CARF Working Paper

CARF-F-519

Frequency Regression and Smoothing for Noisy Nonstationary Time Series

Seisho Sato
Graduate School of Economics, University of Tokyo

Naoto Kunitomo
Tokyo Keizai University

August 24, 2021

CARF is presently supported by Nomura Holdings, Inc., Sumitomo Mitsui Banking Corporation, The Dai-ichi Life Insurance Company, Limited, The Norinchukin Bank, MUFG Bank, Ltd. and Ernst & Young ShinNihon LLC. This financial support enables us to issue CARF Working Papers.

Frequency Regression and Smoothing for Noisy Nonstationary Time Series *

Seisho Sato †
and
Naoto Kunitomo ‡

August 24, 2021

Abstract

We develop a new regression method called frequency regression and smoothing. This method is based on the separating information maximum likelihood developed by Kunitomo and Sato (2021) and Sato and Kunitomo (2020) for estimating the hidden states of random variables and handling noisy nonstationary (small sample) time series data. Many economic time series include not only the trend-cycle, seasonal, and measurement error components, but also factors such as structural breaks, abrupt changes, trading-day effects, and institutional changes. Frequency regression and smoothing can be applied to handle such factors in nonstationary time series. The proposed method is simple and applicable to several problems when analyzing nonstationary economic time series and handling seasonal adjustments. An illustrative empirical analysis of the macroconsumption in Japan is provided.

Key words

Noisy nonstationary time series, Trend-cycle, Measurement error and seasonality, SIML filtering, Regression smoothing, Structural and institutional change, Seasonal adjustment

---

*Version 2021-8-24. It is a revised version of the paper (SDS-20, MIMS=RSP Statistics and Data Science Series, Meiji University). The research has been supported by JSPS-Grant JP17H02513 and CARF. We would like to thank Editage (www.editage.com) for English language editing.

†Graduate School of Economics, University of Tokyo, Bunkyo-ku,Hongo7-3-1, Tokyo113-0033, seisho@e.u-tokyo.ac.jp

‡Tokyo Keizai University, Minami-cho 1-7-34, Kokubunnji-shi, Tokyo 185-8502, JAPAN, kunitomo@tku.ac.jp

1. Introduction

A considerable amont of research has been published on the use of statistical time series analysis of macroeconomic time series. An important feature of the macroeconomic time series, which is different from the standard time series analysis, is that the observed time series is an apparent mixture of nonstationary and stationary components including apparent seasonarity. The nonstationarity may include not only classical trend-cycle components, but also abrupt changes and outliers in the trend-noise components. A recent (vivid) example may be the macro-effects of Corona-Virus occurring in 2020-2021. Another feature is the fact that measurement errors in the economic time series play an important role because macroeconomic data are constructed from various sources including sample surveys from major official statistics, whereas in the statistical time series analysis often ignores the measurement errors. Further, official agencies apply the X-12-ARIMA program, which uses the univariate reg-ARIMA model to remove seasonality, as the standard filtering procedure to publish the seasonally adjusted data. The last important feature is that the sample size of the macroeconomic data is rather small; we obtain 120 time series observations for each series after collecting quarterly data over 30 years. The quarterly GDP series, which has been the most important data in macroeconomy have been published since 1994 by the cabinet office of Japan. Since the sample size is small, it is important to use an appropriate statistical procedure to extract information on the trend-cycle and noise (or measurement error) components in a systematic manner from data.

In this study, we develop a new regression method called frequency regression and smoothing, based on the separating information maximum likelihood (SIML) filtering (or smoothing). SIML filtering was developed by Kunitomo and Sato (2021) to estimate hidden states of random variables and handle multiple time series data. In particular, it can be applicable to small sample economic time series. Kunitomo and Sato (2017), Kunitomo and Sato (2021), and Sato and Kunitomo (2020) developed the SIML method for estimating nonstationary errors-in-variables models. They discussed the asymptotic and finite sample properties of the estimation of unknown parameter and developed the filtering method. We utilize their results to develop the linear regression methods in the frequency domain for nonstationary economic time series. Macroeconomic variables include important factors such as structural breaks, trading-day effects, and institutional changes in macro-economic variables in addition to trend, cycle, and seasonal components as well as the measurement errors. Since there are many factors in nonstationary time series, statistical method that can habdle them in a systematic and coherent manner yet to be developed. The proposed frequency regression method can be applied to handle these factors in a nonstationary time series. Our method is simple and applicable to several problems when analyzing a nonstationary economic time series.

As a classical study, Granger and Hatanaka (1964) reported the spectral analysis of economic time series. Engle (1974) introduced band spectrum regression for stationary time series. Our method of frequency regression could be regarded as extensions of their analyses to nonstationary time series.

The rest of the manuscript is organized as follows. In Section 2, we explain the nonstationary errors-in-variables model and SIML filtering (or smoothing) method. Then, in Section 3, we introduce the frequency regression method and as an application, we mention the method developed by M$\ddot{u}$ller and Watson (2018) briefly. In Section 4, we discuss the regression smoothing method based on SIML smoothing. In Section 5, we discuss the likelihood function. In Section 6, we discuss an illustrative empirical analysis of the macroconsumption of durable goods in Japan as an illustrative case. In Section 7, we provise some concluding remarks. Some details of the mathematical derivations of the results on frequency regression and the corresponding figures are presented in the Appendix.

2. Nonstationary Errors-in-variables models and SIML Filtering

2.1 Nonstationary Errors-in-variables models

Let y_{ji} be the $i-$th observation of the $j-$th time series at i for $i = 1, \cdots, n; j = 1, \cdots, p$. Let $\mathbf{y}_i = (y_{1i}, \cdots, y_{pi})'$ be a $p \times 1$ vector and $\mathbf{Y}_n = (\mathbf{y}_i^{'}) \ (= (y_{ij}))$ be an $n \times p$ matrix of observations, further let $\mathbf{y}_0$ be the initial $p \times 1$ vector. We estimate the model when the underlying nonstationary trend-cycle component $\mathbf{x}_i \ (= (x_{ji})) \ (i = 1, \cdots, n)$, the vector of the seasonal component $\mathbf{s}_i^{'} = (s_{1i}, \cdots, s_{pi})$, and the vector of the noise (or measurement error) component $\mathbf{v}_i^{'} = (v_{1i}, \cdots, v_{pi})$, which are independent of $\mathbf{x}_i$. We use the nonstationary errors-in-variables representation

$$\mathbf{y}_i = \mathbf{x}_i + \mathbf{s}_i + \mathbf{v}_i \ \ (i = 1, \cdots, n), \tag{2.1}$$

where $\mathbf{x}_i$, $\mathbf{s}_i$, and $\mathbf{v}_i$ $(i = 1, \cdots, n)$ are sequences of non-stationary I(1), stationary I(0) seasonal process, and stationary I(0) noise process. The trend-cycle component $\mathbf{x}_i$ satisfies

$$\Delta \mathbf{x}_i = (1 - \mathcal{L})\mathbf{x}_i = \mathbf{v}_i^{(x)} \tag{2.2}$$

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

$$\mathbf{v}_i^{(x)} = \sum_{j=0}^{\infty} \mathbf{C}_j^{(x)} \mathbf{e}_{i-j}^{(x)} \ , \tag{2.3}$$

where $\mathbf{e}_i^{(x)}$ denotes a sequence of i.i.d. random vectors with $\mathbf{E}(\mathbf{e}_i^{(x)}) = \mathbf{0}$ and $\mathbf{E}(\mathbf{e}_i^{(x)} \mathbf{e}_i^{(x)'}) = \mathbf{\Sigma}_e^{(x)}$ (a positive-semi-definite matrix). 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 $0 \le \rho < 1$ and $\|\mathbf{C}_j^{(x)}\| = \max_{k,l=1,\cdots,p} |c_{kl}^{(x)}(j)|$.
The random noise component $\mathbf{v}_i$ satisfies

$$\mathbf{v}_i = \sum_{j=0}^{\infty} \mathbf{C}_j^{(v)} \mathbf{e}_{i-j}^{(v)} , \tag{2.4}$$

where the $p \times p$ coefficient matrices $\mathbf{C}_j^{(v)}$ are absolutely summable and $\|\mathbf{C}_j^{(v)}\| = O(\rho^j)$, where $0 \le \rho < 1$ and $\mathbf{e}_i^{(v)}$ represents a sequence of i.i.d. random vectors with $\mathbf{E}(\mathbf{e}_i^{(v)}) = \mathbf{0}$, $\mathbf{E}(\mathbf{e}_i^{(v)} \mathbf{e}_i^{(v)'}) = \mathbf{\Sigma}_e^{(v)}$ (positive definite matrix).
The seasonal component $\mathbf{s}_i$ $(i = 1, \cdots, n)$ is a sequence of stationary processes satisfying

$$\mathbf{s}_i = \sum_{j=0}^{\infty} \mathbf{C}_{sj}^{(s)} \mathbf{e}_{i-sj}^{(s)} , \tag{2.5}$$

where the lag operator is defined by $\mathcal{L}^s \mathbf{s}_i = \mathbf{s}_{i-s}$ $(s \ge 2)$, and $\mathbf{e}_i^{(s)}$ represents a a sequence of i.i.d. random vectors with $\mathbf{E}(\mathbf{e}_i^{(s)}) = \mathbf{0}$ and $\mathbf{E}(\mathbf{e}_i^{(s)} \mathbf{e}_i^{(s)'}) = \mathbf{\Sigma}_e^{(s)}$ (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 $0 \le \rho < 1$.

Thus, we obtain the observations of an $n \times p$ matrix $\mathbf{Y}_n = (\mathbf{y}_i^{'})$ and set the $np \times 1$ random vector $(\mathbf{y}_1^{'}, \cdots, \mathbf{y}_n^{'})^{'}$. When there is no seasonal component and each pair of vectors $\Delta \mathbf{x}_i$ and $\mathbf{v}_i$ are independently, identically, and normally distributed (i.i.d.) as $N_p(\mathbf{0}, \mathbf{\Sigma}_x)$ and $N_p(\mathbf{0}, \mathbf{\Sigma}_v)$, respectively, $\mathbf{\Sigma}_x = \mathbf{\Sigma}_e^{(x)}$ and $\mathbf{\Sigma}_v = \mathbf{\Sigma}_e^{(v)}$. Then, given the initial condition $\mathbf{y}_0$, $\mathrm{vec}(\mathbf{Y}_n) \sim N_{n \times p}\left(\mathbf{1}_n \cdot \mathbf{y}_0^{'}, \mathbf{I}_n \otimes \mathbf{\Sigma}_v + \mathbf{C}_n \mathbf{C}_n^{'} \otimes \mathbf{\Sigma}_x\right)$, where $\mathbf{1}_n^{'} = (1, \cdots, 1)$ and

$$\mathbf{C}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ 1 & \cdots & 1 & 1 & 0 \\ 1 & \cdots & 1 & 1 & 1 \end{pmatrix}_{n \times n} . \tag{2.6}$$

In the general case with (2.1)-(2.5), we introduce the K_n^*-transformation from $\mathbf{Y}_n$ to $\mathbf{Z}_n$ $(= (\mathbf{z}_k^{'}))$ using

$$\mathbf{Z}_n = \mathbf{K}_n^* \left(\mathbf{Y}_n - \bar{\mathbf{Y}}_0\right) , \mathbf{K}_n^* = \mathbf{P}_n \mathbf{C}_n^{-1} , \tag{2.7}$$

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

$$\mathbf{C}_n^{-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}_{n \times n} , \tag{2.8}$$

and

$$\mathbf{P}_n = (p_{jk}^{(n)})\ ,\ p_{jk}^{(n)} = \sqrt{\frac{2}{n+\frac{1}{2}}}\cos\left[\frac{2\pi}{2n+1}(k-\frac{1}{2})(j-\frac{1}{2})\right]\ . \tag{2.9}$$

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

$$a_{kn}^{*}\ (= d_k)\ = 4\sin^2\left[\frac{\pi}{2}\left(\frac{2k-1}{2n+1}\right)\right]\ (k = 1, \cdots, n)\ . \tag{2.10}$$

2.2 SIML filtering method

We consider the general filtering procedure based on the $\mathbf{K}_n^{*}-$transformation (2.7). It is easy to interpret the role of the elements of the resulting $n \times p$ random matrix $\mathbf{Z}_n$ in the data analysis because they are obtained by the transformation that considers real values in the frequency domain. We consider the inversion of the transformation of orthogonal frequency processes. Let an $n \times p$ matrix

$$\hat{\mathbf{X}}_n = \mathbf{C}_n\mathbf{P}_n\mathbf{Q}_n\mathbf{P}_n\mathbf{C}_n^{-1}(\mathbf{Y}_n - \bar{\mathbf{Y}}_0)\ , \tag{2.11}$$

where $\mathbf{Z}_n = \mathbf{P}_n\mathbf{C}_n^{-1}(\mathbf{Y}_n - \bar{\mathbf{Y}}_0)$ and $\mathbf{Q}_n$denotes an $n \times n$ filtering matrix.
The stochastic process $\mathbf{Z}_n$ represents the orthogonal decomposition of the original time series $\mathbf{Y}_n$ (Section 5 of Kunitomo and Sato (2021)). We provide explicit forms of useful examples including the trend-cycle filtering procedure and seasonal filtering procedure for macro time series. Kunitomo and Sato (2021) developed the filtering (or smoothing) method in the form of (2.11), and we provide the following examples.

