CIRJE-F-977

Trend, Seasonality and Economic Time Series: the Nonstationary Errors-in-variables Models

Naoto Kumitomo
The University of Tokyo

Seisho Sato
The University of Tokyo

June 2015

Trend, seasonality and economic time series : the nonstationary errors-in-variables models *

Naoto Kunitomo †
and
Seisho Sato ‡

May 25, 2015

Abstract

The use of seasonally adjusted (official) data may have statistical problem because it is a common practice to use *X-12-ARIMA* in *the official seasonal adjustment*, which adopts the univariate ARIMA time series modeling with some refinements. Instead of using the seasonally adjusted data, for estimating the structural parameters and relationships among non-stationary economic time series with seasonality and noise, we propose a new method called the Separating Information Maximum Likelihood (SIML) estimation. We show that the SIML estimation can identify the nonstationary trend, the seasonality and the noise components, which have been observed in many macro-economic time series, and recover the structural parameters and relationships among the non-stationary trends with seasonality. The SIML estimation is consistent and it has the asymptotic normality when the sample size is large. Based on simulations, we find that the SIML estimator has reasonable finite sample properties and thus it would be useful for practice.

Key Words

Nonstationary economic time series, Errors-variables models, Non-stationary trend and seasonality, Official Seasonal Adjustment, Structural relationships, Seasonal frequency, SIML method, Asymptotic properties.

*2015-1-8. We thank Yajima Yoshihiro and Shimotsu Katsumi for their comments.

†Graduate School of Economics, University of Tokyo, Bunkyo-ku, Hongo 7-3-1, Tokyo 113-0033, JAPAN, kunitomo@e.u-tokyo.ac.jp

‡Graduate School of Economics, University of Tokyo, Bunkyo-ku, Hongo 7-3-1, Tokyo 113-0033, JAPAN, seishou@e.u-tokyo.ac.jp.

1. Introduction

There have been vast literatures on the use of statistical time series analysis of macro economic time series. One important distinction of macro-economic time series from the standard time series analysis in other areas has been the mixture of non-stationarity and measurement errors including the apparent seasonality although the analysis of seasonality in economic time series has been often ignored. Although there have been many attempts to deal with the stationarity, the non-stationarity and seasonality separately in macro-economic time series analysis, there have been some need to incorporate these different aspects of economic time series in a unifying way.

For an expository purpose, we illustrate two macro time series in Figure 1-1, which gives the original quarterly data of the real GDP and fixed investment published by the Cabinet Office of Japanese Government. We have standardized two time series such that the data in scale have the similar value such that we can observe clear common trends, common seasonality and noise in two important time series, which are quite typical in Japanese quarterly GDP data. An interesting empirical question here would be to find reasonable estimates of correlations of trends and *seasonalities* among two nonstationary macro time series we observe quarterly.

The use of seasonally adjusted data has been a common practice among many economists in business, but then we have to cope with the problem of *the official seasonal adjustments method* producing the published data for macro-economic variables. It has been a common practice to use *X-12-ARIMA* in many official agencies including the Cabinet office of the Japanese Government (i.e. they produce the official GDP in Japan), but usually they use the univariate ARIMA time series modeling with some refinements, which has been a common practice since Box and Jenkins (1970). (See Findley et al. (1998) for the details of *X-12-ARIMA*.)

In this paper, instead of using the seasonally adjusted data and investigate the statistical relationships among macro time series, we propose to use the separating information maximum likelihood (SIML) estimation method, which is new to the macro-time series analysis. We shall show that this macro-SIML method is useful to identify the trend, the seasonal, the cycle, and the irregular noise components in the non-stationary errors-in-variables model. The statistical time series model we shall use is an extension of the univariate decomposition of its components by Akaike (1980) and Kitagawa (2010) in different perspectives.

There have been many studies on the errors-in-variables models, which are closely related to the classical multivariate analysis including the factor models and simultaneous equations models. (See Anderson (1984, 2003) and Fuller (1987) for the related issues.) It has been known that there are serious identification problems occurred in the classical errors-in-variables models and the estimation problem of unknown parameters in the underlying hidden variables has some difficulty. In this paper we shall show that in the mixture of non-stationary and stationary components including the seasonal factor we can identify the unknown parameters generating the hidden time series components. The typical examples are the variance-covariance matrix of the hidden trend variables which follow the random walks and the variance-covariance matrix of the hidden seasonal variables. We shall show that the SIML estimation can estimate the trend, the seasonality and noise components from the observed time series, and recover the structural relationships among the non-stationary trend and seasonality. Also we show that the SIML estimator is consistent and it has the asymptotic normality when the sample size is large. Based on a set of simulations, we find that the SIML estimator has reasonable finite sample properties and thus it would be useful for practice.

