality with a narrow band. We have taken  $\alpha = .99$  since we wanted to remove some contribution of high frequency, but we could have used other choices and the results are not much different from Figure 8 in most cases. We compared the filtered time series using our method and the official (published) seasonally adjusted time series. We found that the differences in two time series (i.e., the published time series and the SIML filtered time series) are rather small and they are often of negligible magnitude. Although our filtering procedure is simple, this empirical

example suggests the usefulness of our method developed in this study.

As the second example, we have analyzed three consumption (monthly) time series and the quarterly consumption time series, which are mentioned in Section 2. Then this is an empirical example when  $p = 4$  and  $s = 12$  with missing observations because the quarterly one series cannot give monthly information. As we have seen in Figure 3, three monthly consumption series have similarities and some differences in their components. In our example, our goal is to construct the monthly consumption index, which is consistent with the observed quarterly consumption time series in trend-cycle component. Because of non-stationary trend-cycle, seasonal and measurement errors, it may not be obvious to construct a useful consumption index using existing statistical tools.

Let  $Y_i$  ( $i = 1, \dots, n$ ) be the target (quarterly) time series and  $Z_{jt}$  ( $j = 1, 2, 3; t = 3(i-1) + l, l = 1, 2, 3$ ) be the  $j$ -th monthly time series ( $t = 0$  is the initial period and we fix the initial values  $Y_0, Z_{j0}$ ). Then the criterion function is

(7.53)

$$\text{MSE}(m, m_1, m_2, m_3, w_1, w_2, w_3) = \sum_{i=1}^n \left[ \Delta \hat{Y}_i^{(T)} - \sum_{j=1}^3 w_j \left( \sum_{l=1}^3 \Delta Z_{j,3(i-1)+l}^{(T)} \right) \right]^2,$$

where  $\Delta \hat{Y}_i^{(T)} = \hat{Y}_i^{(T)} - \hat{Y}_{i-1}^{(T)}$ , (the trend part of the estimated  $\Delta Y_i$  because we observe quarterly data on  $Y_i$ ),  $\Delta Z_{jt}^{(T)} = Z_{jt}^{(T)} - Z_{j,t-1}^{(T)}$  (the trend parts of  $\Delta Z_{jt}$ ), and  $w_j$  ( $j = 1, 2, 3$ ) are (unknown) weight coefficients and  $m, m_j$  ( $j = 1, 2, 3$ ) are the numbers of trend-cycle filtering. In the above formulation, we need to measure the prediction errors based on differenced data because we have non-stationary trend-cycle component.

Using the least squares method, we minimized the MSE criterion with respect to the underlying parameters. The estimated  $w_j$  ( $j = 1, 2, 3$ ) are 3.69, 5.19, and 1.64 (while the measurement units are different), but their magnitudes are comparable to the published quarterly consumption level at 2002Q1-2016Q4), which are statistically significant at 1%. We have chosen  $m = 29, m_1 = 36, m_2 = 23$  and  $m_3 = 33$ . Although it may be possible to use other possibilities, but in our limited experiments, we found some improvements in prediction error over other cases with different combinations of  $m$  and  $m_j$  ( $j = 1, 2, 3$ ).

The black curves are the original series and the red curves are the estimated trend curves in Figures 10 and 11 for two monthly series. By taking relatively large  $m_j$  ( $j = 1, 2, 3$ ), we can recover the cycle components of each series, which are crucial as the indicators of macro-business condition. In Figure 12, the green curve shows the predicted state value calculated from the latest observed (quarterly) data plus the predicted monthly part based on the estimated parameters. As there is no monthly observation of quarterly published consumption, we draw their latest (quarterly) level with the black curve and the estimated SIML (filtered) values with

the red curve <sup>8</sup>. Overall, we found that while our procedure is much simpler than the X-12-ARIMA seasonal adjustment with reg-ARIMA model, the estimated filtered series are consistent with the (published) quarterly series given the fact we have non-stationary trend-cycle, seasonal and measurement errors components even when we do not have large samples. In Figure 13. we have drawn the prediction errors in terms of the differenced value  $Y_i$  ( $i = 1, \dots, n$ ) based on our procedure. This figure illustrates the usefulness of the procedure because the macro-economic time series are non-stationary with measurement errors.

The empirical data analyses in this subsection are presented for illustrations. We are currently investigating the consumption data and the results will be reported on another occasion.

## 8 Concluding Remarks

When the observed non-stationary multivariate time series contain seasonality and noises, it may be difficult to disentangle the effects of trend-cycle, seasonal and measurement errors. In real (seasonally unadjusted) times series, we often observe non-stationary trend-cycle, seasonality and measurement errors while the X-12-ARIMA program in official agencies uses the univariate reg-ARIMA model to remove the seasonality from original time series. In this study we investigate a new procedure to decompose time series into non-stationary trend-cycle components, stationary seasonal and noise (or measurement errors) components. The resulting method for non-stationary multivariate series is simple and free from the underlying distributions of components. Hence, it is robust against possible misspecification in the non-stationary multivariate economic time series. An important conclusion is that it is useful to transform the observed time series using the  $\mathbf{K}_n$ -transformation and investigate the transformed  $\mathbf{Z}_n$  series.

In empirical example in Section 7, we have illustrated our method to analyze quarterly and monthly macro-consumption data in Japan. We presented a way to construct the monthly consumption index as the second example, which is consistent with the published or official (GDP-)consumption quarterly data. Although the problem is practically complicated, we have shown that our method gives a useful way for practical purpose.

There can be several further problems. Although it is easy to handle the  $\mathbf{K}_n$ -transformations of non-stationary multivariate time series and construct the transformed  $\mathbf{Z}_n$ -data, there is a non-trivial initial value problem of filtering. Another direction would be to handle some abrupt changes in multiple time series, such as the changes of consumption tax in Japan. Some progress on these problem has been

---

<sup>8</sup>One notable event was the introduction of consumption tax in April 2014 and a sharp deviation of trend. In the present study, we do not focus on this aspect and it will be in Sato and Kunitomo (2020b).

made (Sato and Kunitomo (2020a,b)), which will be reported in other occasions. As there are many important empirical applications, we need to develop a systematic statistical procedure.