Example 1 : Trend Smoothing : Let an $m \times n$ choice matrix $(0 < m < n)$ $\mathbf{J}_m = (\mathbf{I}_m, \mathbf{O})$, and let $n \times p$ matrix

$$\hat{\mathbf{X}}_n = \mathbf{C}_n\mathbf{P}_n\mathbf{J}_m^{'}\mathbf{J}_m\mathbf{P}_n\mathbf{C}_n^{-1}(\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \tag{2.12}$$

and an $n \times n$ matrix

$$\mathbf{Q}_n = \mathbf{J}_m^{'}\mathbf{J}_m\ . \tag{2.13}$$

We will construct an estimator of the $n \times p$ hidden state matrix $\mathbf{X}_n$ in the frequency domain using the inverse transformation of $\mathbf{Z}_n$. We can recover the trend-cycle component by deleting the estimated noise parts in the high-frequency. (See Nishimura, Sato and Takahashi (2019) as a financial application.) Let the $[m + (n-m)] \times [m + (n-m)]$ partitioned matrix

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

and

$$(2.14) \qquad \mathbf{P}_n \mathbf{J}_m^{'} \mathbf{J}_m \mathbf{P}_n = \begin{pmatrix} \mathbf{P}_{11}^{'} \\ \mathbf{P}_{12}^{'} \end{pmatrix} (\mathbf{P}_{11}, \mathbf{P}_{12}) = \mathbf{I}_n - \begin{pmatrix} \mathbf{P}_{21}^{'} \\ \mathbf{P}_{22}^{'} \end{pmatrix} (\mathbf{P}_{21}, \mathbf{P}_{22}) \ .$$

Then the $(j, j^{'})$-th element of $\mathbf{A}_n = \mathbf{P}_n \mathbf{J}_m^{'} \mathbf{J}_m \mathbf{P}_n$ $(= (a_{j,j^{'}}))$ is given by

$$(2.15) \quad a_{j,j} = \frac{2m}{2n+1} + \frac{1}{2n+1} \left[\frac{\sin \frac{2m\pi}{2n+1}(2j-1)}{\sin \frac{\pi}{2n+1}(2j-1)} \right] ,$$

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

Example 2 : Band Smoothing : We consider the band filtering based on the $\mathbf{K}_n^{*}-$ transformation in (2.7) and use the inversion of only low-frequency parts from the random matrix $\mathbf{Z}_n$. A leading example is the seasonal frequency in the discrete time series, we consider s (> 1) to be a positive integer in this case. Let an $m_2 \times [m_1 + m_2 + (n - m_1 - m_2)]$ choice matrix $\mathbf{J}_{m_1,m_2,n} = (\mathbf{O}, \mathbf{I}_{m_2}, \mathbf{O})$, and let the $n \times p$ matrix

$$(2.16) \qquad \hat{\mathbf{X}}_n = \mathbf{C}_n \mathbf{P}_n \mathbf{J}_{m_1,m_2,n}^{'} \mathbf{J}_{m_1,m_2,n} \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0)$$

and the $n \times n$ matrix

$$(2.17) \qquad \mathbf{Q}_n = \mathbf{J}_{m_1,m_2,n}^{'} \mathbf{J}_{m_1,m_2,n} \ .$$

As an example, when we have the seasonal frequency λ_s $(0 \leq \lambda_s \leq \frac{1}{2})$, we take $m_1 = [2n/s] - h$ and $m_2 = 2h+1$. The $(j, j^{'})$-th element of $\mathbf{A}_n = \mathbf{P}_n \mathbf{J}_{m_1,m_2,n}^{'} \mathbf{J}_{m_1,m_2,n} \mathbf{P}_n$ $(= (a_{j,j^{'}}))$ is given by

$$(2.18) a_{j,j} = \frac{2m_2}{2n+1} + \frac{1}{2n+1} \left[\frac{\sin \frac{2(m_1+m_2)\pi}{2n+1}(2j-1) - \sin \frac{2(m_1)\pi}{2n+1}(2j-1)}{\sin \frac{\pi}{2n+1}(2j-1)} \right] ,$$

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

When $m_1 = 0$ and $m_2 = m$, (2.17) becomes (2.13) in Example 1. However, when we have seasonality, there is a complication in the data analysis and we need to use several transformations. (We shall discuss examples in Section 4 in details.) For quarterly data, a 1 year (4 quarters) cycle cannot be distinguished from the 2 quarters cycle. For monthly data, the 1 year cycle cannot be distinguished from the 6, 4, 3, 2.4, and 2 months cycles.

3. Frequency Regression

In this section, we consider a linear regression model based on observations of $q \times p$ matrix $\mathbf{Z}_m^*$ by

$$\mathbf{Z}_m^* = \mathbf{F}_q \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) = [\mathbf{z}_{1m}^*, \mathbf{Z}_{2m}^*] \ , \tag{3.1}$$

where $\mathbf{F}_q$ denotes a $q \times n$ matrix and $q\ (> p)$ depends on n as $q = q_n$.
There are several interesting examples. Since we consider the case when the rank of $\mathbf{F}_q$ is p $(p < q)$, let us investigate this case.
When we have nonstationary time series, we often have trend, cycle, seasonal, and noise components. To handle these components, we can use a more complicated transformation $\mathbf{F}_q$. Further, there are trading-day components, leap year effects, structural changes such as the 2008 financial crisis and 2020 corona-virus crisis, and institutional changes such as the consumption tax in Japan. When we generate seasonal adjusted data, it is important to handle these effects in meaningful ways. Since there are many components, it is known that an ad hoc method may be followed to handle these effects in official statistics from a standard statistical view.

The case investigated by Kunitomo and Sato (2021) considered the transformation when $\mathbf{F}_q = \mathbf{J}_m$. We first investigate this case and assume that the rank of $\mathbf{F}_q$ is p $(p < q)$. We define $p \times p$ matrices

$$\mathbf{G}_m^* = \frac{1}{m} \mathbf{Z}_m^{*'} \mathbf{Z}_m^* \ , \mathbf{G}_n = \frac{1}{n} \mathbf{Z}_n^{'} \mathbf{Z}_n \tag{3.2}$$

and we denote their probability limits as $m = m_n \to \infty$ $(n \to \infty, m_n/n \to 0)$ when they exist as

$$\mathrm{plim}_{n\to\infty} \mathbf{G}_m^* = \mathbf{\Sigma}_x \ , \mathrm{plim}_{n\to\infty} \mathbf{G}_n = \mathbf{\Sigma}_{\Delta y} \ , \tag{3.3}$$

where

$$\mathbf{\Sigma}_x = (\sum_{j=0}^{\infty} \mathbf{C}_j^{(x)}) \mathbf{\Sigma}_e^{(x)} (\sum_{j=0}^{\infty} \mathbf{C}_j^{(x)'}) \ (= \mathbf{f}_{\Delta x}(0)) \ , \tag{3.4}$$

($\mathbf{\Sigma}_x$ represents the spectral density matrix of $\Delta \mathbf{x}_i$ at zero frequency) and $\mathbf{\Sigma}_{\Delta y}$ is different from $\mathbf{\Sigma}_x$.

We partition $\mathbf{G}_m^*$ and $\mathbf{\Sigma}_x$ into $(1+k) \times (1+k)$ $(k = p-1)$ submatrices as

$$\mathbf{G}_m^* = \begin{bmatrix} g_{11}^* & \mathbf{g}_{12}^* \\ \mathbf{g}_{21}^* & \mathbf{G}_{22}^* \end{bmatrix} , \mathbf{\Sigma}_x = \begin{bmatrix} \sigma_{11} & \boldsymbol{\sigma}_{12} \\ \boldsymbol{\sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix} . \tag{3.5}$$

Then, we investigate the least squares estimator

$$\hat{\boldsymbol{\beta}}_m = \mathbf{G}_{22}^{*-1} \mathbf{g}_{21}^* \ , \tag{3.6}$$

which is an estimator of vector $\boldsymbol{\beta}_m = \boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\sigma}_{21}$ under the assumption that the inverse matrices of $\mathbf{G}_{22}^*$ and $\boldsymbol{\Sigma}_{22}$ exist. (We need to assume that $\boldsymbol{\Sigma}_{22}$ has a full rank.)
We write

$$\hat{\boldsymbol{\beta}}_m - \boldsymbol{\beta} = [\mathbf{Z}_{2m}^{*'}\mathbf{Z}_{2m}^*]^{-1}\mathbf{Z}_{2m}^{*'}\mathbf{Z}_m^* \begin{pmatrix} 1 \\ -\boldsymbol{\beta} \end{pmatrix} , \tag{3.7}$$

where we partitioned $\mathbf{Z}_m^*$ into $q \times (1+k)$ submatrices $\mathbf{Z}_m^* = (\mathbf{z}_{1m}^*, \mathbf{Z}_{2m}^*)$.
Then we have the next result on the asymptotic properties of the least squares estimator and the proof is presented in the Appendix.

Theorem 3.1 : Let $m = m_n = [n^\alpha]$ and $m \to \infty$ $(n \to \infty)$. In (2.1)-(2.5), assume that the fourth-order moments of $\mathbf{e}_i^{(x)}$, $\mathbf{e}_i^{(s)}$, and $\mathbf{e}_i^{(v)}$ are bounded.
(i) For $0 < \alpha < 1.0$, $\boldsymbol{G}_m^*$ is a consistent estimator of $\boldsymbol{\Sigma}_x$.
(ii) Assume that the rank of $\boldsymbol{\Sigma}_{22}$ is k $(= p-1)$. Let $m = m_n = [n^\alpha]$ and $0 < \alpha < 0.8$. Then when $m \to \infty$ $(n \to \infty)$, $\sqrt{m_n}[\hat{\boldsymbol{\beta}}_m - \boldsymbol{\beta}]$ is asymptotically and normally distributed as $N(\mathbf{0}, \sigma_{11.2}\boldsymbol{\Sigma}_{22}^{-1})$ and $\sigma_{11.2} = \sigma_{11} - \boldsymbol{\sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\sigma}_{21}$.

Then, we can rewrite $\mathbf{u}_m = \mathbf{z}_{1m}^* - \mathbf{Z}_{2m}^*\boldsymbol{\beta}$, that is,

$$\mathbf{z}_{1m}^* = \mathbf{Z}_{2m}^*\boldsymbol{\beta} + \mathbf{u}_m \tag{3.8}$$

and $\mathbf{E}[\mathbf{u}_m] = \mathbf{0}$. This is a linear regression equation, however, the error term of $\mathbf{u}_m$ has a specific form of heteroscedasticity.
Theorem 3.1 is used for the case when $\mathbf{G}_n$ is used for $\boldsymbol{\Sigma}_{\Delta y}$, which is different from $\boldsymbol{\Sigma}_x$.

One application of Theorem 3.1 would be Müller and Watson (2018), who proposed the *so-called* long-run co-variability of macroeconomic time series. They investigated many nonstationary time series using their method and obtained some interesting findings. Kunitomo and Sato (2021) have suggested an interpretation of their method as the relationships among long-run trends in our framework when $p = 2$. Let 2×2 matrices $\boldsymbol{\Sigma}_e^{(x)} = (\sigma_{ij}^{(x)})$; then, we define the regression coefficient $\boldsymbol{\beta} = [\sigma_{22}^{(x)}]^{-1}\sigma_{21}^{(x)}$ under the assumption that $\sigma_{22}^{(x)} > 0$.
Further, let $\boldsymbol{G}_m^* = (\hat{g}_{ij}^{(x)})$, and an $n \times 2$ matrix

$$(\mathbf{a}_{1n}, \mathbf{a}_{2n}) = \mathbf{C}_n^{-1}(\mathbf{Y}_n - \mathbf{Y}_0) . \tag{3.9}$$

For estimating $\boldsymbol{\beta}$, we define the estimated regression coefficient as

$$\hat{\boldsymbol{\beta}} = [\hat{g}_{22}^{(x)}]^{-1}\hat{g}_{21}^{(x)} = [\mathbf{a}_{2n}^{'}\mathbf{P}_n\mathbf{J}_m\mathbf{J}_m^{'}\mathbf{P}_n\mathbf{a}_{2n}]^{-1}[\mathbf{a}_{2n}^{'}\mathbf{P}_n\mathbf{J}_m\mathbf{J}_m^{'}\mathbf{P}_n\mathbf{a}_{1n}] . \tag{3.10}$$

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 a $n \times 2$ matrix $\mathbf{Y}_n = (\mathbf{y}_{1n}, \mathbf{y}_{2n})$;

that is, essentially the same as the one proposed by Müller and Watson (2018) [1]. Then, from Theorem 3.1, we immediately obtain the following result; the proof is reported in Kunitomo and Sato (2021).