A motivation of our study is the fact that it is not a trivial task to handle the exact likelihood function and calculate the exact ML estimator of structural relationships among trends from non-stationary time series data when the observed time series have seasonality and noise in the nonstationary errors-in-variables models. This aspect is quite important for the analysis of multivariate macro-economic time series because the modeling the seasonality and noise could have possible misspecifications. In this paper we regard the seasonality and noise as the measurement errors. Instead of calculating the exact likelihood function, we try to separate the information of the signal part and the measurement errors part from the likelihood function and then use each information separately. This procedure approximates the maximization of the likelihood function and make the estimation procedure applicable to multivariate non-stationary time series data in a straightforward manner. We denote our estimation method as the Separating Information Maximum Likelihood (SIML) estimator because it gives an extension of the standard ML estimation method. The main merit of the SIML estimation is its simplicity and then it can be practically used for the multivariate non-stationary economic time series.

Earlier and related literatures in econometrics are Engle and Granger (1987) and Johansen (1995), which have dealt with the multivariate nonstationary and stationary time series and developed the notion of co-integration. The problem of the present paper is related to their work, but has different aspects because of our analysis on the non-stationary seasonality and measurement errors in the nonstationary errors-in-variable models. Also our approach is related to the earlier studies of Engle (1974) and Phillips (1991) on the spectral regression analysis since our estimation method is related to the spectral analysis of trend, seasonal and noise frequencies because the former considred the stationary time series while the latter investigated only the non-stationary trends. In this sense our analysis could be regarded as an extension of their earlier works.

In Section 2 we present a general formulation of the problem and give simple examples to illustrate the problem in this paper. Then in Section 3 we develop the nonstationary model with random walk plus noise, and in Section 4 we develop the macro-SIML estimation method. In Section 5 we discuss our method to analyze the seasonal components. In Section 6 we discuss some simulation results and then we have some concluding remarks in Section 7. The proofs of Theorems in this paper are based on the modifications of the results by Kunitomo and Sato (2008, 2011, 2013) for the financial-SIML estimation. Since they are often quite similar to their mathematical proofs, we omit the details in this version.

2. The general problem and some examples

2.1 The general problem

Let y_{ij} be the $i-$th observation of the $j-$th time series at i for $i = 1,\cdots,n; j = 1,\cdots,p$. We set $\mathbf{y}_i = (y_{1i},\cdots,y_{pi})^{'}$ be a $p \times 1$ vector and $\mathbf{Y}_n = (\mathbf{y}_i^{'})$ $(= (y_{ij}))$ be an $n \times p$ matrix of observations. ($\mathbf{y}_0$ is the initial observation vector.) We consider the situation when the underlying non-stationary trends $\mathbf{x}_i$ $(= (x_{ji}))$ $(i = 1,\cdots,n)$ are not necessarily the same as the observed time series and let $\mathbf{s}_i^{'} = (s_{1i},\cdots,s_{pi})$ and $\mathbf{v}_i^{'} = (v_{1i},\cdots,v_{pi})$ be the vectors of the seasonal components, and the noise components, respectively, which are independent of $\mathbf{x}_i$. Then we use the additive decomposition model (see Kitagawa (2010))

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

where $\mathbf{x}_i$ $(i = 1, \cdots, n)$ are a sequence of non-stationary trend components satisfying

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

with $\mathcal{L}\mathbf{x}_i = \mathbf{x}_{i-1}$, $\Delta = 1 - \mathcal{L}$, $\mathcal{E}(\mathbf{w}_i^{(x)}) = \mathbf{0}$, $\mathcal{E}(\mathbf{w}_i^{(x)}\mathbf{w}_i^{(x)'}) = \mathbf{\Sigma}_x$, and $\mathbf{s}_i$ $(i = 1, \cdots, n)$ are a sequence of seasonal components satisfying

$$(1 - \mathcal{L}^s)\mathbf{s}_i = \mathbf{w}_i^{(s)} \tag{2.3}$$

or alternatively

$$(1 + \mathcal{L} + \cdots + \mathcal{L}^{s-1})\mathbf{s}_i = \mathbf{w}_i^{(s)} \tag{2.4}$$