## References

- [1] Anderson, T.W. (1971), *The Statistical Analysis of Time Series*, Wiley.
- [2] Brillinger, D. and M. Hatanaka (1969) "An Harmonic Analysis of Nonstationary Multivariate Economic Processes," *Econometrica*, Vol. 35, 131-141.
- [3] Brillinger, D. (1980) *Time Series : data analysis and theory*, Expanded Edition, Holden-Day.
- [4] Brockwell, P. and R. Davis (1990), *Time Series : Theory and Methods*, 2nd Edition, Wiley.
- [5] Baxter, H. and R. King, (1999) "Measuring Business Cycles : Approximate Band-Pass Filters for Economic Time Series," *Review of Economics and Statistics*, Vol.81-4, 575-593.
- [6] Christiano, L. and T. Fitzgerald (2003) "The Band Pass Filter," *International Economic Review*, Vol.44-2, 435-465.
- [7] Doob, J.L (1953), *Stochastic Processes*, Wiley.
- [8] Durrett, R. (1991), *Probability : Theory and Examples*, Duxbury Press.
- [9] Engle, R. and C.W.J. Granger (1987), "Co-integration and Error Correction," *Econometrica*, Vol.55, 251-276.
- [10] Findley, D., Monsell, B., Bell, W., Otto, M. and Chen, B. (1998),"New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program," *Journal of Business and Economic Statistics*, 16, 127-177.
- [11] Fuller, W. (1987), *Measurement Error Models*, John-Wiley.
- [12] Granger, C. W. J. and Hatanaka, M. (1964), *Spectral Analysis of Economic Time Series*, Princeton U.P.
- [13] Granger, C. W. J. and Newbold, P. (1977), *Forecasting Economic Time Series*, Academic Press.
- [14] Johansen, S. (1995), *Likelihood Based Inference in Cointegrated Vector Autoregressive Models*, Oxford UP.

- [15] Harvey, A. and T. Trimbur (2008), "Trend Estimation and The Hodrick-Prescott Filter," J. Japan Statistical Society, 38-1, 41-49.
- [16] Hayashi, F. (2000), *Econometrics*, Princeton University Press.
- [17] Kitagawa, G. (2010), *Introduction to Time Series Analysis*, CRC Press.
- [18] Kunitomo, N. , S. Sato and D. Kurisu (2018), *Separating Information Maximum Likelihood Estimation for High Frequency Financial Data*, Springer.
- [19] Kunitomo, N. and S. Sato (2017), "Trend, Seasonality and Economic Time Series : the Non-stationary Errors-in-variables Models," SDS-4, MIMS, Meiji University, <http://www.mims.meiji.ac.jp/publications/2017-ds>.
- [20] Kunitomo, N., N. Awaya and D. Kurisu (2019), "Some Properties of Estimation Methods for Structural Relationships in Non-stationary Errors-in-Variables Models," *Japanese Journal of Statistics and Data Science*, Springer (Open Access),
- [21] Morgenstern, Oskar (1950). *On the accuracy of economic observations*, (1966, revised edition), Princeton University Press.
- [22] Nerlove, M., D.M. Grether and J.L. Carvalho (1995), *Analysis of Economic Time Series : A Synthesis*, Revised Edition, Academic Press.
- [23] Müller, U. and M. Watson (2018), "Long-run Covariability," *Econometrica*, 86-3, 775-804.
- [24] Nishimura, G. K., Seisho Sato, Akihiko Takahashi (2019), "Term Structure Models During the Global Financial Crisis: A Parsimonious Text Mining Approach," *Asia-Pacific Financial Markets*.
- [25] Sato, S. and N. Kunitomo (2020a), "On Backward Filtering for Noisy Non-stationary Time Series," MIMS-RBP Statistics & Data Science Series (SDS-15), MIMS, Meiji University.
- [26] Sato, S. and N. Kunitomo (2020b), "Frequency Regression and Smoothing for Noisy Non-stationary Time Series," In preparation.
- [27] Yamada, H. (2019), "A Smoothing Method that looks like the Hodrick-Prescott Filter," Unpublished Manuscript.

## APPENDIX A : Mathematical Derivations

We now present some details of derivations that we have omitted in the previous sections. We refer to Kunitomo, Sato and Kurisu (2018) as KSK(2018).

(i) **On (4.24) and (4.27)** : When we take  $m_1 = 0$  and  $m_2 = m$ , then (4.27) reduces (4.24). Hence, we show (4.27).

For  $\theta_{jk} = \frac{2\pi}{2n+1}(j - \frac{1}{2})(k - \frac{1}{2})$  ( $j, k = 1, \dots, n$ ), we use the relation that

$$\theta_{jk} + \theta_{j',k} = \frac{2\pi}{2n+1}(j + j' - 1)(k - \frac{1}{2}), \quad \theta_{jk} - \theta_{j',k} = \frac{2\pi}{2n+1}(j - j')(k - \frac{1}{2}).$$

We then have

$$(A.54) \quad 4 \sum_{k \in \mathbf{I}_n^{(1)}} [\cos \theta_{jk} \cos \theta_{j',k}] \\ = \sum_{k \in \mathbf{I}_n^{(1)}} [e^{i(\theta_{jk} + \theta_{j',k})} + e^{-i(\theta_{jk} + \theta_{j',k})}] + \sum_{k \in \mathbf{I}_n^{(1)}} [e^{i(\theta_{jk} - \theta_{j',k})} + e^{-i(\theta_{jk} - \theta_{j',k})}],$$

where  $\mathbf{I}_n^{(1)} = [1, \dots, m]$  (or  $\mathbf{I}_n^{(2)} = [m_1 + 1, \dots, m_1 + m_2]$ ) is the index set for  $j$  and  $k$ .

For  $\mathbf{I}_n^{(2)} = [m_1 + 1, \dots, m_1 + m_2]$ , by rewriting

$$\theta_{jk} + \theta_{j',k} = (m_1 - \frac{1}{2}) \frac{2\pi}{2n+1}(j + j' - 1) + \frac{2\pi}{2n+1}(j + j' - 1)(k - m_1),$$

and

$$\theta_{jk} - \theta_{j',k} = (m_1 - \frac{1}{2}) \frac{2\pi}{2n+1}(j - j') + \frac{2\pi}{2n+1}(j - j')(k - m_1),$$

the summation of the first two terms in (A.54) becomes

$$e^{i(m_1 + \frac{1}{2}) \frac{2\pi}{2n+1}(j+j'-1)} \times \frac{1 - e^{i \frac{2\pi}{2n+1}(j+j'-1)m_2}}{1 - e^{i \frac{2\pi}{2n+1}(j+j'-1)}} + e^{-i(m_1 + \frac{1}{2}) \frac{2\pi}{2n+1}(j+j'-1)} \times \frac{1 - e^{-i \frac{2\pi}{2n+1}(j+j'-1)m_2}}{1 - e^{-i \frac{2\pi}{2n+1}(j+j'-1)}}.$$