Corollary 3.1 : When $p = 2$, we assume that $\mathbf{\Sigma}_e^{(x)}$ is positive semi-definite, $\mathbf{\Sigma}_e^{(s)}$ is positive semi-definite, $\mathbf{\Sigma}_e^{(v)}$ is positive definite, and that that the fourth-order moments of $\mathbf{e}_i^{(x)}$, $\mathbf{e}_i^{(s)}$, and $\mathbf{e}_i^{(v)}$ $(i = 1, \cdots, n)$ are bounded.
(i) We consider a sequence of integers $m = m_n$. Then $\hat{\beta}$ cannot be consistent when $n \to \infty$.
(ii) Set $m_n = [n^\alpha]$ and $0 < \alpha < 1$, then, as $n \longrightarrow \infty$, $\hat{\boldsymbol{\beta}}_m - \boldsymbol{\beta} \xrightarrow{p} \mathbf{0}$.
(iii) Set $m_n = [n^\alpha]$ and $0 < \alpha < 0.8$, then, as $n \longrightarrow \infty$, $\sqrt{m_n}[\hat{\boldsymbol{\beta}}_m - \boldsymbol{\beta}]$ is asymptotically normal.

It may be rather straight-forward to incorporate the regression effects of dummy variables in trend relations such as structural breaks.

4. Regression Smoothing

When we have noisy-nonstationary multivariate time series, we often need to remove the seasonality and/or low frequency component. However, in some applications of official statistics, we need to construct the seasonally adjusted data after removing additional effects such as trading-day components including the leap year effect, structural changes such as the 2008 financial crisis and institutional changes such as the introduction of consumption tax in Japan. These effects are can be defined in deterministic ways.

Let the observed vector times series $\mathbf{y}_i$ be decomposed as

$$\mathbf{y}_i = \mathbf{x}_i + \mathbf{SCO}_i + \mathbf{S}_i + \mathbf{v}_i \quad (i = 1, \cdots, n), \tag{4.1}$$

and $\mathbf{SCO}_i = \mathbf{SC}_i + \mathbf{O}_i$, where $\mathbf{x}_i$ denotes the trend-cycle component, $\mathbf{SC}_i$ denotes the structural break component, $\mathbf{S}_i$ represents the seasonal component, $\mathbf{v}_i$ denotes the noise component, and $\mathbf{O}_i$ represents the outlier component.
In this section we consider the case where in $\mathbf{SC}_i$ and $\mathbf{O}_i$ can be expressed as $\mathbf{SC}_i + \mathbf{O}_i = \mathbf{SCO}_i(w)$, where $\mathbf{w}$ denotes the set of instrumental variables. If these terms can be expressed as linear relationships, we write

$$\mathbf{y}_i = \mathbf{B}^{'}\mathbf{w}_i + \mathbf{u}_i \quad (i = 1, \cdots, n), \tag{4.2}$$

where $\mathbf{B}^{'}$ denotes a $p \times r$ matrix, $\mathbf{w}_i$ denotes a $r \times 1$ vector of instrumental variables, $\mathbf{z}_i$ and $\mathbf{u}_i = \mathbf{x}_i + \mathbf{S}_i + \mathbf{v}_i$ represents a sequence of I(1) process. Hence, the model is a multivariate regression model when the noise terms are $I(1)$ process with stationary

[1] In their notation, m corresponds to q, which is fixed. They did use (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 (2.11) in our notation.

noise term and seasonal terms. We extend the SIML smoothing method developed by Kunitomo and Sato (2021) and Sato and Kunitomo (2020) and incorporate extraneous information such as dummy variables to extract or delete some components from the observed time series based on (4.1).
To find the regression and smoothing procedure of trend and seasonal components, we use the K_n^*-transformation of data and rewrite (4.2) as

$$\mathbf{Y}_n^* = \mathbf{W}_n^* \mathbf{B} + \mathbf{U}_n^* \ , \tag{4.3}$$

where $\mathbf{Y}_n^* = \mathbf{P}_n \mathbf{C}_n^{-1}(\mathbf{Y}_n - \mathbf{Y}_0)$ and $\mathbf{W}_n^* = \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{W}_n$ ($\mathbf{W}_n = (\mathbf{w}_t'))$ represent $n \times p$ and $n \times r$ matrices of the explained variables and explanatory variables, respectively, and $\mathbf{U}_n^* = \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{U}_n$ denotes an $n \times p$ disturbance matrix. (We fix the initial condition $\mathbf{y}_0$ $(= \mathbf{x}_0)$ and the state variables $\mathbf{x}_i^* = \mathbf{x}_i - \mathbf{x}_0$. See Kunitomo and Sato (2021) and Sato and Kunitomo (2021) for details.)
As a consequence of the K_n^*-transformation, we have the disturbance terms in (4.3), that are stationary processes.

Because (4.2) is a linear regression equation, it is possible to apply Theorem 3.1 by defining a $(p + r) \times 1$ vector

$$\mathbf{y}_i^* = \begin{bmatrix} \mathbf{y}_i \\ \mathbf{w}_i \end{bmatrix} .$$

Then we can estimate the regression coefficients and calculate the residuals from the regression equations. When vectors $\mathbf{w}_i$ $(i = 1, \cdots, n)$ are deterministic, we assume that

$$\lim_{m \to \infty} \frac{1}{m} \mathbf{W}_m^{*\prime} \mathbf{W}_m^* = \mathbf{\Sigma}_{w^*} \ , \tag{4.4}$$

where $\mathbf{\Sigma}_{w^*}$ denotes a positive definite matrix and $\mathbf{W}_m^* = \mathbf{J}_m \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{W}_n$ represents an $m \times r$ matrix.
When the $r \times 1$ instrumental variables $\mathbf{w}_i$ $(i = 1, \cdots, n)$ are exogenous or deterministic, we have the following result from Theorem 3.1.

Corollary 4.1 : We assume the nonsingularity condition (4.4), $\mathbf{w}_i$ $(i = 1, \cdots, n)$ is exogenous or deterministic, and the fourth-order moments of $\mathbf{e}_i^{(x)}$, $\mathbf{e}_i^{(s)}$ and $\mathbf{e}_i^{(v)}$ are bounded. In (4.3), we represent the transformed $\mathbf{Y}_m^* = \mathbf{J}_m \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{Y}_n$ on $\mathbf{W}_m^* = \mathbf{J}_m \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{W}_n$ and $\hat{\mathbf{B}}_m$ denotes the least squares estimator of $\mathbf{B}$. Let $m_n = [n^\alpha]$ $(0 < \alpha < 0.8)$, and then, as $n \longrightarrow \infty$, we have the asymptotic normality

$$\sqrt{m_n}[\hat{\mathbf{B}}_m - \mathbf{B}] \xrightarrow{w} N(\mathbf{0}, \mathbf{\Sigma}_{w^*}^{-1} \otimes \mathbf{\Sigma}_x) \ . \tag{4.5}$$

Define the general transformed instrumental variables

$$\hat{\mathbf{W}}_n = \mathbf{J}_W \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{W}_n \ , \tag{4.6}$$

where $\mathbf{J}_W$ represents a $q \times n$ choice matrix, and we denote the idempotent matrix ($q \times q$ matrix)

(4.7) $$\mathbf{Q}_W = \hat{\mathbf{W}}_n(\hat{\mathbf{W}}_n'\hat{\mathbf{W}}_n)^{-1}\hat{\mathbf{W}}_n' \ .$$

We utilize the regression information on smoothing by utilizing the projection matrix $\mathbf{Q}_W$ to construct

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

There are several possibilities to how we incorporate the extraneous information in the smoothing procedure. It is reasonable to consider the case when $\mathbf{Q}_W$ is an idempotent matrix such as $\mathbf{Q}_W^2 = \mathbf{Q}_W$. In our study, we use two alternative smoothing procedures : *Type-I* and *Type II*. Type-I smoothing may be appropriate for change-point smoothing in the trend component and Type-II smoothing may be appropriate for outlier detection in the noise component.

(i) **Type-I Smoothing** : Type-I is based on Example 1 presented in Section 2. The (trend-cycle) regression part of $\mathbf{Y}_n$ is (4.1) when we take $\mathbf{J}_W = (\mathbf{I}_m, \mathbf{O})$ ($\hat{\mathbf{W}}_n$ represents an $m \times r$ matrix and $\mathbf{J}_m' = (\mathbf{I}_m, \mathbf{O})'$ represents an $n \times m$ matrix) and an $n \times n$ matrix

(4.9) $$\mathbf{Q}_n^{(0)} = \mathbf{J}_m'\hat{\mathbf{W}}_n(\hat{\mathbf{W}}_n'\hat{\mathbf{W}}_n)^{-1}\hat{\mathbf{W}}_n'\mathbf{J}_m \ .$$

If we want to remove the regression effects and use only the trend-cycle part, we need to take $\mathbf{J}_W = \mathbf{J}_m$, $\mathbf{J}_m = (\mathbf{I}_m, \mathbf{O})$ ($m \times n$ choice matrix, $m \leq n$) and

(4.10) $$\mathbf{Q}_n^{(1)} = \mathbf{J}_m'\mathbf{J}_m - \mathbf{Q}_n^{(0)} = \mathbf{J}_m'[\mathbf{I}_m - \hat{\mathbf{W}}_n(\hat{\mathbf{W}}_n'\hat{\mathbf{W}}_n)^{-1}\hat{\mathbf{W}}_n']\mathbf{J}_m \ .$$

Then we have the decomposition

(4.11) $$\begin{aligned} \hat{\mathbf{X}}_n &= \mathbf{C}_n\mathbf{P}_n\mathbf{J}_m'\mathbf{J}_m\mathbf{P}_n\mathbf{C}_n^{-1}(\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \\ &= \mathbf{C}_n\mathbf{P}_n\mathbf{J}_m'[\mathbf{Q}_n^{(0)} + \mathbf{Q}_n^{(1)}]\mathbf{J}_m\mathbf{P}_n\mathbf{C}_n^{-1}(\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \ . \end{aligned}$$

In this case, we have the property $\mathbf{Q}_n^2 = \mathbf{Q}_n = \mathbf{Q}_n^{(0)} + \mathbf{Q}_n^{(1)} = \mathbf{J}_m'\mathbf{J}_m$, and we have the decomposition of the trend-cycle part and the regression part. There is a simple interpretation of this smoothing because we use only the regression part at m low frequencies. We first remove the regression part from $\mathbf{Y}_n$ by taking

(4.12) $$\mathbf{X}_n^{(1)} = \mathbf{C}_n\mathbf{P}_n[\mathbf{I}_n - \mathbf{Q}_n^{(0)}]\mathbf{P}_n\mathbf{C}_n^{-1}(\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \ .$$

We apply the 2nd smoothing to $\mathbf{Y}_n^{(1)}$ as

(4.13) $$\mathbf{X}_n^{(2)} = \mathbf{C}_n\mathbf{P}_n\mathbf{J}_m'\mathbf{J}_m\mathbf{P}_n\mathbf{C}_n^{-1}\mathbf{X}_n^{(1)}$$

by taking another transformation.
Then, the resulting transformation is (4.8) with $\mathbf{Q}_W = \mathbf{Q}_n^{(1)}$ after an iteration. Since Sato and Kunitomo (2020) developed an iterated smoothing procedure, there should be some mechanism for performing further iterations.

(ii) **Type-II Smoothing** : Type-II smoothing is based on Example 2 presented in Section 2. When we need to estimate not only the trend component, but also the noise component, it is important to estimate structural changes and outlier components consistently. For instance, we need to estimate the seasonal component for obtaining the seasonally adjusted series, and it is related to Example 2. Thus, we construct an $q \times n$ choice matrix $\mathbf{F}_q$ such that the seasonal components can be removed in their frequencies.

When $s = 4$, we want to remove the data with frequencies around $\lambda_s = 1/4, 1/2$ (1/2 corresponds to the cycle of 2 quarters and 1/4 corresponds to the cycle of 4 quarters). However, we cannot distinguish the 4 quarters cycle from the 2 quarters cycle by using quarterly observations. We set $m_1 = [2n/s]$, and an $(n-2h-1) \times n$ choice matrix and an $(n-3h-2) \times (n-2h-1)$ choice matrix as

$$\mathbf{J}_1^Q = \begin{bmatrix} \mathbf{I}_{m_1-(h+1)} & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{I}_{n-m_1-h} \end{bmatrix} , \ \mathbf{J}_2^Q = [\mathbf{I}_{n-3h-2} \ , \ \mathbf{O}] \ . \tag{4.14}$$

Then we take a $q \times n$ matrix

$$\mathbf{F}_q^Q = \mathbf{J}_2^Q \mathbf{J_1}^Q \tag{4.15}$$

with a small positive integer $h > 0$.