with $\mathcal{L}^s\mathbf{s}_i = \mathbf{s}_{i-s}$, $\mathcal{E}(\mathbf{w}_i^{(s)}) = \mathbf{0}$, $\mathcal{E}(\mathbf{w}_i^{(s)}\mathbf{w}_i^{(s)'}) = \mathbf{\Sigma}_s$, and $\mathbf{v}_i$ are a sequence of independent noise components with $\mathcal{E}(\mathbf{v}_i) = \mathbf{0}$, $\mathcal{E}(\mathbf{v}_i\mathbf{v}_i^{'}) = \mathbf{\Sigma}_v$.
We assume that $\mathbf{w}_i^{(x)}, \mathbf{w}_i^{(s)}$ and $\mathbf{v}_i$ are the sequence of i.i.d. random variables with $\mathbf{\Sigma}_v$ being positive definite and finite, and the random variables $\mathbf{w}_i^{(x)}, \mathbf{w}_i^{(s)}$ and $\mathbf{v}_i$ are mutually independent.

The main purpose of this study is to estimate the structural parameters and the structural relationships among the hidden random variables; the trend components and seasonal components in particular when we have stationary and non-stationary errors-in-variables models. Let $\boldsymbol{\beta}$ be a $p \times 1$ vector and we want to estimate

$$\boldsymbol{\beta}^{'}\mathbf{x}_i = O_p(1) \quad (i = 1, \cdots, n), \tag{2.5}$$

when we have the observations of $p \times 1$ vectors $\mathbf{y}_i$ $(i = 1, \cdots, n)$. More generally, let $\mathbf{B}$ be a $q \times p$ $(q \leq p)$ matrix and we want to estimate

$$\mathbf{B}\mathbf{x}_i = O_p(1) \quad (i = 1, \cdots, n) \tag{2.6}$$

when we have the observations of $p \times 1$ vectors $\mathbf{y}_i$ $(i = 1, \cdots, n)$. Similarly, some structural relations among seasonal components can be written as

$$\mathbf{B}_s\mathbf{s}_i = O_p(1) \quad (i = 1, \cdots, n), \tag{2.7}$$

and they imply that the observed multivariate time series have common seasonality.

2.2 Examples

We give simple examples when $p = 2$ for illustrating the problem of nonstationary errors-in-variables models.

Example 1 : Assume that for the sequence of observable random vectors $\mathbf{y}_i = (y_{1i}, y_{2i})^{'}$, the random variables $x_{1i} = \nu_i = \beta_2\mu_i$ and $x_{2i} = \mu_i$ satisfy $\mu_i = \mu_{i-1} + w_i^{(x)}$ $(i = 1, \cdots, n)$ and $w_i^{(x)}$ are i.i.d. random variables with $\mathcal{E}(w_i^{(x)}) = 0$ and $\mathcal{E}(w_i^{(x)2}) = \sigma_x^2$. Then we can write

$$\mathbf{y}_i = \begin{pmatrix} \beta_2 \\ 1 \end{pmatrix} \mu_i + \mathbf{v}_i \ . \tag{2.8}$$

Since μ_i follows the random walk model, the invariance (CLT) principle says that as $n \to \infty$,

$$\frac{1}{n^2}\sum_{i=1}^{n} \mu_i^2 \xrightarrow{p} \sigma_x^2 \int_0^1 B_s^2 ds \ , \tag{2.9}$$

where B_s is the standard Brownian Motion on $[0, 1]$.

If we multiply the vector $\boldsymbol{\beta}^{'} = (1, -\beta_2)$ to (2.7) from the left, we have the relation

$$\boldsymbol{\beta}^{'}\mathbf{y}_i = u_i \; (= \boldsymbol{\beta}^{'}\mathbf{v}_i) \; , \tag{2.10}$$

which is a structural equation.

Example 2 : We take the case when $\mathbf{x}_i = \boldsymbol{\mu}_i$, and $\boldsymbol{\mu}_i = \boldsymbol{\mu}_{i-1} + \mathbf{w}_i^{(x)}$, which has been often called *spurious regression.* It can be written as

$$\mathbf{y}_i = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \boldsymbol{\mu}_i + \mathbf{v}_i \tag{2.11}$$

and the dimension of random walk is 2 and $\boldsymbol{\beta}^{'}\mathbf{y}_i = \boldsymbol{\beta}^{'}\boldsymbol{\mu}_i + u_i$, $u_i = \boldsymbol{\beta}_x^{'}\mathbf{v}_i$ for any $\boldsymbol{\beta} \neq \mathbf{0}$ (the non-stationary term of $\boldsymbol{\beta}^{'}\boldsymbol{\mu}_i$ cannot be disappeared).