For the last two terms in (A.54), we need to evaluate each term when (i)  $j = j'$  and (ii)  $j \neq j'$ , separately. Using similar calculations in (A.54) with the index set  $\mathbf{I}_n^{(2)}$ , when  $j \neq j'$ , the summation of last two terms becomes

$$e^{i(m_1 + \frac{1}{2}) \frac{2\pi}{2n+1}(j-j')} \times \frac{1 - e^{i \frac{2\pi}{2n+1}(j-j')m_2}}{1 - e^{i \frac{2\pi}{2n+1}(j-j')}} + e^{-i(m_1 + \frac{1}{2}) \frac{2\pi}{2n+1}(j-j')} \times \frac{1 - e^{-i \frac{2\pi}{2n+1}(j-j')m_2}}{1 - e^{-i \frac{2\pi}{2n+1}(j-j')}}.$$

When  $j = j'$ ,  $\theta_{jk} - \theta_{j',k} = 0$  and the summation of last two terms with the index set  $\mathbf{I}_n^{(2)}$  becomes  $2m_2$ . Hence, it is possible to evaluate each terms of (4.24) and (4.27).

Using the relation

$$\begin{aligned}
& e^{i(m_1+\frac{1}{2})\frac{2\pi}{2n+1}(j+j'-1)} \times \frac{1 - e^{i\frac{2\pi}{2n+1}(j+j'-1)m_2}}{1 - e^{i\frac{2\pi}{2n+1}(j+j'-1)}} \\
& + e^{-i(m_1+\frac{1}{2})\frac{2\pi}{2n+1}(j+j'-1)} \times \frac{1 - e^{-i\frac{2\pi}{2n+1}(j+j'-1)m_2}}{1 - e^{-i\frac{2\pi}{2n+1}(j+j'-1)}} . \\
= & \frac{e^{i\frac{2\pi}{2n+1}\frac{1}{2}(j+j'-1)(m_1)} - e^{i\frac{2\pi}{2n+1}\frac{1}{2}(j+j'-1)(m_1+m_2)}}{e^{i\frac{2\pi}{2n+1}(-\frac{1}{2})(j+j'-1)} - e^{i\frac{2\pi}{2n+1}(\frac{1}{2})(j+j'-1)}} \\
& + \frac{e^{-i\frac{2\pi}{2n+1}\frac{1}{2}(j+j'-1)(m_1)} - e^{-i\frac{2\pi}{2n+1}\frac{1}{2}(j+j'-1)(m_1+m_2)}}{e^{-i\frac{2\pi}{2n+1}(-\frac{1}{2})(j+j'-1)} - e^{-i\frac{2\pi}{2n+1}(\frac{1}{2})(j+j'-1)}}
\end{aligned}$$

and the corresponding results for  $j - j'$  (there are two cases when (a)  $j = j'$  and (b)  $j \neq j'$ ), we obtain the results of (4.27), and then (4.24) by using  $\mathbf{I}_n^{(1)}$  instead of  $\mathbf{I}_n^{(2)}$ .

(ii) **The proof of Proposition 1** : Essentially, we apply the Central Limit Theorem (Theorem 8.4.3 of Anderson (1971) or Theorem 7.6 of Durrett (1991)) to the sequence of ergodic stationary (discrete) time series. We will give the basic steps of our derivations and mention that the problem here is similar to those explained in Chapters 7-9 of Anderson (1971) in details.

First, we need to show that the resulting variance-covariance terms correspond to those of the limiting Gaussian random variables.

For this purpose, we need to evaluate

(A.55)

$$\mathbf{E} \left[ \mathbf{z}_n(\lambda_k^{(n)}) \mathbf{z}_n(\lambda_{k'}^{(n)})' \right] = \left[ \frac{1}{2n+1} \right] \sum_{j,j'=1}^n (e^{i\theta_{jk}} + e^{-i\theta_{jk}})(e^{i\theta_{j'k'}} + e^{-i\theta_{j'k'}}) \mathbf{E}[\mathbf{r}_j \mathbf{r}_{j'}'] .$$

When  $k \neq k'$ , we find that the right-hand side terms are bounded by using the straightforward calculations. We notice that the right-hand side consists of sums of four terms associated with

$$\begin{aligned}
\text{(A.56)} \quad & (e^{i\theta_{jk}} + e^{-i\theta_{jk}})(e^{i\theta_{j'k'}} + e^{-i\theta_{j'k'}}) \\
= & e^{i(\theta_{jk}+\theta_{j'k'})} + e^{-i(\theta_{jk}+\theta_{j'k'})} + e^{i(\theta_{jk}-\theta_{j'k'})} + e^{-i(\theta_{jk}-\theta_{j'k'})} \\
= & (A) + (B) + (C) + (D) \text{ (, say) .}
\end{aligned}$$

Then we find that the sums of each terms associated with (A) and (B) in (A.55) are bounded, which becomes small when  $n$  is large. We write

$$\text{(A.57)} \quad \sum_{j,j'=1}^n e^{i(\theta_{jk}+\theta_{j'k'})} \mathbf{E}[\mathbf{r}_j \mathbf{r}_{j'}'] = \sum_{h=-(n-1)}^{n-1} \sum_{j' \in S(h)} e^{i(\theta_{h+j',k}+\theta_{j'k'})} \mathbf{\Gamma}(h) ,$$



where  $S(h) = \{1, 2, \dots, n-h\}$  for  $h \geq 0$  and  $S(h) = \{1-h, 2-h, \dots, n\}$  for  $h < 0$ . When  $h \geq 0$  given  $h$ , the second sum is given as

$$\begin{aligned} & \sum_{j'=1}^{n-h} e^{i\frac{2\pi}{2n+1}(h+j'-\frac{1}{2})(k-\frac{1}{2})} e^{i\frac{2\pi}{2n+1}(j'-\frac{1}{2})(k'-\frac{1}{2})} \times \mathbf{\Gamma}(h) \\ = & \left[ \sum_{j'=1}^{n-h} e^{i\frac{2\pi}{2n+1}(k+k'-1)(j'-1)} \right] e^{i\frac{2\pi}{2n+1}(h+1/2)(k-\frac{1}{2})+\frac{1}{2}(k'-\frac{1}{2})} \times \mathbf{\Gamma}(h) . \end{aligned}$$

When  $h < 0$  given  $h$ , the sum can be written

$$\sum_{j'+h-1=0}^{n+h-1} e^{i\frac{2\pi}{2n+1}[(h+j'-1)(k+k'-1)+\frac{1}{2}(k-\frac{1}{2})-(h+\frac{1}{2})(k'-\frac{1}{2})]} \times \mathbf{\Gamma}(h) .$$

Because these sums are finite and we have the condition  $\sum_{h=-\infty}^{+\infty} \|\mathbf{\Gamma}(h)\| < +\infty$ , the sum with (A) and (B) in (A.55) become arbitrarily small when  $n$  is large. (The terms with (B) are the same as those with (A) except their signs in the exponential parts.)