When $s = 12$, we need a more complicated transformation to remove seasonality because we cannot distinguish the 12 month cycle from the 6, 4, 3, 2.4, and the 2 month cycles using monthly observations with frequencies around $\lambda_s = 1/12, 2/12, 3/12, 4/12, 5/12, 6/12$. We set $m_i = i[2n/s]$ and take $(n - i(2h+1)) \times (n - (i-1)(2h+1))$ choice matrices $(i = 1, \cdots, 5)$ and an $(n - 5(2h+1) - (h+1)) \times (n - 5(2h+1))$ choice matrix such that

$$\mathbf{J}_i^M = \begin{bmatrix} \mathbf{I}_{m_i-(i-1)(2h+1)-(h+1)} & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{I}_{n-m_i-h} \end{bmatrix} , \ \mathbf{J}_6^M = [\mathbf{I}_{n-11h-6} \ , \ \mathbf{O}] \ , \tag{4.16}$$

with a small positive integer $h > 0$. To remove the data with seasonal frequencies around λ_{js} $(j = 2, 3, 4, 5)$ using $\mathbf{J}_j^M$ $(j = 1, \cdots, 6)$, we set a $q \times n$ matrix

$$\mathbf{F}_q^M = \prod_{j=1}^{6} \mathbf{J}_{7-j}^M \ . \tag{4.17}$$

Although we do not know the unknown coefficient matrix $\mathbf{B}$, we can incorporate the estimated coefficient by regressing

$$\mathbf{Y}_m^* = \mathbf{F}_q \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \tag{4.18}$$

to

$$\mathbf{W}_m^* = \mathbf{F}_q \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{W}_n - \bar{\mathbf{W}}_0) \ , \tag{4.19}$$

where $\mathbf{F}_q$ is either $\mathbf{F}_q^Q$ or $\mathbf{F}_q^M$.
Type-II smoothing is defined by

$$\mathbf{Q}_n^{(2)} = \mathbf{W}_n^*(\mathbf{W}_n^{*'}\mathbf{W}_n^*)^{-1}\mathbf{W}_n^{*'} \ . \tag{4.20}$$

and

$$\mathbf{Q}_n^{(3)} = \mathbf{F}_q^{'}\mathbf{F}_q - \mathbf{Q}_n^{(2)} \ . \tag{4.21}$$

Then, we have the decomposition

$$\begin{aligned} \hat{\mathbf{X}}_n^* &= \mathbf{C}_n\mathbf{P}_n\mathbf{F}_q^{'}\mathbf{F}_q\mathbf{P}_n\mathbf{C}_n^{-1}(\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \\ &= \mathbf{C}_n\mathbf{P}_n[\mathbf{Q}_n^{(2)} + \mathbf{Q}_n^{(3)}]\mathbf{P}_n\mathbf{C}_n^{-1}(\mathbf{Y}_n - \bar{\mathbf{Y}}_0) \ . \end{aligned} \tag{4.22}$$

In this case, we have the decomposition $\mathbf{Q}_n^{(2)} + \mathbf{Q}_n^{(3)} = \mathbf{F}_q^{'}\mathbf{F}_q$ and the corresponding decomposition of the trend-cycle and regression parts.

Examples of Dummy Variables :
There are some examples of outlier and trend dummies. For nonstationary time series, we should be careful about normalization because there can be significant effects on smoothing. Although there are many other possible dummy variables, we provide some examples that have been used in official data handling such as official seasonal adjustment. Let w_s $(s = 1, \cdots, n)$ be the dummy variable.

Example 1 :
The level shift (LS) variable can be defined as $w_s = 0$ if $s < t$ and $w_t = 1$ if $s \geq t$ for $s = 1, \cdots, n$. This can be handled by Type-I smoothing.

Example 2 :
The outlier variable can be defined as $w_s = 1$ if $s = t$ and $w_t = 0$ if $s \neq t$ for $s = 1, \cdots, n$. This variable is often called additive outlier (AO).

Example 3 :
The ramp variable can be defined by $w_s = 1$ if $s < t_0$, $w_s = 1 - (t - t_0)/(t_1 - t_0)$ if $t_0 \leq t \leq t_1$, and $w_t = 0$ if $s \geq t_1$.

Example 4 :
The double ramp variable can be defined by $w_s = 1$ if $s < t_0$, $w_s = 1-(t-t_0)/(t_1-t_0)$ if $t_0 \leq t \leq t_1$, $w_s = (t - t_1)/(t_2 - t_1)$ if $t_1 \leq t \leq t_2$, and $w_t = c$ if $s \geq t_2$.

5. Frequency Domain and Likelihood

We consider the additive decomposition model $\mathbf{y}_i = \mathbf{x}_i + \mathbf{s}_i + \mathbf{v}_i$ $(i = 1, \cdots, n)$ and take positive integers s $(s > 1)$, N, and $n = sN$ for the resulting simplicity and arguments.

Let $\mathbf{f}_{\Delta x}(\lambda)$, $\mathbf{f}_s(\lambda)$, and $\mathbf{f}_v(\lambda)$ be the spectral density $(p \times p)$ matrices of $\Delta \mathbf{x}_i$, $\mathbf{s}_i$, and $\mathbf{v}_i$ $(i = 1, \cdots, n)$, respectively, such that

$$(5.1) \qquad \mathbf{f}_{\Delta x}(\lambda) = (\sum_{j=0}^{\infty} \mathbf{C}_j^{(x)} e^{2\pi i \lambda j}) \mathbf{\Sigma}_e^{(x)} (\sum_{j=0}^{\infty} \mathbf{C}_j^{(x)'} e^{-2\pi i \lambda j}) \quad (-\frac{1}{2} \leq \lambda \leq \frac{1}{2}) ,$$

$$(5.2) \qquad \mathbf{f}_v(\lambda) = (\sum_{j=0}^{\infty} \mathbf{C}_j^{(v)} e^{2\pi i \lambda j}) \mathbf{\Sigma}_e^{(v)} (\sum_{j=0}^{\infty} \mathbf{C}_j^{(v)'} e^{-2\pi i \lambda j}) \quad (-\frac{1}{2} \leq \lambda \leq \frac{1}{2}) ,$$

and

$$(5.3) \qquad \mathbf{f}_s(\lambda) = (\sum_{j=0}^{\infty} \mathbf{C}_{sj}^{(s)} e^{2\pi i \lambda s j}) \mathbf{\Sigma}_e^{(s)} (\sum_{j=0}^{\infty} \mathbf{C}_{sj}^{(s)'} e^{-2\pi i \lambda s j}) \quad (-\frac{1}{2} \leq \lambda \leq \frac{1}{2}) ,$$

where we set $\mathbf{C}_0^{(x)} = \mathbf{C}_0^{(v)} = \mathbf{C}_0^{(s)} = \mathbf{I}_p$ as normalizations and $i^2 = -1$. Then, the $p \times p$ spectral density matrix of the transformed vector process, which are observable, and the spectral density of the difference series $\Delta \mathbf{y}_i$ $(= \mathbf{y}_i - \mathbf{y}_{i-1})$ can be represented as

$$(5.4) \qquad \mathbf{f}_{\Delta y}(\lambda) = \mathbf{f}_{\Delta x}(\lambda) + (1 - e^{2\pi i \lambda})[\mathbf{f}_s(\lambda) + f_v(\lambda)](1 - e^{-2\pi i \lambda}) \ .$$

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

$$(5.5) \qquad \mathbf{\Sigma}_e^{(x)} = \mathbf{f}_{\Delta x}(0) \ (= (\sigma_{gh}^{(x)})) \ , \ \mathbf{\Sigma}_e^{(v)} = f_v(0) \ = (\sigma_{gh}^{(v)}) \ .$$

Let $f_v^{(SR)}(\lambda_k)$, $f_s^{(SR)}(\lambda_k)$ and $f_{\Delta x}^{(SR)}(\lambda_k)$ be the symmetrized $p \times p$ spectral matrices of $\mathbf{v}_i$, $\mathbf{s}_i$ and $\Delta \mathbf{x}_i$ at λ_k $(= (k - \frac{1}{2})/(2n+1))$ for $k = 1, \cdots, n$, that is, $f_v^{(SR)}(\lambda_k) = (1/2)[f_v^{(SR)}(\lambda_k) + \bar{f}_v^{(SR)}(\lambda_k)]$, $f_s^{(SR)}(\lambda_k) = (1/2)[f_s^{(SR)}(\lambda_k) + \bar{f}_s^{(SR)}(\lambda_k)]$, and $f_{\Delta x}^{(SR)}(\lambda_k) = (1/2)[f_{\Delta x}^{(SR)}(\lambda_k) + \bar{f}_{\Delta x}^{(SR)}(\lambda_k)]$.
Proposition 1 of Kunitomo and Sato (2021) yields the condition that orthogonal processes are approximately distributed as Gaussian distributions. Then, (-2) times the log-likelihood function in the general model can be approximated as

$$(5.6) \quad \begin{aligned} (-2) l_n(\boldsymbol{\theta}) \ = \ & \sum_{k=1}^{n} \log |a_{kn}^{*}(f_v^{(SR)}(\lambda_k) + f_s^{(SR)}(\lambda_k)) + f_{\Delta x}^{(SR)}(\lambda_k))| \\ & + \sum_{k=1}^{n} \mathbf{z}_k^{'} [a_{kn}^{*}(f_v^{(SR)}(\lambda_k) + f_s^{(SR)}(\lambda_k)) + f_{\Delta x}^{(SR)}(\lambda_k)]^{-1} \mathbf{z}_k \ . \end{aligned}$$

We further consider the case when $\Delta \mathbf{x}_i$, $\mathbf{s}_i$, and $\mathbf{v}_i$ are a sequence of independent random vectors. Then we have $\mathbf{\Sigma}_e^{(x)} = f_{\Delta x}^{(SR)}(\lambda_k)$ and $\mathbf{\Sigma}_e^{(v)} = f_v^{(SR)}(\lambda_k)$ for $k = 1, \cdots, n$, and $\mathbf{\Sigma}_e^{(s)} = f_s^{(SR)}(\lambda_k)$ for some index set $k \in I_n(s)$.

Moreover, when we have some dummy variables $\mathbf{W}_n$, we have assumed that they are independent of other noise, cycle, seasonal, and trend components. Then, under Gaussian assumption given the initial condition and the information set of explanatory variables $\mathbf{W}_n$, we can write (-2) times the conditional log-likelihood as

$$(5.7)\quad (-2)l_n(\boldsymbol{\theta}|\mathbf{W}_n) = \sum_{k=1}^{n} \log|\mathbf{\Sigma}_{u^*}(k,w)| + \sum_{k=1}^{n}[\mathbf{y}_k^{*'} - \mathbf{w}^{*`}\mathbf{B}][\mathbf{\Sigma}_{u^*(k,w)}]^{-1}[\mathbf{y}_k^* - \mathbf{B}^{`}\mathbf{w}_k^*]\ ,$$

where $\mathbf{y}_k^*$ and $\mathbf{w}_k^*$ are the transformed explained variables and explanartory variables (using K_n^*transformation from the observed $\mathbf{y}_i$ and $\mathbf{w}_i$ $(i = 1, \cdots, n)$, respectiely, and $\mathbf{\Sigma}_{u^*}(k,w) = a_{kn}^*(\mathbf{\Sigma}_e^{(v)}(w) + \mathbf{\Sigma}_e^{(s)}(w)) + \mathbf{\Sigma}_e^{(x)}(w)$, which is the variance-covariance matrix of $\mathbf{u}_k^*$ $(k = 1, \cdots, n)$. If we regard $a_{k,n}$ as constants with respect to k, the log-likelihood function becomes a standard form. This analysis is useful because the likelihood function is a complicated function and we oten need some approximation.

When we use the explanatory variables $\mathbf{W}_n$, we can estimate the unknown matrix $\mathbf{B}$ by Corollary 4.1 consistently. Let $\hat{\mathbf{B}}$ be the SIML estomator and $\mathbf{z}_k^* = \mathbf{y}_k^* - \hat{\mathbf{B}}^{`}\mathbf{w}_k^*$ $(k = 1, \cdots, n)$, which depend son $\mathbf{W}_n$ and denote $\mathbf{z}_k^*(w)$ $(k = 1, \cdots, n)$.
To estimate $\mathbf{\Sigma}_x$, it is reasonable to use

$$(5.8)\qquad \mathbf{G}_m^*(w) = \frac{1}{m_n}\sum_{k=1}^{m_n} \mathbf{z}_k^*(w)\mathbf{z}_k^{*'}(w)$$

with $m/n \to 0$ and $m \to \infty$.

There are two remarks. First, the ML estimation in the nonstationary errors-in-variables models may have some difficulty when $p > 1$ without some strong restrictions of the parameter space as illustrated by Kunitomo, Awaya and Kurisu (2019). Second, the likelihood functions in this section can be interpreted as the *Whittle-type* likelihood function that does not depend on the Gaussian distributions as utilized by Hosoya (1997).