Example 3 : Assume that the random vectors $\mathbf{s}_i = (s_{1i}, s_{2i})^{'}$ with $s_{1i} = \nu_i^{(s)} = \beta_2^{(s)}\mu_i^{(s)}$ and $s_{2i} = \mu_i^{(s)}$ satisfy $\mu_i^{(s)} = \mu_{i-s}^{(s)} + w_i^{(s)}$ ($s \geq 1$; $i = 1, \cdots, n$) and $w_i^{(s)}$ are i.i.d. random variables with $\mathcal{E}(w_i^{(s)}) = 0$ and $\mathcal{E}(w_i^{(s)2}) = \sigma_s^2$. Then we can write

$$\mathbf{y}_i = \begin{pmatrix} \beta_2^{(s)} \\ 1 \end{pmatrix} \mu_i + \mathbf{v}_i \; . \tag{2.12}$$

If we multiply the vector $\boldsymbol{\beta}_s^{'} = (1, -\beta_2^{(s)})$ to (2.11) from the left, we have the relation

$$\boldsymbol{\beta}_s^{'}\mathbf{y}_i = u_i \; (= \boldsymbol{\beta}_s^{'}\mathbf{v}_i) \tag{2.13}$$

and $\mathbf{y}_i$ has the common seasonal components.

Example 4 : We consider the situation when $\mathbf{x}_i = \boldsymbol{\mu}_i$, $\boldsymbol{\mu}_i = \boldsymbol{\mu}_{i-1} + \mathbf{w}_i^{(x)}$ with $\boldsymbol{\Sigma}_x = \sigma_x^2\mathbf{I}_2$ (which is proportional to the identity) as the nonstationary trends and $\mathbf{s}_i = (s_{1i}, s_{2i})'$ with $s_{1i} = \nu_i^{(s)} = \beta_2^{(s)}\mu_i^{(s)}$, $s_{2i} = \mu_i^{(s)}$, $\mu_i^{(s)} = \mu_{i-s}^{(s)} + w_i^{(s)}$ ($w_i^{(s)}$ are i.i.d. random variables) and $\mathbf{\Sigma}_s \geq 0$ (non-negative definite) as the nonstationary seasonals. In this case nonstationary trends do not have any common trend, but there is a common nonstationary seasonals. The standard regression of one nonstationary variable on another nonstationary variable does not necessarily give any meaningful information on the underlying relationships with trends and seasonals.

3. The Case without Seasonality

Let $p \geq 2$ and $\mathbf{s}_i = \mathbf{0}$. We consider the multivariate time series model having the representation

$$\mathbf{y}_i = \mathbf{x}_i + \mathbf{v}_i = \mathbf{\Pi}\boldsymbol{\mu}_i + \mathbf{v}_i \ , \tag{3.1}$$

where $\mathbf{w}_i^{(x)} = \Delta\mathbf{x}_i$, $\mathcal{E}(\mathbf{w}_i^{(x)}) = \mathbf{0}$, and $\mathcal{E}(\mathbf{w}_i^{(x)}\mathbf{w}_i^{(x)'}) = \mathbf{\Sigma}_x$. We assume that the rank of $p \times q$ matrix $\mathbf{\Pi}$ is q ($\leq p$), $\boldsymbol{\mu}_i$ are $q \times 1$ vectors, and there exists a $q \times p$ matrix $\mathbf{B}$ such that $\mathbf{B}\mathbf{y}_i = \mathbf{u}_i$ ($= \mathbf{B}\mathbf{v}_i$), which are the set of q structural equations.

We consider the situation when $\Delta\mathbf{x}_i$ and $\mathbf{v}_i$ ($i = 1, \cdots, n$) are independent and each component vectors are independently, identically and normally distributed as $N_p(\mathbf{0}, \mathbf{\Sigma}_x)$ and $N_p(\mathbf{0}, \mathbf{\Sigma}_v)$, respectively. We use an $n \times p$ matrix $\mathbf{Y}_n = (\mathbf{y}_i')$ and consider the distribution of $np \times 1$ random vector $(\mathbf{y}_1', \cdots, \mathbf{y}_n')'$. Given the initial condition $\mathbf{y}_0$, we have

$$(3.2) \qquad \mathbf{Y}_n \sim N_{n\times p}\left(\mathbf{1}_n \cdot \mathbf{y}_0^{'}, \mathbf{I}_n \otimes \boldsymbol{\Sigma}_v + \mathbf{C}_n\mathbf{C}_n^{'} \otimes \boldsymbol{\Sigma}_x\right) ,$$