Second, we need to show that the sums of each terms with (C) and (D) in (A.55) when  $k = k'$  are dominant terms. We utilized the relation

$$(A.58) \quad \sum_{j,j'=1}^n e^{i(\theta_{jk}-\theta_{j'k'})} \mathbf{E}[\mathbf{r}_j \mathbf{r}'_{j'}] = \sum_{h=-(n-1)}^{n-1} \sum_{j' \in S(h)} e^{i(\theta_{h+j',k}-\theta_{j'k'})} \times \mathbf{\Gamma}(h) ,$$

When  $h \geq 0$  given  $h$ , the second sum is given as

$$\sum_{j'=1}^{n-h} e^{i\frac{2\pi}{2n+1}(j'-1)(k-k')} e^{i\frac{2\pi}{2n+1}[(h+\frac{1}{2})(k-\frac{1}{2})-\frac{1}{2}(k'-\frac{1}{2})]} \times \mathbf{\Gamma}(h)$$

and when  $h < 0$  given  $h$ , the sum can be written

$$\sum_{j'+h-1=0}^{n+h-1} e^{i\frac{2\pi}{2n+1}[(h+j'-1)(k-k')+\frac{1}{2}(k-\frac{1}{2})-(h-\frac{1}{2})(k'-\frac{1}{2})]} \times \mathbf{\Gamma}(h) .$$

The sums of the above terms are bounded when  $k \neq k'$  by using the same argument as we have (A.57). (The terms with (D) are the same as those with (C) except their signs in the exponential parts.) On the other hand, when  $k = k'$ ,  $\sum_{j'=1}^{n-h} e^{i2\pi\frac{k-k'}{2n+1}(j'-1)} = n-h$ . Then the dominant sums with (C) and (D) in (A.55) become

$$(A.59) \quad \left[ \frac{n}{2n+1} \right] \sum_{h=-(n-1)}^{n-1} \left[ \cos 2\pi \frac{k-1/2}{2n+1} h \right] [\mathbf{\Gamma}(h) + \mathbf{\Gamma}(-h)] .$$

We note that under the assumption of stationarity of  $\mathbf{r}_j$ , it has been known that  $\frac{1}{n} \sum_{j'=1}^n \mathbf{r}_{h+j'} \mathbf{r}_{j'}' \xrightarrow{p} \Gamma(h)$ . (See Chapter 8 of Anderson (1971), and Brockwell and Davis (1990), for instance.)

Third, we consider the situation that  $\lambda_k^{(n)} \rightarrow s$ ,  $\lambda_{k'}^{(n)} \rightarrow t$  as  $n \rightarrow \infty$  for  $0 < s < t < \frac{1}{2}$ . Since  $\sum_{h=-\infty}^{+\infty} \|\Gamma(h)\| < +\infty$  and  $\|\Gamma(h)\|$  is small as  $h$  is large, in the situation that  $\lambda_k^{(n)} \rightarrow \lambda$  for  $0 < \lambda < \frac{1}{2}$  as  $n \rightarrow \infty$ , we have (5.31). (We can take  $k$  such that  $k/(2n) = \lambda + o(1/n^\epsilon)$  ( $\epsilon > 0$ ) such that  $\sum_{|h|>n^\epsilon} \|\Gamma(h)\|$  is arbitrary small.)

For the asymptotic normality, we set a sequence of random variables

$$(A.60) \quad W_{k,k'}^{(n)} = (\boldsymbol{\alpha}'_1, \boldsymbol{\alpha}'_2) \begin{bmatrix} \mathbf{Z}'_n \mathbf{e}_k^{(n)} \\ \mathbf{Z}'_n \mathbf{e}_{k'}^{(n)} \end{bmatrix}, k, k' = 1, \dots, n,$$

and  $\mathbf{Z}_n = \mathbf{P}_n \mathbf{R}_n$ ,  $\mathbf{R}_n = (\mathbf{r}'_j)$  ( $n \times p$  matrix), where  $\boldsymbol{\alpha}_i$  ( $i = 1, 2$ ) are  $p \times 1$  (non-zero) constant vectors and  $\mathbf{e}_k^{(n)} = (0, \dots, 0, 1, 0, \dots, 0)'$  ( $k = 1, \dots, n$ ) are  $n \times 1$  unit vectors.

We can also construct a sequence of random variables  $W_{s,t}^{(n)}$  by using

$$p_{s,j}^{(n)} = \sqrt{2/(2n+1)} \cos 2\pi s(j-1/2) \text{ instead of } p_{jk}^{(n)} \text{ in } \mathbf{P}_n.$$

Then  $\mathbf{E}[|W_{k,k'}^{(n)} - W_{s,t}^{(n)}|^2] \xrightarrow{p} 0$  as  $n \rightarrow \infty$  ( $\lambda_k^{(n)} \rightarrow s$ ,  $\lambda_{k'}^{(n)} \rightarrow t$  ( $0 < s < t < \frac{1}{2}$ )).

Finally, we utilize the truncation argument used in the proof of Theorem 8.4.3 of Anderson (1971) and approximate (5.30) by the (finite order) moving-average representation for a large  $h > 0$

$$(A.61) \quad \mathbf{r}_i^{(n,h)} = \boldsymbol{\mu} + \sum_{j=0}^h \boldsymbol{\Gamma}_j \mathbf{w}_{i-j}$$

and  $\mathbf{z}_{n,h}(\lambda_k^{(n)}) = \sum_{j=1}^n p_{jk}^{(n)} \mathbf{r}_j^{(n,h)}$ . By applying the standard method of proof in Theorem 8.4.3 of Anderson (1971), i.e., first we fix a  $h > 0$  and we show the CLT for  $\mathbf{z}_{n,h}(\lambda_k^{(n)})$  and then by taking  $h \rightarrow \infty$  and we can show that the effects of  $h$  are negligible because  $\sum_{s=-\infty}^{+\infty} \|\Gamma(s)\| < +\infty$ . (**Q.E.D.**)