6. Example of macroconsumption of durable goods

We use the official macroconsumption data of durable goods in Japan from 1994Q1 to 2019Q4 to illustrate the regression smoothing [2]. Further, we need to estimate the trend, seasonal, and noise components to deal with the original quarterly time series. We applied the SIML smoothing procedure with $m = 29$ and $h = 2$, which yields the minimum numbers of AIC. Kunitomo and Sato (2021) and Sato and Kunitomo (2020) explained some aspects of the choice problem of m. All corresponding figures are presented in Appendix B.

Figure 1 shows a summary of SIML smoothing for log-transformed data. This is done because the original series has a significant heteroscedastic seasonality. In

[2] We have taken data from https://www.esri.cao.go.jp/jp/sna/menu.html (Economic and Social Research Institute (ESRI), Cabinet Office, Japan). They are original series in real terms and ESRI uses the X-12-ARIMA smoothing program for constructing seasonal-adjusted official data.

Figures 1-4, “org ” stands for the original series, “trend ”, “seasonal ”, “noise ” mean the estimated trend, seasonal, and noise components,]while “adj ” means the estimated seasonally adjusted series, i.e., the observed series minus the estimated seasonal component. “Z ” means the transformed series. In Figures 2-4, “reg ” stands for the dummy variable.

The original time series has typical characteristics of macroeconomic time series in Japan, i.e., it is a realization of nonstationary time series and exhibits rather clear trend, cycle, seasonal and irregular components. We applied SIML filtering with $m = 29$; red curve indicates the estimated trend-cycle component. Using Z-transfprmed data, we capture the segnificant effects at the seaeasonal frequencis. Although the estimated seasonal component gives regular seasonal pattern, the estimated trend-cycle and noise components suggest there are some abrupt changes around the year of 2008-2009, 2011, and 2014, which may be different from the usual noise component.

Then, we applied two AO-dummy variables at 2011Q1 and 2014Q1. In these periods, there was a large effect caused by the 2011 earthquake in Japan, in addition to introduction of consumption tax, both these events had significant effects on the macroeconomy and consumption in Japan. We applied the ramp-dummy variable from 2008Q3 to 2009Q1. In this period, there was a rapid downward effect attributed by the 2008 financial crisis, and we can consider it to be appropriate to use the ramp-dummy at 2008Q3-2009Q1. Figures 2 and 3 summarize SIML smoothing and frequency regression results for the cases.

Finally, Figure 4 represents a summary of SIML smoothing and frequency regression with three dummy variables considered simultaneously. Based on the criteria of AIC, we selected the last case for the best modelling for the macroconsumption of durable goods; these effects are captured by our method. By using the transformed data of (4.14) and (4.15) and the dummy variables, the AIC(w) was calculated based on the regression equation by

$$AIC(w) = n \log \hat{\sigma}_w^2 + 2r \tag{6.1}$$

where we use $\hat{\sigma}_w^2$ calculated from the residuals of the dummy regression ((4.2) with $p = 1$) and r denotes the number of dummy variables [3].
In our example we have $p = 1$, and two AICs were calculated: the first AIC in figures was calculated using m low frequencies while the AIC in the parenthesis was calculated using all frequency data except data around the seasonal frequency.

By using the model selection criteria for minimizing these AICs, we find that SIML smoothing with three dummy variables (i.e., two AOs and a double ramp) is the best model. We have reasonable result on the decomposition of original

[3]This $AIC(w)$ is based on (4.3) and (5.7), which can be implemented easily. However, we have taken the case as if a_{kn} were constant with respect to k because we use the procedure, that is free from the maximum likelihood (ML) estimation of unknown parameters needed. In this sense, our $AIC(w)$ is an appromimate one.

time series into trend-cycle, seasonal, and noise components. In the selected model, the trend-cycle component includes one structural change and the noise component includes two outliers.

This empirical analysis illustrates that we need to consider the important role of incorporating the effects of the change point problem and abrupt changes in the seasonal adjustment procedure.

7. Concluding Remarks

In many official time series, it is common to observe nonstationary trend, cycles, seasonal, and measurement errors simultanepusly. In addition to these components, we sometimes observe abrupt changes, trading-day effects, and other irregular components. Thus, it seems difficult to remove the seasonal component and construct macro-index, which involve multiple nonstationary time series. This paper presents a new approach to handle nonstationary time series using frequency regression based on SIML modelling in a systematic manner. Our method sheds new light on some practical approaches to handle published economic time series, which have been practically used in official seasonal adjustments without formal justifications. There are many empirical examples, and we reported an application of constructing a monthly macroconsumption index in Kunitomo, Sato and Sakurai (2021) in some detail as a real illustration. Another application would be the macro effects of Corona-Virus occurring in 2020-2021, which is currently under investigation.

There are some problems that stilll need to be investigated. The present study is based on the standard time series decomposition in (2.1) and (5.1), and assume that $\mathbf{v}_i^{(x)}$, $\mathbf{v}_i^{(s)}$ and $\mathbf{v}_i^{(x)}$ are i.i.d. sequences of random variables with mean zero and variance-covariance matrix. This implies that their spectral densities are constant in the frequency domain. There may be another approach to formulate the problem and decompose the time series. For instance, it may be reasonable to consider the case when the spectral density of $\mathbf{v}_i^{(s)}$ is zero except the region around the zero frequency.

Another issue would be the computation of the procedure we explained in this paper. We have developed R-programs (Sato (2020), Kunitomo, Sato and Sakurai (2021)), which will be available in the future.

References

[1] Anderson, T. W. (1971), *The Statistical Analysis of Time Series*, Wiley.

[2] Brockwell, P. and R. Davis (1990), *Time Series : Theory and Methods*, 2nd Edition, Wiley.

[3] Engle, R. (1974), "Band spectrum regression," *Econometrica*, Vol. 15-1, 1-11.

[4] Granger, C. and M. Hatanaka (1964), *Spectral Analysis of Economic Time Series*, Princeton University Press.

[5] Hosoya, Y. (1997), "A limit theory for long range dependence and statistical inference on related models," *The Annals of Statistics*, vol.25, no.1, pp.105-137.

[6] Kitagawa, G. (2010), *Introduction to Time Series Analysis*, CRC Press.

[7] Kunitomo, N., N. Awaya and D. Kurisu (2020), "Some properties of estimation methods for sructural relationships in nonstationary errors-in-variables models," Open Access, *Japanese Journal of Statistics and Data Science*, Springer.

[8] Kunitomo, N. and S. Sato (2017), "Trend, seasonality and economic time series : the nonstationary errors-in-variables models," SDS-4, MIMS, Meiji University, http://www.mims.meiji.ac.jp/publications/2017-ds.

[9] Kunitomo, N. , S. Sato and D. Kurisu (2018), *Separating Information Maximum Likelihood Estimation for High Frequency Financial Data*, Springer.

[10] Kunitomo, N. and S. Sato (2021), "A robust filtering method for noisy nonstationary multivariate time series with an application to macroconsumption," Open Access, *Japanese Journal of Statistics and Data Science*, Springer.

[11] Kunitomo, N., S. Sato and T. Sakurai (2021), "Estimating economic time series state and constructing a macro index," (in japenese) SDS-19, MIMS, Meiji University.

[12] Sato, S. (2020), "x12siml6," In preparation (Unpublished Program).

[13] Sato, S. and N. Kunitomo (2020), "On backward smoothing for noisy nonstationary time series," SDS-15, MIMS, Meiji University.

[14] Müller, U. and M. Watson (2018), "Long-run covariability," *Econometrica*, Vol. 86-3, 775-804.

APPENDIX A : Mathematical Derivations

We present the details for the derivation of Theorem 3.1 as an application of Theorem A.1 below, that is an extension of Proposition 2 in Kunitomo and Sato (2021) and Chapter 5 of Kunitomo, Sato and Kurisu (2018). Since some details are essentially the same in the existing literature, we omit some of them and only refer to them. We denote them as KS (2021) and KSK (2018), and use notations and some arguments in their proof. We first provide the intuition for our result based on KS (2021) and then present our proof.

A-I A Heuristic Derivation : We use the arguments in Section 5 of KS (2021). 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 $\mathbf{z}_k$ $(k = 1, \cdots, n)$ as

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

which is a (real-valued) Fourier type transformation and $\mathbf{y}_0$ is fixed.
Then, we find that $\mathbf{z}_n(\lambda_k^{(n)})$ $(k = 1, \cdots, n)$ 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 $\mathbf{z}(\lambda)$ $(0 \le \lambda \le 1/2)$, which is continuous in the frequency domain. They are asymptotically uncorrelated random variables. (See Chapters 8-9 of Anderson (1971) or Section 5 of KS (2021).)
Then, using a similar argument as in Proposition 1 of KS (2021), we find that for $k \neq k^{'}$

$$\text{(A.2)} \qquad \mathbf{E}[\mathbf{z}_{in}(\lambda_k^{(n)}) \mathbf{z}_{jn}(\lambda_k^{(n)}) \mathbf{z}_{hn}(\lambda_{k'}^{(n)}) \mathbf{z}_{ln}(\lambda_{k'}^{(n)})] = \sigma_{ij}(\lambda_k^{(n)}) \sigma_{hl}(\lambda_{k'}^{(n)}) + o(1)$$

and for $k = k^{'}$

$$\begin{aligned} \text{(A.3)} \qquad & \mathbf{E}[\mathbf{z}_{in}(\lambda_k^{(n)}) \mathbf{z}_{jn}(\lambda_k^{(n)}) \mathbf{z}_{hn}(\lambda_k^{(n)}) \mathbf{z}_{ln}(\lambda_k^{(n)})] \\ = \ & \sigma_{ij}(\lambda_k^{(n)}) \sigma_{hl}(\lambda_k^{(n)}) + \sigma_{ih}(\lambda_k^{(n)}) \sigma_{jl}(\lambda_k^{(n)}) + \sigma_{il}(\lambda_k^{(n)}) \sigma_{jh}(\lambda_k^{(n)}) + o(1) \ , \end{aligned}$$

where

$$\text{(A.4)} \qquad \sigma_{ij}(\lambda_k^{(n)}) = \sum_{h=-(n-1)}^{n-1} \left[\cos 2\pi \lambda_k^{(n)} h\right] \mathbf{\Gamma}_{ij}(h) \ ,$$

$\mathbf{E}[\mathbf{r}_j \mathbf{r}_{j-h}^{'}] = \mathbf{\Gamma}(h)$, $\mathbf{r}_j = \mathbf{y}_j - \mathbf{y}_{j-1}$ $(j = 1, \cdots, n)$ and $\mathbf{y}_0$ is a fixed vector.
As $n \to \infty$ and $m/n \to 0$, we have $\lambda_k^{(n)} \to 0$ for $1 \le k \le m$. We write for $k = 1, \cdots, m$ and as $m/n \to 0$,

$$\text{(A.5)} \qquad \lim_{n\to\infty} \sigma_{ij}(\lambda_k^{(n)}) = \sigma_{ij}^{(x)} \ (i, j = 1, \cdots, p)$$

and $\mathbf{\Sigma}_x = (\sigma_{ij}^{(x)})$. Then

$$\text{(A.6)} \qquad \mathbf{Var}[\frac{1}{m}\sum_{k=1}^{m}\mathbf{z}_{in}(\lambda_k^{(n)})\mathbf{z}_{jn}(\lambda_k^{(n)})] \longrightarrow \sigma_{ii}^{(x)}\sigma_{jj}^{(x)} + \sigma_{ij}^{(x)2} \ .$$

We construct a sequence of random variables, which are approximately uncorrelated (see Proposition 1 of KS (2021)) and for $i,j = 1,\cdots,p$

$$s_{ij}(t) = \mathbf{z}_{in}(\lambda_t^{(n)})\mathbf{z}_{jn}(\lambda_t^{(n)}) - \mathbf{E}[\mathbf{z}_{in}(\lambda_t^{(n)})\mathbf{z}_{jn}(\lambda_t^{(n)})]$$

and

$$M_{ij}(n,k) = \sum_{t=1}^{k} s_{ij}(t) \ .$$

Then, heuristically, we can apply the central limit theorem (CLT) for the stationary process to obtain the asymptotic normality. However, to show this argument in a rigorous way, we need further developments.

A-II Proof of Main Results : We first prepare a general result of the consistency and asymptotic normality of the SIML estimation in nonstationary time series; this is new and an extension of Proposition 2 of KS (2021).