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

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

Then given the initial condition $\mathbf{y}_0$ the maximum likelihood (ML) estimator can be defined as the solution of maximizing the log-likelihood function except a constant as

$$\begin{aligned} L_n^* &= \log |\mathbf{I}_n \otimes \boldsymbol{\Sigma}_v + \mathbf{C}_n\mathbf{C}_n^{'} \otimes \boldsymbol{\Sigma}_x|^{-1/2} \\ &\quad -\frac{1}{2}[vec(\mathbf{Y}_n - \bar{\mathbf{Y}}_0)^{'}]^{'}[\mathbf{I}_n \otimes \boldsymbol{\Sigma}_v + \mathbf{C}_n\mathbf{C}_n^{'} \otimes \boldsymbol{\Sigma}_x]^{-1}[vec(\mathbf{Y}_n - \bar{\mathbf{Y}}_0)^{'}] \end{aligned}$$

and

$$(3.4) \qquad \bar{\mathbf{Y}}_0 = \mathbf{1}_n \cdot \mathbf{y}_0^{'} .$$

We transform $\mathbf{Y}_n$ to $\mathbf{Z}_n$ $(= (\mathbf{z}_k^{'}))$ by

$$(3.5) \qquad \mathbf{Z}_n = \mathbf{P}_n\mathbf{C}_n^{-1}\left(\mathbf{Y}_n - \bar{\mathbf{Y}}_0\right)$$

where

$$(3.6) \qquad \mathbf{C}_n^{-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}_{n\times n} ,$$

and

$$(3.7)\quad \mathbf{P}_n = (p_{jk}^{(n)})\ ,\ p_{jk}^{(n)} = \sqrt{\frac{2}{n+\frac{1}{2}}}\cos\left[\frac{2\pi}{2n+1}(k-\frac{1}{2})(j-\frac{1}{2})\right]\ .$$

By 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

$$d_k = 2[1-\cos(\pi(\frac{2k-1}{2n+1}))]\ (k=1,\cdots,n)\ .$$

Then the log-likelihood function is proportional to

$$(3.8)\quad L_n = \sum_{k=1}^{n}\log|a_{kn}\mathbf{\Sigma}_v + \mathbf{\Sigma}_x|^{-1/2} - \frac{1}{2}\sum_{k=1}^{n}\mathbf{z}_k^{'}[a_{kn}\mathbf{\Sigma}_v + \mathbf{\Sigma}_x]^{-1}\mathbf{z}_k\ ,$$

where

$$(3.9)\qquad a_{kn}\ (= d_k)\ = 4\sin^2\left[\frac{\pi}{2}\left(\frac{2k-1}{2n+1}\right)\right]\ (k=1,\cdots,n)\ .$$

Since we are dealing with the errors-in-variables model, there is an issue if we can identify the structural equation of our interest. When $\mathbf{x}_i$ are i.i.d. random variables, for instance, the coefficient parameters are not identified without some further restrictions. In the classical case when the observed random vectors $\{\mathbf{y}_i\}$ are independent, we need to impose some conditions on the covariance matrix (such as the homoscedasticity and zero covariance) when we have the functional relationship model in Example 1 except $x_i\ (= \mu_i\ ; i=1,\cdots,n)$ with

$$(3.10)\qquad \frac{1}{n}\sum_{i=1}^{n}\mu_i^2 \xrightarrow{p} \sigma_x^2\ .$$

(See Fuller (1987) for the details of such conditions.) Here we say that the parameter vector $\boldsymbol{\theta}\ (= (\theta_j))$ is identified if $\boldsymbol{\theta} \neq \boldsymbol{\theta}^{'}$ implies that $L_n(\theta) \neq L_n(\theta^{'})$.