(iii) **The proof of Proposition 2**: Let  $\mathbf{z}_k^{(x)} = (z_{kj}^{(x)})$  and  $\mathbf{Z}_k^{(s+v)} = (z_{kj}^{(s+v)})$  ( $k = 1, \dots, n$ ) be the  $k$ -th row vector elements of  $n \times p$  matrices

$$(A.62) \quad \mathbf{Z}_n^{(x)} = \mathbf{K}_n^* (\mathbf{X}_n - \bar{\mathbf{Y}}_0), \mathbf{Z}_n^{(s+v)} = \mathbf{K}_n^* (\mathbf{S}_n + \mathbf{V}_n), \mathbf{K}_n^* = \mathbf{P}_n \mathbf{C}_n^{-1},$$

respectively, where we denote  $\mathbf{X}_n = (\mathbf{x}'_k) = (x_{kg})$ ,  $\mathbf{S}_n = (\mathbf{s}'_k) = (s_{kg})$ ,  $\mathbf{V}_n = (\mathbf{v}'_k) = (v_{kg})$ ,  $\mathbf{Z}_n = (\mathbf{z}'_k) = (z_{kg})$  are  $n \times p$  matrices with  $z_{kg} = z_{kg}^{(x)} + z_{kg}^{(s+v)}$ . We write  $z_{kg}, z_{kg}^{(x)}, z_{kg}^{(s+v)}$  as the  $g$ -th component of  $\mathbf{z}_k, \mathbf{z}_k^{(x)}, \mathbf{z}_k^{(s+v)}$  ( $k = 1, \dots, n; g = 1, \dots, p$ ). We use  $z_{kg}^{(f)}$  ( $f = x, s+v$ ) and decompose  $\hat{\Sigma}_x - \Sigma_x = (\hat{\sigma}_{gh}^{(x)} - \sigma_{gh}^{(x)})_{gh}$  for  $g, h = 1, \dots, p$ .

We re-write

$$\begin{aligned}
(A.63) \quad & \frac{1}{m_n} \sum_{k=1}^{m_n} \mathbf{z}_k \mathbf{z}_k' - \Sigma_x \\
&= \left[ \frac{1}{m_n} \sum_{k=1}^{m_n} \mathbf{z}_k^{(x)} \mathbf{z}_k^{(x)'} - \Sigma_x \right] + \frac{1}{m_n} \sum_{k=1}^{m_n} \mathbf{E}[\mathbf{z}_k^{(s+v)} \mathbf{z}_k^{(s+v)'}] \\
&\quad + \frac{1}{m_n} \sum_{k=1}^{m_n} \left[ \mathbf{z}_k^{(s+v)} \mathbf{z}_k^{(s+v)'} - \mathbf{E}[\mathbf{z}_k^{(s+v)} \mathbf{z}_k^{(s+v)'}] \right] + \frac{1}{m_n} \sum_{k=1}^{m_n} \left[ \mathbf{z}_k^{(x)} \mathbf{z}_k^{(s+v)'} + \mathbf{z}_k^{(s+v)} \mathbf{z}_k^{(x)'} \right].
\end{aligned}$$

Then we will investigate the conditions that three terms except the first one of (A.63) are  $o_p(1)$  and

$$(A.64) \quad \frac{1}{m_n} \sum_{k=1}^{m_n} \mathbf{z}_k^{(x)} \mathbf{z}_k^{(x)'} - \Sigma_x \xrightarrow{p} \mathbf{O}$$

as  $m_n \rightarrow \infty$  ( $n \rightarrow \infty$ ). Since  $\Delta \mathbf{x}_i$  is stationary, it is not difficult to show the second assertion via Chapter 5 of KSK(2018). By using (A.59), the expected (matrix) value of  $(1/m) \sum_{k=1}^{m_n} \mathbf{z}_k \mathbf{z}_k'$  is given by

$$\frac{n}{2n+1} \sum_{h=-(n-1)}^{n-1} \left[ \frac{1}{m} \sum_{k=1}^n \cos 2\pi \frac{k-1/2}{2n+1} h \right] [\Gamma(h) + \Gamma(-h)].$$

Then by Lemma 5.1 of KSK(2018), the middle sum is  $[1/2m] \sin[2\pi mh/(2n+1)] / \sin[\pi h/(2n+1)]$ , which becomes 1 as  $m \rightarrow \infty, m/n \rightarrow 0$  for a fixed  $h$ . Then the limit is  $f_{SR}(0) = \sum_{-\infty}^{+\infty} \Gamma(h)$  because it is finite. Here we use the notation  $c_{ij}$  ( $i, j = 1, \dots, n$ ) of Chapter 5 of KSK(2018) as  $c_{ij} = (2/m) \sum_{k=1}^m \theta_{ik} \theta_{jk}$ . Then for (non-zero  $p \times 1$ ) constant vector  $\boldsymbol{\alpha}$  we can evaluate

$$\begin{aligned}
\mathbf{E} \left[ \frac{1}{m_n} \sum_{k=1}^{m_n} (\boldsymbol{\alpha}' \mathbf{z}_k)^2 - \boldsymbol{\alpha}' \Sigma_x \boldsymbol{\alpha} \right]^2 &= \left( \frac{2}{2n+1} \right)^2 \mathbf{E} \left[ \sum_{j,j'=1}^n c_{jj'} \boldsymbol{\alpha}' (\mathbf{r}_j \mathbf{r}_{j'}' - \mathbf{E}(\mathbf{r}_j \mathbf{r}_{j'}')) \boldsymbol{\alpha} \right]^2 \\
&\leq K_1 \left[ \frac{2}{2n+1} \right]^2 \sum_{j,j'=1}^n c_{j,j'}^2,
\end{aligned}$$

where  $K_1$  is a positive constant and we used the boundedness of fourth moments. Again by using Lemma 5.2 of KSK(2018), we find that  $\sum_{j,j'=1}^n c_{j,j'}^2 = (n+1/2)^2/m$  and we have shown the first assertion as  $m \rightarrow \infty, m/n \rightarrow 0$ .