Theorem A.1: Assume that the fourth order moments of each element of $\mathbf{v}_i^{(x)}$ and $\mathbf{v}_i$ in (2.1)-(2.5) are bounded. Let

$$\text{(A.7)} \qquad \hat{\mathbf{\Sigma}}_x (= (\hat{\sigma}_{gh}^{(x)})) = \frac{1}{m}\mathbf{Z}_m^{*'}\mathbf{Z}_m^{*} \ ,$$

which is $\mathbf{G}_m^*$ in (3.2). Then
(i) For $m_n = [n^\alpha]$ and $0 < \alpha < 1$, as $n \longrightarrow \infty$

$$\text{(A.8)} \qquad \hat{\mathbf{\Sigma}}_x - \mathbf{\Sigma}_x \xrightarrow{p} \mathbf{O} \ .$$

(ii) We set $\mathbf{\Sigma}_x = (\sigma_{gh}^{(x)})$. For $m_n = [n^\alpha]$ and $0 < \alpha < 0.8$, as $n \longrightarrow \infty$

$$\text{(A.9)} \qquad \sqrt{m_n}\left[\hat{\sigma}_{gh}^{(x)} - \sigma_{gh}^{(x)}\right] \xrightarrow{\mathcal{L}} N\left(0, \sigma_{gg}^{(x)}\sigma_{hh}^{(x)} + \left[\sigma_{gh}^{(x)}\right]^2\right) \ .$$

The covariance of the limiting distributions of $\sqrt{m_n}[\hat{\sigma}_{gh}^{(x)} - \sigma_{gh}^{(x)}]$ and $\sqrt{m_n}[\hat{\sigma}_{kl}^{(x)} - \sigma_{kl}^{(x)}]$ is given by $\sigma_{gk}^{(x)}\sigma_{hl}^{(x)} + \sigma_{gl}^{(x)}\sigma_{hk}^{(x)}$ $(g,h,k,l = 1,\cdots,p)$.

Proof of Theorem A.1 : The proof consists of two steps.
(**Step 1**) : Let $\mathbf{z}_k^{(x)} = (z_{kj}^{(x)})$ and $Z_k^{(s+v)} = (z_{kj}^{(s+v)})$ $(k = 1,\cdots,n)$ be the k-th row vector elements of $n \times p$ matrices

$$\text{(A.10)} \qquad \mathbf{Z}_n^{(x)} = \mathbf{K}_n^*(\mathbf{X}_n - \bar{\mathbf{Y}}_0) \ , \ \mathbf{Z}_n^{(s+v)} = \mathbf{K}_n^*(\mathbf{S}_n + \mathbf{V}_n) \ , \ \mathbf{K}_n^* = \mathbf{P}_n\mathbf{C}_n^{-1} \ ,$$

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

$$
\begin{aligned}
\text{(A.11)} \quad & \sqrt{m_n}\left[\frac{1}{m_n}\sum_{k=1}^{m_n}\mathbf{z}_k\mathbf{z}_k^{'} - \mathbf{\Sigma}_x\right] \\
&= \sqrt{m_n}\left[\frac{1}{m_n}\sum_{k=1}^{m_n}\mathbf{z}_k^{(x)}\mathbf{z}_k^{(x)'} - \mathbf{\Sigma}_x\right] + \frac{1}{\sqrt{m_n}}\sum_{k=1}^{m_n}\mathbf{E}[\mathbf{z}_k^{(s+v)}\mathbf{z}_k^{(s+v)'}] \\
&+ \frac{1}{\sqrt{m_n}}\sum_{k=1}^{m_n}\left[\mathbf{z}_k^{(s+v)}\mathbf{z}_k^{(s+v)'} - \mathbf{E}[\mathbf{z}_k^{(s+v)}\mathbf{z}_k^{(s+v)'}]\right] + \frac{1}{\sqrt{m_n}}\sum_{k=1}^{m_n}\left[\mathbf{z}_k^{(x)}\mathbf{z}_k^{(s+v)'} + \mathbf{z}_k^{(s+v)}\mathbf{z}_k^{(x)'}\right] .
\end{aligned}
$$

Then we can show that three terms except the first one of (A.11) are $o_p(1)$ (as in Theorem 4.1 of KS (2021), also see Chapter 5 of KSK (2018)) and the dominant term

$$
\text{(A.12)} \qquad \sqrt{m_n}\left[\frac{1}{m_n}\sum_{k=1}^{m_n}\mathbf{z}_k^{(x)}\mathbf{z}_k^{(x)'} - \mathbf{\Sigma}_x\right]
$$

is asymptotically normal as $m_n \to \infty$ $(n \to \infty)$, where $\mathbf{\Gamma}(h) = \mathbf{E}[\Delta\mathbf{x}_t\Delta\mathbf{x}_{t-h}^{'}]$ and

$$
\text{(A.13)} \qquad \mathbf{\Sigma}_x = \mathbf{f}_{\Delta x}(0) = \sum_{h=-\infty}^{+\infty}\mathbf{\Gamma}(h) .
$$

Then, the second term of (A.5) is $o_p(1)$ if $m = [n^{\alpha}]$ $(0 < \alpha < 0.8)$.

We show that the second, third and fourth terms on the right-hand-side of (A.11) are asymptotically negligible as $n \to \infty$. By modifying the proof of Proposition 2 of KS (2021), it is straightforward to show these conditions because of the independence assumption among $\mathbf{\Delta x}_i$, $\mathbf{s}_i$ and $\mathbf{v}_i$ $(i = 1, \cdots, n)$. We utilized the relation that

$$
\text{(A.14)} \qquad \sqrt{m_n}\frac{1}{m_n}\sum_{k=1}^{m_n}a_{kn}^{*} = \frac{1}{\sqrt{m_n}}2\sum_{k=1}^{m_n}\left[1 - \cos(\pi\frac{2k-1}{2n+1})\right] = O(\frac{m_n^{5/2}}{n^2}) ,
$$

and

$$
\text{(A.15)} \qquad \frac{1}{m}\sum_{k=1}^{m}2\cos(\pi\frac{2k-1}{2n+1}) = \frac{1}{m}\sum_{k=1}^{m}[e^{i\frac{2\pi}{2n+1}(k-\frac{1}{2})} + e^{-i\frac{2\pi}{2n+1}(k-\frac{1}{2})}] = \frac{1}{m}\frac{\sin(\frac{2\pi}{2n+1}m)}{\sin(\frac{\pi}{2n+1})} .
$$

Then, we find that (A.14) is $o(1)$ when $m_n = [n^{\alpha}]$ $(0 < \alpha < 0.8)$ while (A.15) is bounded when $m_n/n \to 0$ and $n \to \infty$.
Because of (A.14), $(1/m)\sum_{k=1}^{m}a_{kn}^{*} = O([m/n]^2)$. Uing (A.15), we have the consistency result of $\hat{\Sigma}_x$ in (i) under the condition $m/n \to 0$ as $n \to \infty$ as in Kunitomo

and Sato (2017) and KSK (2018).
(**Step 2**) : When we have the condition $0 < \alpha < 0.8$, we need to evaluate the limiting distribution of the first term of (A.11) because of (A.14). Instead of (A.12), we consider the asymptotic distribution of

$$s_{ij}^{(m)*} = \frac{1}{\sqrt{m}}[g_{ij}^{(m*)} - \mathbf{E}(g_{ij}^{(m*)})] \tag{A.16}$$

and

$$g_{ij}^{(m*)} = (\frac{1}{m}\sum_{k=1}^{m}\mathbf{z}_k^{(x)}\mathbf{z}_k^{(x)'})_{ij}\ (i,j = 1,\cdots,p)\ . \tag{A.17}$$

Then, we decompose

$$\begin{aligned} s_{ij}^{(m)*} &= \frac{1}{\sqrt{m}}\sum_{k=1}^{m}[\sum_{s=t=1}^{n} p_{ks}^2(r_{is}^*r_{js}^* - \mathbf{E}(r_{is}^*r_{js}^*))] \\ &\quad + \frac{1}{\sqrt{m}}\sum_{k=1}^{m}[\sum_{s\neq t=1}^{n} p_{ks}^2(r_{is}^*r_{jt}^* - \mathbf{E}(r_{is}^*r_{jt}^*))] \end{aligned} \tag{A.18}$$

and

$$\mathbf{r}_i^* = \Delta\mathbf{x}_i = \sum_{s=0}^{\infty}\mathbf{\Gamma}_s\mathbf{w}_{i-s}\ , \tag{A.19}$$

where $\mathbf{\Gamma}_s$ $(= (\gamma_{is}))$ are $p\times p$ matrices with $\mathbf{\Gamma}(h) = O(\rho^{|h|})$ $(0\leq\rho<1)$, and we consider $\mathbf{w}_i$ $(= \mathbf{e}_i^{(x)})$ as a sequence of mutually independent random variables with $\mathbf{E}[\mathbf{w}_i] = 0$, $\mathbf{E}[\mathbf{w}_i\mathbf{w}_i^{'}] = \mathbf{\Sigma}_v^{(x)}$ (>0).
When we have the condition $m_n/n \to 0$ as $n\to\infty$, the proof of Proposition 1 of KS (2021) implies that

$$\frac{1}{\sqrt{m_n}}[\sigma_{ij}^{(x)} - \mathbf{E}(g_{ij}^{(m*)})] = o(1)\ . \tag{A.20}$$

The evaluation of the limiting distribution of (A.16) or (A.18) is considerably simpler than that for (A.12).

We set $c_{st} = [(2n+1)/2m]a_{st}$ for a_{st} $(s,t = 1,\cdots,n)$ in (2.15). Then the first term of (A.18) is asymptotically negligible because $\sum_{k=1}^{n}p_{ks}^2 = [2m/[2n+1]c_{ss}$ and $\sum_{s=1}^{n}c_{ss}^2 = O(n)$ as given in Chapter 5 of KSK (2018). We show the asymptotic normality of the leading term

$$s_{ij}^{(m)**} = \frac{2\sqrt{m}}{2n+1}\sum_{s\neq t=1}^{n}c_{st}[r_{is}^*r_{jt}^* - \mathbf{E}(r_{is}^*r_{jt}^*)]\ , \tag{A.21}$$

where

$$c_{st} = \frac{2}{m}\sum_{k=1}^{m}s_{sk}s_{tk}\ \ ,\ s_{jk} = \cos\frac{2\pi}{2n+1}(j-\frac{1}{2})(k-\frac{1}{2})\ . \tag{A.22}$$

Under the stationarity condition of (A.19), the difference between (A.21) and the second term of (A.18) is asymptotically negligible.
The proof of the asymptotic normality requires a lengthy derivation, which is a modification of the method for the spectral density estimation used in the proof of Theorem 9.4.1 presented by Anderson (1971). Because some of our arguments are quite similar, we only repeat the essential arguments and differences. We provide the proof for the case when $p = 1$ and use the notation $\mathbf{\Gamma}_s = \gamma_s$ $(s = 0, 1, \cdots)$ and $s^{(m)**} = s_{ij}^{(m)**}$ because the proof of the general case when $p \geq 1$ can be obtained by using the standard device of $r_j^{**} = \mathbf{a}^{'}\mathbf{r}_j^{*}$ $(j = 1, \cdots, n)$ with an arbitrary $(p \times 1$ non-zero constant) vector $\mathbf{a}$.

Let $K_n = [n/m]$ be a sequence of positive integers and $K_n \to \infty$ $(n \to \infty)$. Then, given s, $c_{st} \to 0$ for $t - s > K_n$ as $m, n \to \infty$ and $m/n \to 0$. Then, by taking $t = s + k$ $(k = 1, \cdots, n - s)$ we rewrite

$$
\text{(A.23)} \quad s^{(m)**} = \frac{4\sqrt{m}}{2n+1} \sum_{t>s=1}^{n} c_{st}[r_s^* r_t^* - \mathbf{E}(r_s^* r_t^*)] \; ,
$$

$$
= \frac{4\sqrt{m}}{2n+1} \sum_{l,l'=1}^{\infty} \gamma_l \gamma_{l'} \sum_{s=1}^{n} \sum_{t=s+1, s-l \neq t-l'}^{n} c_{st} w_{s-l} w_{t-l'} \; .
$$

$$
= \frac{4\sqrt{m}}{2n+1} \sum_{l,l'=1}^{\infty} \gamma_l \gamma_{l'} \sum_{s=1}^{n} \sum_{k=1}^{n+1-s} \sum_{s=1, s-l \neq s+k-l'}^{n} c_{s,s+k} w_{s-l} w_{s+k-l'} \; .
$$

We truncate the sum $\sum_{l,l'=1}^{\infty}[\,\cdot\,]$, which decomposes as $(\sum_{l=1}^{r_n} + \sum_{l=r_n+1}^{\infty})\,(\sum_{l'=1}^{r_n} + \sum_{l'=r_n+1}^{\infty})[\,\cdot\,]$ by a sequence of sums $\sum_{l,l'=1}^{r_n}[\,\cdot\,]$ such that $r_n \to \infty$ and $\sum_{l=r_n+1}^{\infty} |\gamma_l| \to 0$. We can approximate the infinite sum by a finite sum because the remaing terms are of smaller order. The main term is

$$
\text{(A.24)} \qquad s_1^{(m)**} = \frac{4\sqrt{m}}{2n+1} \sum_{l,l'=1}^{r_n} \gamma_l \gamma_{l'} \sum_{s=1}^{n} \sum_{k=1, s-l \neq s+k-l'}^{n+1-s} c_{s,s+k} w_{s-l} w_{s+k-l'}
$$

$$
= \frac{4\sqrt{m}}{2n+1} \sum_{l,l'=1}^{r_n} \gamma_l \gamma_{l'} \sum_{h=l-l'+1, h \neq 0}^{n-q-l^{'}} \sum_{q=1-l}^{n-l} c_{q+l,q+h+l'} w_q w_{q+h} \; .
$$

Since some parts of the above summation (i.e., the terms in $\sum_{l-l' \leq h < 1}[\,\cdot\,]$, $\sum_{n-q-l' \leq h < n-q}[\,\cdot\,]$, $\sum_{1-l \leq q < 0}[\,\cdot\,]$ and $\sum_{n-l \leq q \leq n-1}[\,\cdot\,]$) can be of negligible order, we can approximate the summation as

$$
\text{(A.25)} \qquad s_{11}^{(m)**} = \frac{4\sqrt{m}}{2n+1} [\sum_{l=1}^{r_n} \gamma_l \sum_{l'=1}^{r_n} \gamma_{l'}] \sum_{h=l-l'+1}^{n-q-l^{'}} \sum_{q=1-l}^{n-l} c_{q+l,q+h+l'} w_q w_{q+h}
$$

$$
\sim \frac{4\sqrt{m}}{2n+1} [\sum_{l=1}^{r_n} \gamma_l \sum_{l'=1}^{r_n} \gamma_{l'}] \sum_{h=1}^{n-q} \sum_{q=1}^{n} c_{q+l,q+h+l'} w_q w_{q+h} \; ,
$$

where we denote $c_{s,t} = 0$ ($s > n$ or $t > n$) for the resulting notational convenience.