We illustrate our arguments on the likelihood function when $p = 2$ and $q = 1$. If $\mathbf{\Sigma}_x$ is degenerate (i.e. $\text{rank}(\mathbf{\Sigma}_x) < p$) and we set $\mathbf{\Sigma}_v \to 0$ in (3.8), it is not obvious to have the (finite) maximum of the likelihood function. We take $\boldsymbol{\theta}$ $(= \boldsymbol{b})$ and apply the matrix formulae that for a positive definite $\mathbf{A}$ we have

$$|\mathbf{A} + \mathbf{b}\mathbf{b}^{'}| = |\mathbf{A}|[1 + \mathbf{b}^{'}\mathbf{A}^{-1}\mathbf{b}]$$

and

$$[\mathbf{A} + \mathbf{b}\mathbf{b}^{'}]^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{b}[1 + \mathbf{b}^{'}\mathbf{A}^{-1}\mathbf{b}]^{-1}\mathbf{b}^{'}\mathbf{A}^{-1}$$

for $\mathbf{A} = a_{kn}\mathbf{\Sigma}_v$ $(k = 1, \cdots, n)$, $\mathbf{\Sigma}_x = \mathbf{b}\mathbf{b}^{'}$, $\mathbf{b} = \sigma_\mu \mathbf{\Pi}$, $\sigma_\mu^2 = \mathcal{E}[(\Delta\boldsymbol{\mu}_i)^2]$, and $\mathbf{b}_* = \mathbf{\Sigma}_v^{-1}\mathbf{b}$.

Then L_n is proportional to (-1/2) times

$$\begin{aligned} L_{1n} &= \sum_{k=1}^{n}\left[\log|\mathbf{\Sigma}_v| + \log(a_{kn} + \mathbf{b}^{'}\mathbf{\Sigma}_v^{-1}\mathbf{b}) + a_{kn}^{-1}\mathbf{z}_k^{'}\mathbf{\Sigma}_v^{-1}\mathbf{z}_k - \frac{a_{kn}^{-1}(\mathbf{z}_k^{'}\mathbf{\Sigma}_v^{-1}\mathbf{b})^2}{a_{kn} + \mathbf{b}^{'}\mathbf{\Sigma}_v^{-1}\mathbf{b}}\right] \\ &= n\log|\mathbf{\Sigma}_v| + \sum_{k=1}^{n} a_{kn}^{-1}\mathbf{z}_k^{'}\mathbf{\Sigma}_v^{-1}\mathbf{z}_k + \sum_{k=1}^{n}\left[\log(a_{kn} + c) - \frac{a_{kn}^{-1}(\mathbf{z}_k^{'}\mathbf{b}_*)^2}{a_{kn} + c}\right], \end{aligned}$$

where we take $c = \mathbf{b}^{'}\mathbf{\Sigma}_v^{-1}\mathbf{b}$ as the normalization.

Because L_{1n} is a concave function of $\mathbf{\Sigma}_v^{-1}$ and the last term becomes

$$\sum_{k=1}^{n}\frac{(\mathbf{z}_k^{'}\mathbf{\Sigma}_v^{-1}\mathbf{b})^2}{a_{kn}^2 + a_{kn}\mathbf{b}^{'}\mathbf{\Sigma}_v^{-1}\mathbf{b}} = \mathbf{b}_*^{'}\sum_{k=1}^{n}[\frac{1}{a_{kn}(a_{kn} + c)}\mathbf{z}_k\mathbf{z}_k^{'}]\mathbf{b}_*$$

because $\mathbf{b}_* = \mathbf{\Sigma}_v^{-1}\mathbf{b}$. Then given $\mathbf{b}_*^{'}\mathbf{\Sigma}_v\mathbf{b}_* = c$, it is a quadratic form and its maximum (or the likelihood function is maximized) is the larger characteristic root. Given the initial condition $\mathbf{y}_0$, the unconditional maximum likelihood (ML) estimator can be defined as the solution of maximizing the log-likelihood function. In the general case when

$p > q \geq 1$, we have the next result.

Theorem 3.1 : Assume $\mathbf{\Sigma}_v$ is non-singular and rank$(\mathbf{\Sigma}_x) = p-q$ $(p > q \geq 1)$ in (2.2) and (3.1). Then there exists a unique ML estimator for $\mathbf{B}$.