Next we show that the second, third and fourth terms in the right-hand-side of (A.63) are asymptotically negligible as  $n \rightarrow \infty$ . and we could estimate the variance

and covariance of the underlying processes consistently as if there were no seasonal and noise terms

Let  $\mathbf{b}_k = (b_{kj}) = \mathbf{e}'_k \mathbf{P}_n \mathbf{C}_n^{-1} = (b_{kj})$  and  $\mathbf{e}_k^{(n)'} = (0, \dots, 1, 0, \dots)$  be an  $n \times 1$  vector. We write  $z_{kg}^{(s+v)} = \sum_{j=1}^n b_{kj}(s_{jg} + v_{jg})$  for the seasonal and noise part and use the relation

$$(A.65) \quad (\mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{C}_n'^{-1} \mathbf{P}'_n)_{k,k'} = \delta(k, k') 4 \sin^2 \left[ \frac{\pi}{2n+1} \left( k - \frac{1}{2} \right) \right].$$

Then under the conditions that  $\|\mathbf{C}_j^{(s)}\| = O(\rho^j)$  ( $0 \leq \rho < 1$ ) and  $\|\mathbf{C}_j^{(v)}\| = O(\rho^j)$  ( $0 \leq \rho < 1$ ), we can find  $K_2$  (a positive constant) such that

$$(A.66) \quad \mathbf{E}[(z_{kg}^{(s+v)})]^2 = \mathbf{E} \left[ \sum_{i=1}^n b_{ki}(v_{ig}^{(s)} + v_{ig}) \sum_{j=1}^n b_{kj}(v_{jg}^{(s)} + v_{jg}) \right] \leq K_2 \times a_{kn}^*.$$

It is because

$$\mathbf{E}[(z_{kg}^{(s+v)})]^2 = \sum_{i,j=1}^n b_{ki} b_{kj} [\sigma_{gg}^{(v)}(i-j) + \sigma_{gg}^{(s)}(i-j)],$$

where  $\sigma_{gg}^{(v)}(i-j) + \sigma_{gg}^{(s)}(i-j)$  is the  $(i-j)$ -th autocovariance of  $v_{ig} + v_{ig}^{(s)}$  and  $v_{jg} + v_{jg}^{(s)}$ . (We denote  $b_{ki} = 0$  for  $i < 0$  and  $i > n$ .) Then

$$\begin{aligned} \mathbf{E}[(z_{kg}^{(s+v)})]^2 &= \sum_{l=-(n-1)}^{n-1} \left[ \sum_{j=1}^n b_{kj} b_{k,j+l} (\sigma_{gg}^{(v)}(l) + \sigma_{gg}^{(s)}(l)) \right] \\ &\leq \left[ \sum_{j=1}^n b_{kj}^2 \right] \sum_{l=-\infty}^{\infty} [|\sigma_{gg}^{(v)}(l)| + |\sigma_{gg}^{(s)}(l)|]. \end{aligned}$$

Because  $\|\mathbf{C}_j^{(s)}\| = O(\rho^j)$  and  $\|\mathbf{C}_j^{(v)}\| = O(\rho^j)$ ,  $\sum_{l=-\infty}^{\infty} [|\sigma_{gg}^{(v)}(l)| + |\sigma_{gg}^{(s)}(l)|]$  is bounded. Also it is straight-forward to find that

$$\frac{1}{m_n} \sum_{k=1}^{m_n} a_{kn}^* = \frac{1}{m_n} 2 \sum_{k=1}^{m_n} \left[ 1 - \cos \left( \pi \frac{2k-1}{2n+1} \right) \right] = O \left( \frac{m_n^2}{n^2} \right),$$

by using the relation

$$\sum_{k=1}^m 2 \cos \left( \pi \frac{2k-1}{2n+1} \right) = \sum_{k=1}^m [e^{i \frac{2\pi}{2n+1} (k-\frac{1}{2})} + e^{-i \frac{2\pi}{2n+1} (k-\frac{1}{2})}] = \frac{\sin \left( \frac{2\pi}{2n+1} m \right)}{\sin \left( \frac{\pi}{2n+1} \right)}.$$

Then the second term of (A.63) becomes

$$(A.67) \quad \frac{1}{m_n} \sum_{k=1}^{m_n} \mathbf{E}[z_{kg}^{(s+v)}]^2 \leq K_3 \frac{1}{m_n} \sum_{k=1}^{m_n} a_{kn}^* = O \left( \frac{m_n^2}{n^2} \right),$$

which is  $o(1)$  if we set  $\alpha$  such that  $0 < \alpha < 1$  and  $K_3$  is a positive constant. (The arguments here are similar to the derivations in Chapter 5 of KSK(2018), but there is a major difference on the conditions because their asymptotic arguments are based on the situation when the length of observed intervals decrease. In the present situation, there is no  $\sqrt{n}$  factor in  $\mathbf{K}_n$ -transformation and we use  $\mathbf{K}_n^*$  and  $a_{kn}^*$  while they have used  $a_{kn} = na_{kn}^*$ .) For the fourth term,

$$\mathbf{E} \left[ \frac{1}{m_n} \sum_{j=1}^{m_n} z_{kg}^{(x)} z_{kg}^{(v)} \right]^2 = \frac{1}{m_n^2} \sum_{k,k'=1}^{m_n} \mathbf{E} \left[ z_{kg}^{(x)} z_{k',g}^{(x)} z_{kg}^{(v)} z_{k',g}^{(v)} \right] = O\left(\frac{m_n}{n^2}\right).$$

In the above evaluation we have used the evaluation that if we set  $s_{jk} = \cos[\frac{2\pi}{2n+1}(j - \frac{1}{2})(k - \frac{1}{2})]$  ( $j, k = 1, 2, \dots, n$ ), then we have the relation

$$\left| \sum_{j=1}^n s_{jk} s_{j,k'} \right| \leq \left[ \sum_{j=1}^n s_{jk}^2 \right] = \frac{n}{2} + \frac{1}{4} \text{ for any } k \geq 1.$$

(See Chapter 5 of Kunitomo, Sato and Kurisu (2018).)

Finally, for the third term, we need to consider the variance of

$$\begin{aligned} & (z_{kg}^{(s+v)})^2 - \mathbf{E}[(z_{kg}^{(s+v)})^2] \\ &= \sum_{j,j'=1}^n b_{kj} b_{k,j'} \left[ (v_{jg} + v_{jg}^{(s)})(v_{j',g} + v_{j',g}^{(s)}) - \mathbf{E}[(v_{jg} + v_{jg}^{(s)})(v_{j',g} + v_{j',g}^{(s)})] \right]. \end{aligned}$$