Let $m = [n^{\alpha}]$ ($0 < \alpha < 0.8$), $K_n = [n/m]$, $N_n = [n^{\delta/2}]$ ($\delta > 0$), and $M_n = [n^{1-\delta/2}]$ such that $1 - \delta/2 > 0$ and $\alpha + \delta/2 > 1$. Then, $K_n/N_n \to 0$, $N_n/n \to 0$ $\sqrt{m}/n \sim [1/\sqrt{n}][1/\sqrt{K_n}]$ and $M_n \sim n/N_n$ as $n \to \infty$. In the following we utilize the relation $c_{q+l,q+h+l'} - c_{q,q+h} = o(1)$ for $l, l' = 1, \cdots, r_n$ if we take r_n such that $r_n \times m_n/n \to 0$ as $n, m_n \to \infty$. This is because

$$\begin{aligned} & \sin 2\pi m[\frac{2q+h+l+l'}{2n+1}] - \sin 2\pi m[\frac{2q+h}{2n+1}] \\ = & \sin 2\pi m[\frac{2q+h}{2n+1}][\cos 2\pi m(\frac{l+l'}{2n+1}) - 1] + \cos 2\pi m[\frac{2q+h}{2n+1}] \sin 2\pi m[\frac{l+l'}{2n+1}] \to 0 \end{aligned}$$

as $n \to \infty$.
Furthermore, by using that some parts of (A.25) are of smaller orders as $n \to \infty$ (the terms in $\sum_{h=K_n+1}[\,\cdot\,]$), we can apply the CLT to

$$s_{11}^{(m)***} = 2[\sum_{l=1}^{r_n} \gamma_l]^2 \frac{1}{\sqrt{n}} \frac{1}{\sqrt{K_n}} \sum_{q=1}^{n} \sum_{h=1}^{K_n} c_{q,q+h} w_q w_{q+h} \ , \tag{A.26}$$

where we denote $c_{q,q+h} = 0$ ($q + h > n$) for notational convenience. We notice that $c_{q,q+h}$ ($q = 1, \cdots, n$) is a sequence of bounded real numbers.
Let

$$W_{qn} = \frac{1}{\sqrt{K_n}} \sum_{h=1}^{K_n} c_{q,q+h} w_q w_{q+h} \tag{A.27}$$

and

$$U_{jn} = \frac{1}{\sqrt{N_n}} [W_{(j-1)N_n+1,n} + \cdots + W_{jN_n-K_n,n}] \ \ (j = 1, \cdots, M_n) \ . \tag{A.28}$$

Then, we find that $\mathbf{E}[W_{q,n}] = 0$, $\mathbf{E}[W_{q,n}W_{q+h,n}] = 0$ (h is any non-zero integer), $\mathbf{E}[W_{q,n}^2]$ are bounded. Further, we have that $U_{1,n}, \cdots, U_{M_n,n}$ are mutually independent and $\mathbf{E}[U_{i,n}^4]$ ($i = 1, M_n$) are uniformly bounded using the assumption of the boundedness of the 4-th order moments of W_q ($q = 1, \cdots, n$). Since other terms except the leading term are stochastically of the smaller order, we can ignore them when evaluating the limiting distribution, and we apply the Liyaponov-type CLT. By using the relation that

$$\frac{1}{\sqrt{n}} \sum_{q=1}^{n} W_{qn} - \frac{1}{\sqrt{M_n}} \sum_{j=1}^{M_n} U_{jn} \xrightarrow{p} 0 \tag{A.29}$$

as $n \to \infty$, the remaining terms are of smaller order (i.e. K_n terms in each U_{jn} ($j = 1, \cdots, M_n$)) when $m, n \to \infty$ and $m = n^{\alpha}, 0 < \alpha < 0.8$ because of $K_n/N_n \to 0$.

Then we have the asymptotic normality of (A.19) when $p = 1$. By using the relation $\sum_{s,t=1}^{n} c_{st}^2 = (n + 1/2)^2/m$ ((5.16) of KSK (2018)) and

$$\text{(A.30)} \qquad 4[\sum_{j=-\infty}^{\infty} \gamma_j]^2 \sum_{g=1}^{n} \sum_{h=1}^{K_n} c_{g,g+h} [\sigma_v^{(x)}]^4 \sim 2[\sum_{j=-\infty}^{\infty} \gamma_j]^2 \sum_{s,t=1}^{n} c_{st}^2 [\sigma_v^{(x)}]^4 \ ,$$

we have the desired result of the asymptotic variance when $p = 1$.
When $p \geq 1$, we can evaluate the asymptotic covariance by calculating the covariance of $\sum_{a,b,q,h} c_{q,q+h} \boldsymbol{\gamma}_{as} \mathbf{w}_q \boldsymbol{\gamma}_{bt} \mathbf{w}_{q+h}$ and $\sum_{c,d,q',h'} c_{q',q'+h'} \boldsymbol{\gamma}_{cs} \mathbf{w}_{q'} \boldsymbol{\gamma}_{bt} \mathbf{w}_{q'+h'}$, where $\boldsymbol{\gamma}_{ac}$ represents the a-th row vector of $\boldsymbol{\Gamma}_s$. Then, after a straightforward evaluation, we finally find the asymptotic covariance in Theorem A.1 as $\sigma_{ac}^{(x)} \sigma_{bd}^{(x)} + \sigma_{ad}^{(x)} \sigma_{bc}^{(x)}$ $(a, b, c, d = 1, \cdots, p)$. (The argument here is essentially the same as the one in Chapter 5 of KSK (2018) in the high-frequency financial formulation.)
(**Q.E.D**)

Proof of Theorem 3.1 : We use the representation

$$\text{(A.31)} \qquad \hat{\boldsymbol{\beta}}_m - \boldsymbol{\beta} = \mathbf{G}_{22}^{-1} [(\mathbf{0}, \mathbf{I}_k) \mathbf{G} \begin{pmatrix} 1 \\ -\boldsymbol{\beta} \end{pmatrix} \ .$$

Because $(1/m)\mathbf{G}_m \xrightarrow{p} \boldsymbol{\Sigma}_{22}$ $(m/n \to 0, n \to \infty)$ and under the assumption that $\boldsymbol{\Sigma}_{22}$ is a positive definite matrix, we investigate the asymptotic distribution of

$$\text{(A.32)} \qquad \sqrt{m_n} [\hat{\boldsymbol{\beta}}_m^* - \boldsymbol{\beta}] = \boldsymbol{\Sigma}_{22}^{-1} \frac{1}{\sqrt{m}} (\mathbf{0}, \mathbf{I}_k) \mathbf{G} \begin{pmatrix} 1 \\ -\boldsymbol{\beta} \end{pmatrix} \ .$$

Then, the asymptotic variance-covariance matrix can be written as

$$\text{(A.33)} \qquad \mathbf{AV}[\hat{\boldsymbol{\beta}}_m] = \boldsymbol{\Sigma}_{22}^{-1} \mathbf{Cov}[(\mathbf{0}, \mathbf{I}_k) \mathbf{S} \mathbf{b} \ , \ \mathbf{b}' \mathbf{S} \begin{pmatrix} \mathbf{0}' \\ \mathbf{I}_k \end{pmatrix}] \boldsymbol{\Sigma}_{22}^{-1} \ ,$$

where $\mathbf{S} = \sqrt{m_n}[\mathbf{G} - \boldsymbol{\Sigma}]$ $(= (s_{jk}))$ and $\mathbf{b} = \begin{pmatrix} 1 \\ -\boldsymbol{\beta} \end{pmatrix}$ $(= (b_j))$.
By using Theorem A.1, we can evaluate the (l, l')-th element $(l, l' = 2, \cdots, k+1)$ as

$$\begin{aligned} \mathbf{Cov}[\sum_{j=1}^{k+1} b_j s_{jl} \sum_{j'=1}^{k+1} b_{j'} s_{j'l'}] &= \sum_{j,j'=1}^{k+1} b_j b_{j'} (\sigma_{j,j'} \sigma_{l,l'} + \sigma_{j,l'} \sigma_{j',l}) \\ &= \sigma_{l,l'} \sum_{j=1}^{k+1} b_j [\sum_{l'=1}^{k+1} b_{j'} \sigma_{j,j'}] + [\sum_{j=1}^{k+1} b_j \sigma_{j,l'}][\sum_{j'=1}^{k+1} b_{j'} \sigma_{l,j'}] \\ &= \sigma_{l,l'} \sigma_{11.2} \end{aligned}$$

because $[\boldsymbol{\sigma}_{21}, \boldsymbol{\Sigma}_{22}]\mathbf{b} = \mathbf{0}$ and

$$\text{(A.34)} \qquad [\boldsymbol{\sigma}_{11}, \boldsymbol{\sigma}_{12}]\mathbf{b} = \sigma_{11} - \boldsymbol{\sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{\sigma}_{21} \ .$$

Then we have the result of the asymptotic variance-covariance matrix.
(**Q.E.D**)

Proof of Corollary 4.1 : We use the representation

$$\hat{\mathbf{B}}_m - \mathbf{B} = (\mathbf{W}_n^{*'}\mathbf{W}_n^*)^{-1}\mathbf{W}_n^{*'}\mathbf{U}_n^* \ , \tag{A.35}$$

where $\mathbf{U}_n^* = \mathbf{J}_m\mathbf{P}_n\mathbf{C}_n^{-1}\mathbf{U}_n$. By using a similar argument as the proof of Theorem 3.1 under the assumption of (4.4), we find that

$$\mathbf{AV}[\hat{\mathbf{B}}_m] = \mathbf{\Sigma}_{w^*}^{-1}\mathbf{Cov}[\frac{1}{\sqrt{m}}\mathbf{W}_n^{*'}\mathbf{U}_n^*, \frac{1}{\sqrt{m}}\mathbf{W}_n^{*'}\mathbf{U}_n^*]\mathbf{\Sigma}_{w^*}^{-1} \ . \tag{A.36}$$

Then, by using Theorems A.1 and 3.1 we have the result.
(**Q.E.D**)

APPENDIX B : Figures

In this Appendix, we gather some figures cited in Section 6. All computations have been done by x12siml6 written in R (Sato (2020)), which will be available in the near future.