As a consequence of this result, under the above conditions the structural parameter $\boldsymbol{\theta}$ (i.e. $\mathbf{B}$) is identified (up to a normalization). When we take the normalization $\mathbf{b}^{'}\mathbf{\Sigma}_v^{-1}\mathbf{b} = \mathbf{b}_*^{'}\mathbf{\Sigma}_v\mathbf{b}_* = c$ (*a constant*), the maximum likelihood estimator of $\mathbf{\Sigma}_v$ is given by

$$\hat{\mathbf{\Sigma}}_{v,ML} = \frac{1}{n}\sum_{k=1}^{n} a_{kn}^{-1}\mathbf{z}_k\mathbf{z}_k^{'} \ . \tag{3.11}$$

In the present setting

$$\mathcal{E}[\hat{\mathbf{\Sigma}}_{v,ML}] = \mathbf{\Sigma}_v + (\frac{1}{n}\sum_{k=1}^{n} a_{kn}^{-1})\mathbf{\Sigma}_x \tag{3.12}$$

and thus $\hat{\mathbf{\Sigma}}_{v,ML}$ cannot be a consistent estimator of $\mathbf{\Sigma}_v$. This is one of the consequence of the errors-in-variables models although it has been known in the standard errors-in-variables models. (See Anderson (1984) for the details.) Also one may think that as an estimator of $\mathbf{\Sigma}_x$ we may have

$$\mathbf{S}_n = \frac{1}{n}\sum_{k=1}^{n} \mathbf{z}_k\mathbf{z}_k^{'} \ . \tag{3.13}$$

Since

$$\mathcal{E}[\mathbf{S}_n] = \mathbf{\Sigma}_x + (\frac{1}{n}\sum_{k=1}^{n} a_{kn})\mathbf{\Sigma}_v \ , \tag{3.14}$$

then $\mathbf{S}_n$ cannot be a consistent estimator of $\mathbf{\Sigma}_x$.

It is straightforward to extend the above likelihood analysis to the cases for more general q $(q \leq p)$ and we have the corresponding results. Then it is not obvious to find a general way to construct the consistent estimator of $\mathbf{\Sigma}_x$ and $\mathbf{\Sigma}_v$ at the same time even if we do not have the seasonality component in the nonstationary errors-in-variable models.

4. Macro-SIML Estimation

Although we have considered the likelihood function in the errors-in-variables models under the Gaussianity, we need a simple robust procedure such that the assumptions of Gaussianity and the specifications of each components are not crucial for the estimating results.

We denote $a_{k_n,n}$ and we notice that $a_{k_n,n} \to 0$ as $n \to \infty$ when $k_n = O(n^{\alpha})$ $(0 < \alpha < 1)$ since $\sin x \sim x$ as $x \to 0$. When k_n is small, we expect that $a_{k_n,n}$ is small. Then the separating information maximum likelihood (SIML) estimator of $\hat{\mathbf{\Sigma}}_x$ is defined by

$$\hat{\mathbf{\Sigma}}_{x,SIML} = \frac{1}{m_n} \sum_{k=1}^{m_n} \mathbf{z}_k \mathbf{z}_k^{'} \ . \tag{4.1}$$

(We need to use a consistent estimator for $\mathbf{\Sigma}_v$.) For $\hat{\mathbf{\Sigma}}_x$, the number of terms m_n should be dependent on n. Then we need the order requirement that $m_n = O(n^{\alpha})$ and $0 < \alpha < 1$.

Asymptotic properties of SIML

For the estimation of the variance-covariance matrix $\mathbf{\Sigma}_x = (\sigma_{gh}^{(x)})$, we have the next result.

Theorem 4.1 : We assume (2.2) and (3.1) and also assume that

$\mathbf{w}_i^{(x)} = (w_{ji}^{(x)})$ and $\mathbf{v}_i = (v_{ji})$ are a sequence of independent random variables with $\mathcal{E}[w_{ig}^{(x)4}] < \infty$ and $\mathcal{E}[v_{ig}^4] < \infty$ $(i,j = 1,\cdots,n; g,h = 1,\cdots,p)$.

Then (i) For $m_n = n^\alpha$ and $0 < \alpha < 1$, as $n \longrightarrow \infty$

$$\hat{\mathbf{\Sigma}}_x - \mathbf{\Sigma}_x \xrightarrow{p} \mathbf{O} . \tag{4.2}$$

(ii) For $m_n = n^\alpha$ and $0 < \alpha < 0.8$, as $n \longrightarrow \infty$

$$\sqrt{m_n}\left[\hat{\sigma}_{gh}^{(x)} - \sigma_{gh}^{(x)}\right] \xrightarrow{\mathcal{L}} N\left(0, \sigma_{gg}^{(x)}\sigma_{hh}^{(x)} + \left[\sigma_{gh}^{(x)}\right]^2\right) . \tag{4.3}$$

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 $\sigma_{gk}^{(x)}\sigma_{hl}^{(x)} + \sigma_{gl}^{(x)}\sigma_{hk}^{(x)}$ $(g,h,k,l = 1,\cdots,p)$.