Then by using the assumption we have after some evaluations, we find a positive constant  $K_4$  such that

$$\begin{aligned} \text{(A.68)} \quad & \mathbf{E} \left[ \frac{1}{m_n} \sum_{k=1}^{m_n} ((z_{kg}^{(s+v)})^2 - \mathbf{E}[(z_{kg}^{(s+v)})^2]) \right]^2 \\ &= \frac{1}{m_n^2} \sum_{k_1, k_2=1}^{m_n} \mathbf{E} \left[ \sum_{j_1, j_2, j_3, j_4=1}^n b_{k_1, j_1} b_{k_1, j_2} \left( (v_{j_1, g}^{(s)} + v_{j_1, g})(v_{j_2, g}^{(s)} + v_{j_2, g}) \right. \right. \\ & \quad \left. \left. - \mathbf{E}(v_{j_1, g}^{(s)} + v_{j_1, g})(v_{j_2, g}^{(s)} + v_{j_2, g}) \right) \right. \\ & \quad \left. \times b_{k_2, j_3} b_{k_2, j_4} \left( (v_{j_3, g}^{(s)} + v_{j_3, g})(v_{j_4, g}^{(s)} + v_{j_4, g}) - \mathbf{E}(v_{j_3, g}^{(s)} + v_{j_3, g})(v_{j_4, g}^{(s)} + v_{j_4, g}) \right) \right] \\ &\leq K_4 \frac{1}{m_n^2} \left[ \sum_{k=1}^{m_n} a_{kn}^* \right]^2 = O\left(\frac{1}{m_n^2} \times \left(\frac{m_n^3}{n^2}\right)^2\right), \end{aligned}$$

which is  $O(m_n^4/n^4)$  by straight-forward calculations.

Here we just give an illustration of our derivations when  $p = 1$  and  $\mathbf{s}_i = 0$  ( $i =$

$1, \dots, n$ ). We re-write (3.12) as  $v_i = \sum_{j=0}^{\infty} c_j e_{i-j}^{(v)}$  and we evaluate

$$\begin{aligned} & \left(\frac{1}{m}\right)^2 \sum_{k_1, k_2=1}^m \sum_{j_1, j_2, j_3, j_4} b_{k_1, j_1} b_{k_1, j_2} b_{k_2, j_3} b_{k_2, j_4} \\ & \quad \times \mathbf{E}\{[v_{j_1} v_{j_2} - \mathbf{E}(v_{j_1} v_{j_2})][v_{j_3} v_{j_4} - \mathbf{E}(v_{j_3} v_{j_4})]\} \\ = & \left(\frac{1}{m}\right)^2 \sum_{k_1, k_2=1}^m \sum_{j_1, j_2, j_3, j_4} b_{k_1, j_1} b_{k_1, j_2} b_{k_2, j_3} b_{k_2, j_4} \sum_{l_1, l_2, l_3, l_4=0}^{\infty} c_{l_1} c_{l_2} c_{l_3} c_{l_4} \\ & \quad \times \mathbf{E}\{[e_{j_1-l_1} e_{j_2-l_2} - \mathbf{E}(e_{j_1-l_1} e_{j_2-l_2})][e_{j_3-l_3} e_{j_4-l_4} - \mathbf{E}(e_{j_3-l_3} e_{j_4-l_4})]\}. \end{aligned}$$

We need to evaluate the corresponding terms for four cases when (i)  $j_1 - l_1 = j_2 - l_2 = j_3 - l_3 = j_4 - l_4$ , (ii)  $j_1 - l_1 = j_2 - l_2 \neq j_3 - l_3 = j_4 - l_4$ , (iii)  $j_1 - l_1 = j_3 - l_3 \neq j_2 - l_2 = j_4 - l_4$ , (iv)  $j_1 - l_1 = j_4 - l_4 \neq j_2 - l_2 = j_3 - l_3$ . By using the condition in the general case that  $\|\mathbf{C}_j^{(s)}\| = O(\rho^j)$ ,  $\|\mathbf{C}_j^{(v)}\| = O(\rho^j)$  ( $j \geq 0, 0 \leq \rho < 1$ ), we have  $\sum_{j=0}^{\infty} |c_j| < \infty$  in this special case. We also utilize the relation such as  $\sum_{j=1}^n b_{kj} b_{k'j} = \delta(k, k') a_{kn}^*$  in the general case and we have used the notation that  $b_{k,j} = 0$  for  $k = 1, \dots, m$ ,  $j < 0, j > n$  and  $c_j = 0$  ( $j < 0$ ).

Then in each (i)-(iv) case, we can take a positive constant  $K_5$  such that (A.68) is less than

$$K_5 \left(\frac{1}{m}\right)^2 \sum_{k_1, k_2=1}^m \left[\sum_{j_1=1}^n b_{k_1, j_1}^2\right]^{1/2} \left[\sum_{j_2=1}^n b_{k_1, j_2}^2\right]^{1/2} \left[\sum_{j_3=1}^n b_{k_2, j_3}^2\right]^{1/2} \left[\sum_{j_4=1}^n b_{k_2, j_4}^2\right]^{1/2}.$$

The above evaluation method works in the general case with a complication of notations. Therefore, by using (68), the third term of (A.63) is negligible if we set  $\alpha$  such that  $0 < \alpha < 1$ . **(Q.E.D)**

## APPENDIX B : Some Figures

We now present figures used in Section 7. As we have explained in Section 2, all data are official data published by ESRI (Economic and Social Research Institute), Cabinet Office of Japan, METI, and Statistics Bureau, Ministry of Internal Affairs and Communications. They are available from the government official website : e.stat.

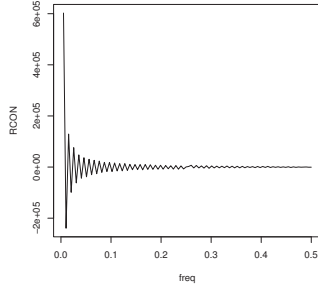


Figure 5: **Orthogonal Process  $\mathbf{z}_k$  for Original Series.** (Without differencing  $\mathbf{C}_n^{-1}$  we used the transformation  $\mathbf{P}_n$  of data and then calculated  $\mathbf{z}_k$  ( $k = 1, \dots, n$ ). Data are the quarterly real consumption between 1994Q1-2018Q2, published in 2018 by the Economic Social Research Institute (ESRI), Cabinet Office, Japan.)

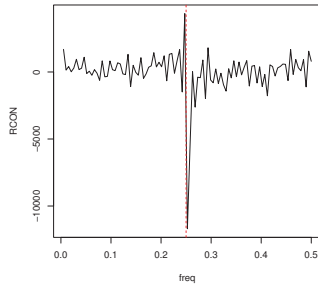


Figure 6: **Orthogonal Process  $\mathbf{z}_k$  for Quarterly Consumption.** (We used the transformation  $\mathbf{K}_n$  of data and then calculated  $\mathbf{z}_k$  ( $k = 1, \dots, n$ ). Data are the quarterly real consumption between 1994Q1-2018Q2, published in 2018 by the Economic Social Research Institute (ESRI), Cabinet Office, Japan.)