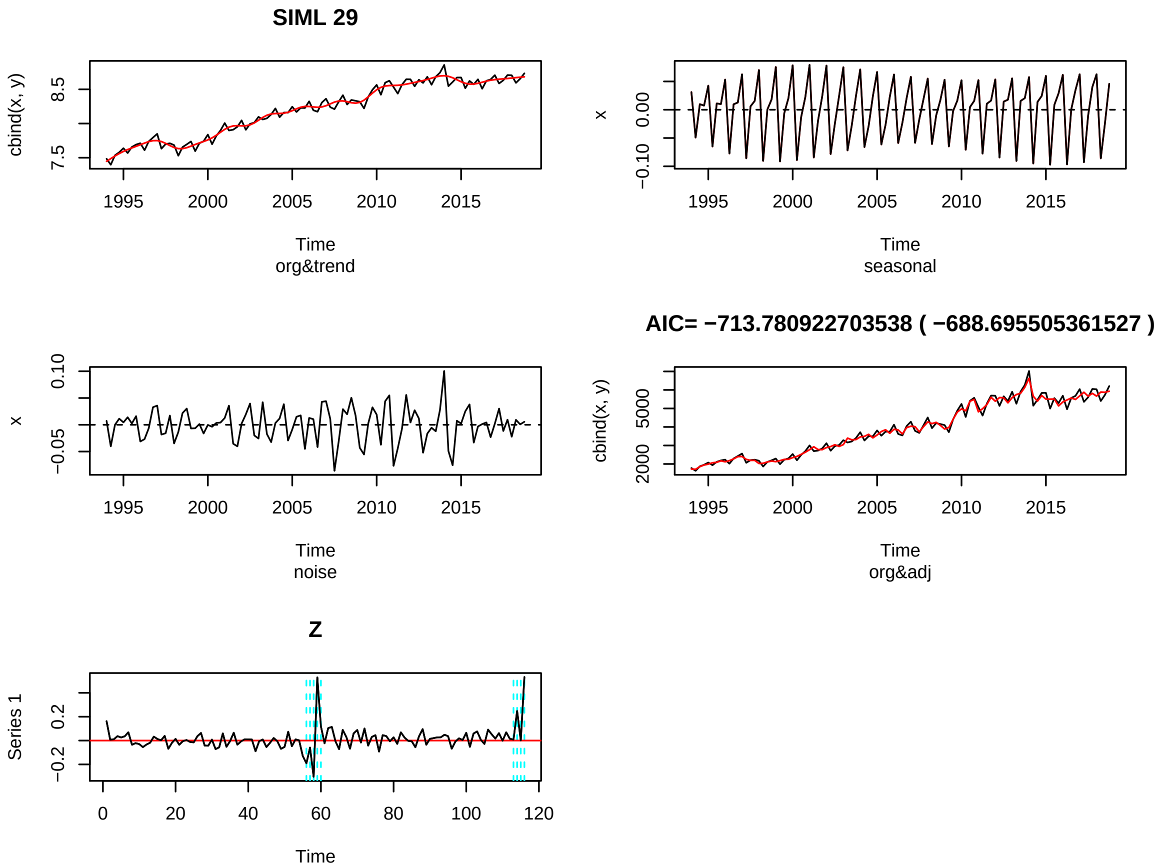

Figure 1: Macroconsumption (Data are the Quarterly real consumption of durable goods (after log-transformation) between 1994Q1-2018Q4 published by the Economic Social Research Institute (ESRI), Cabinet Office, Japan.)

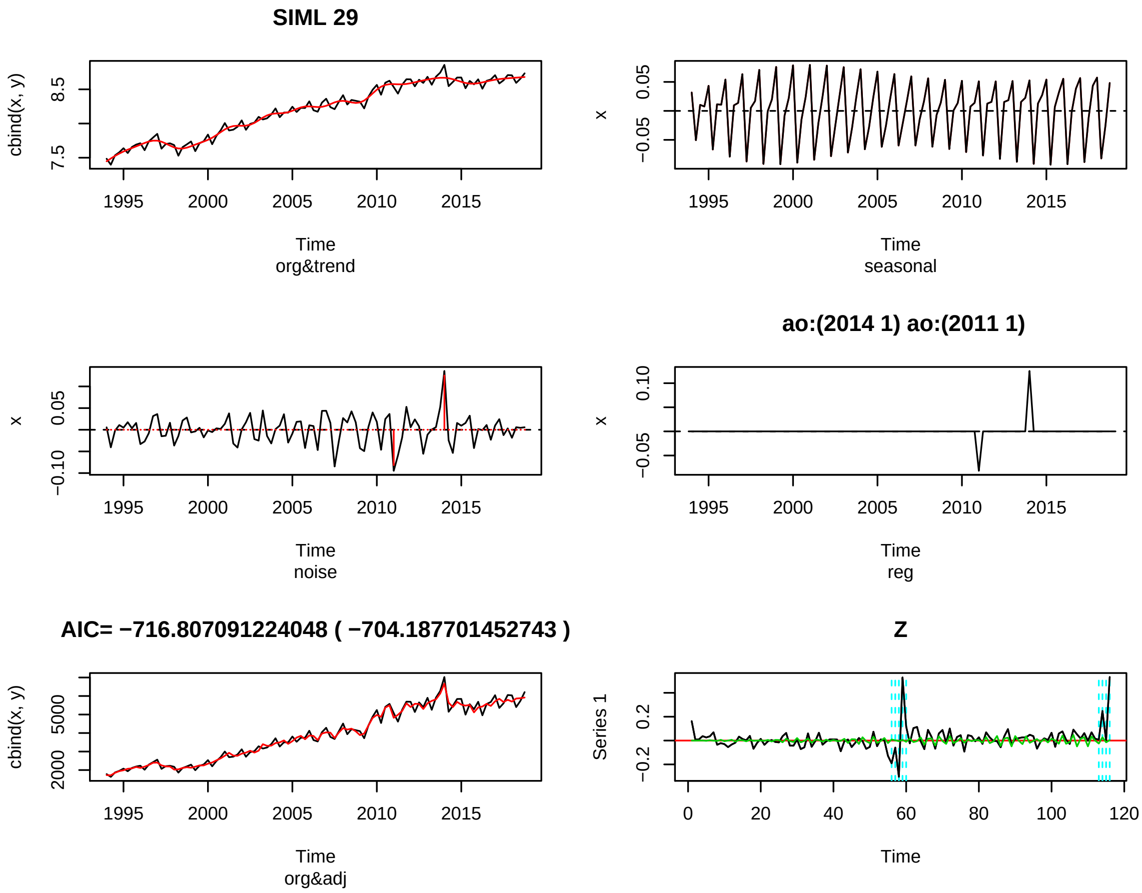

Figure 2: Macroconsumption (Data are the Quarterly real consumption of durable goods (after log-transformation) between 1994Q1-2018Q4 published by the Economic Social Research Institute (ESRI), Cabinet Office, Japan.)

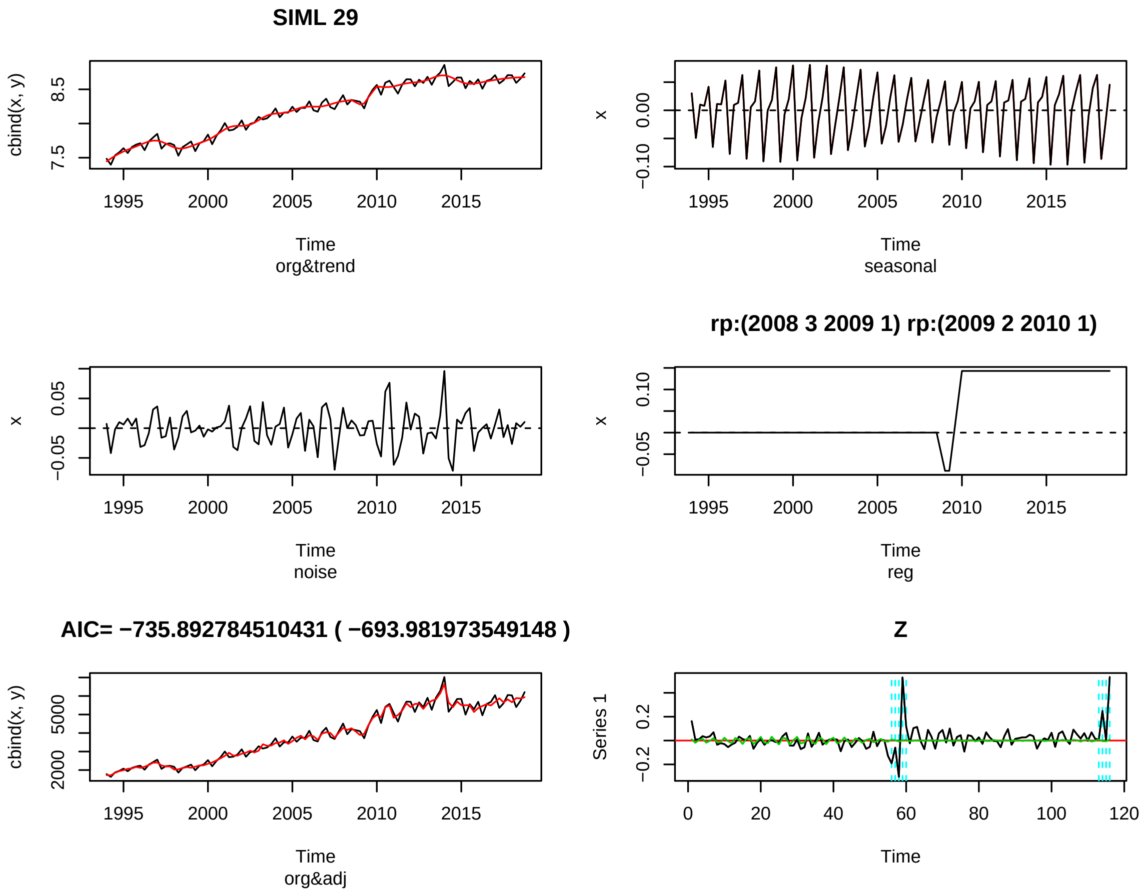

Figure 3: Macroconsumption (Data are the Quarterly real consumption of durable goods (after log-transformation) between 1994Q1-2018Q4 published by the Economic Social Research Institute (ESRI), Cabinet Office, Japan.)

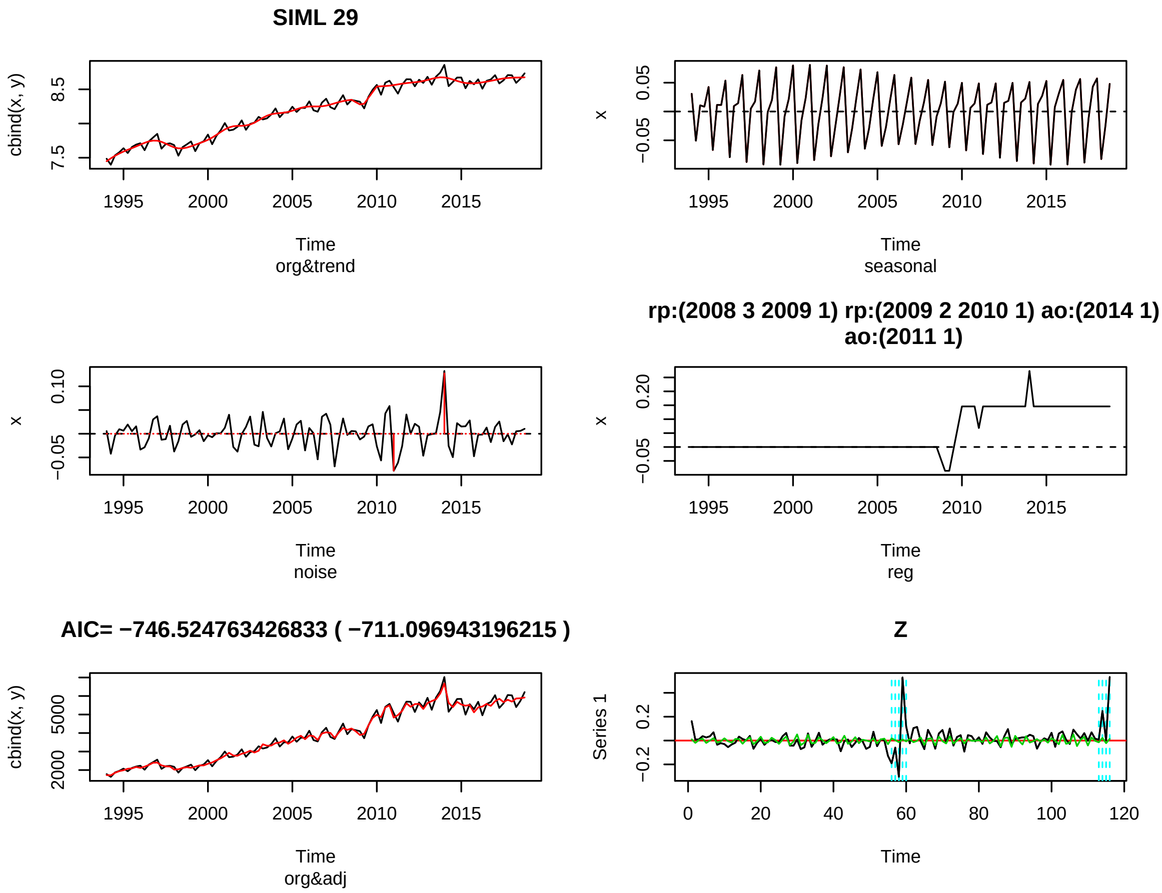

Figure 4: Macroconsumption (Data are the Quarterly real consumption of durable goods (after log-transformation) between 1994Q1-2018Q4 published by the Economic Social Research Institute (ESRI), Cabinet Office, Japan.)