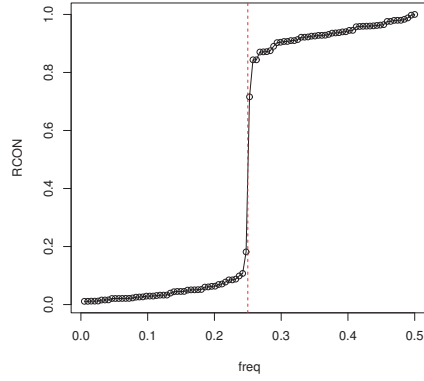


Figure 7: **Squared Orthogonal Process  $z_k^2$  for Quarterly Consumption.** (We calculated the cumulative sum of squares of the transformed orthogonal  $z_n^2$  from the data.)

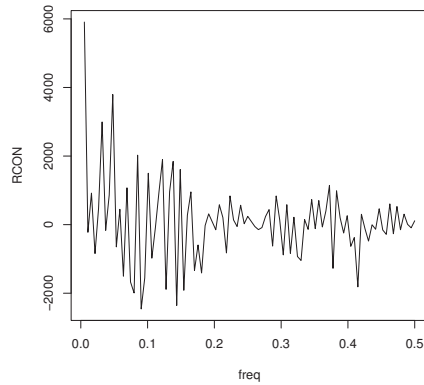


Figure 8: **Orthogonal Process  $z_k$  for Seasonally Differenced Consumption.** (After seasonal differencing instead of  $C_n^{-1}$ , we used the transformation  $\mathbf{P}_n$  on the differenced data and calculated the orthogonal elements.)



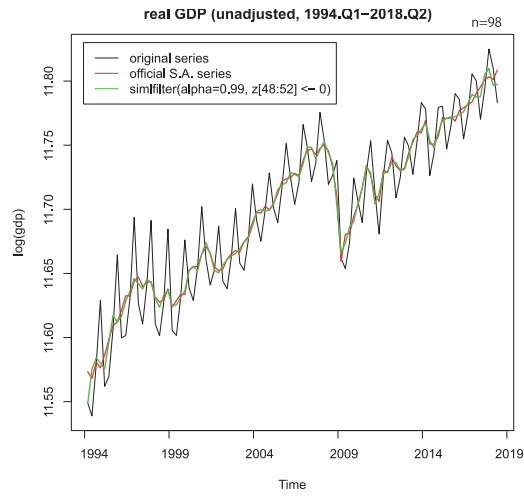


Figure 9: **Seasonally Adjusted Series.** (The official seasonally adjusted (S.A.) series are constructed using X-12-ARIMA and we compared the SIML filtered value. Data are the quarterly real GDP between 1994Q1-2018Q2, published in 2018 by the Economic Social Research Institute (ESRI), Cabinet Office, Japan.)

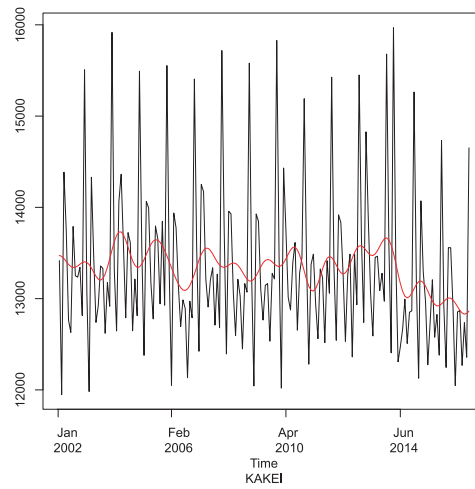


Figure 10: **Monthly Consumption Series.** (We compared the original series and the SIML filtered value. Data are the monthly real consumption between 2002M1-2016M12, published in 2017 by the Statistics Bureau, Japan.)

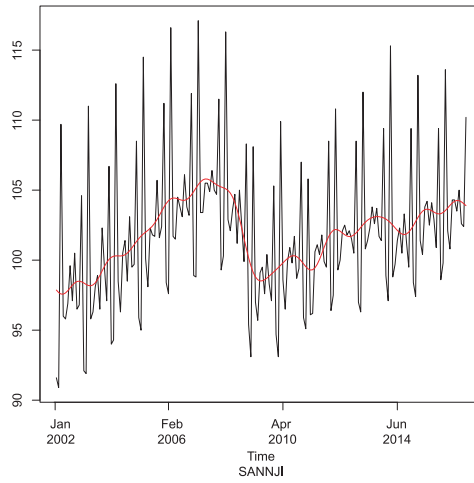


Figure 11: **Monthly Consumption Series.** (We compared the original series and the SIML filtered value. Data are the monthly consumptions between 2002M1-2016M12, published in 2017 by Ministry of Economy, Trade and Industry (METI), Japan.)

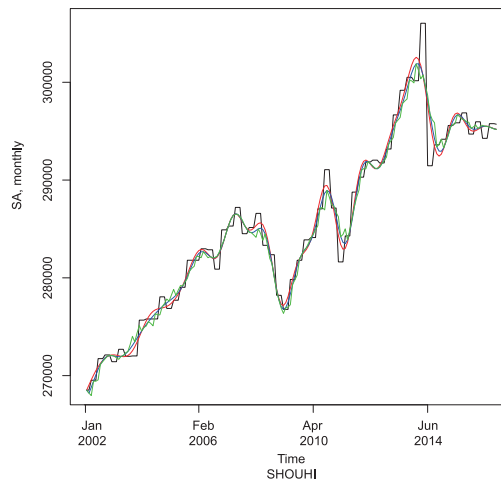


Figure 12: **Consumption Series.** (Data are the quarterly real consumption between 1994Q1-2018Q2, published in 2018 by the Economic Social Research Institute (ESRI), Cabinet Office, Japan.)

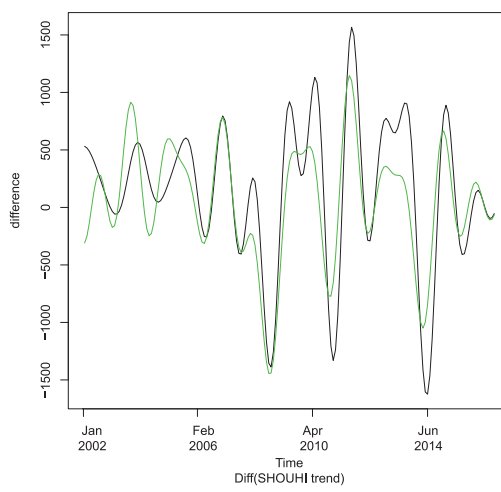


Figure 13: **Predicted Series.** (The predicted and monthly consumption state variable are shown between 1994Q1-2018Q2.)