For estimating the variance-covariance matrix $\mathbf{\Sigma}_x = (\sigma_{gh}^{(x)})$,the number of terms m_n should be dependent on n because we need the resulting desirable asymptotic properties. Then we need the order requirement that $m_n = O(n^\alpha)$ $(0 < \alpha < 1)$. Because the properties of the SIML estimation method depend on the choice of m_n, which are dependent on n, we have investigated the asymptotic effects as well as the small sample effects of several possibilities. We can obtain an optimal choice of m_n.

Theorem 4.2 : In the setting of Theorem 4.1, an optimal choice of $m_n = n^\alpha$ $(0 < \alpha < 1)$ with respect to the asymptotic mean squared error when n is large is given by $\alpha^* = 0.8$.

It may be natural to use the sample quantities

$$\hat{\mathbf{\Sigma}}_x = (\frac{1}{m_n}\sum_{k=1}^{m_n} z_{ik}z_{jk}) \tag{4.4}$$

in order to make statistical inference on $\mathbf{\Sigma}_x$. The estimation of the Pearson-correlation coefficients among the trend variables is a typical case, which is given by

$$\hat{\rho}_{ij} = \frac{\sum_{k=1}^{m_n} z_{ik}z_{jk}}{\sqrt{\sum_{k=1}^{m_n} z_{ik}^2}\sqrt{\sum_{k=1}^{m_n} z_{jk}^2}} \ . \tag{4.5}$$

Furthermore, we consider the estimation of the structural relationships in the non-stationary time series process satisfying (3.1). Here we notice that the present statistical problem could be regarded as the estimation of structural relationships with the covariance matrix $\mathbf{\Sigma}_x(\boldsymbol{\theta})$ with $\boldsymbol{\theta}$ being the vector of parameters. In the standard statistical multivariate analysis, Anderson (1984, 2004) have discussed the statistical models of estimating structural relationships among a set of variables and we have n independent observations on the underlying variables.

We consider the estimation of the structural parameter vector $\boldsymbol{\beta}$ in the structural equation

$$\boldsymbol{\beta}^{'}\mathbf{y}_i = u_i \ , \tag{4.6}$$

where u_i $(= \boldsymbol{\beta}^{'}\mathbf{v}_i)$ in (3.1). It is a simle case when $q = 1$. By using the arguments on the likelihood function, it may be natural to consider the characteristic equation

$$\left[\hat{\mathbf{\Sigma}}_x - \lambda\mathbf{\Sigma}_v\right]\hat{\boldsymbol{\beta}} = \mathbf{0} \ . \tag{4.7}$$

where $\hat{\boldsymbol{\Sigma}}_x$ is given by (4.1) and λ is the (scalar) characteristic root. Here we need to use a consistent estimator for $\boldsymbol{\Sigma}_v$. When we take the smallest eigenvalue λ_1 in (4.7) and we take $\hat{\boldsymbol{\Sigma}}_{v,SIML}$, we have the $\hat{\boldsymbol{\beta}}_{SIML}$, which is called the SIML estimator of $\boldsymbol{\beta}$.

Theorem 4.3 : We assume (3.1), (2.2), (4.6) and rank$(\boldsymbol{\Sigma}_x) = p-1$. Let $\hat{\boldsymbol{\beta}}$ be the characteristic vector with the corresponding minimum characteristic root of (4.7), which is the SIML estimator of $\boldsymbol{\beta}$. We assume that $\mathbf{w}_i^{(x)} = (w_{ji}^{(x)})$ and $\mathbf{v}_i = (v_{ji})$ $(i = 1, \cdots, n; j = 1, \cdots, p)$ are a sequence of independent random variables with $\mathcal{E}[w_{ig}^{(x)4}] < \infty$ and $\mathcal{E}[v_{ig}^4] < \infty$ $(i, j = 1, \cdots, n; g, h = 1, \cdots, p)$. We further assume that we have a consistent estimator $\hat{\boldsymbol{\Sigma}}_v = \boldsymbol{\Sigma}_v + O_p(m_n^{-1/2})$.
Then (i) For $m_n = n^\alpha$ and $0 < \alpha < 1$, as $n \longrightarrow \infty$

(4.8) $$\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \stackrel{p}{\longrightarrow} \mathbf{0}\ .$$

(ii) For $m_n = n^\alpha$ and $0 < \alpha < 0.8$, as $n \longrightarrow \infty$

(4.9) $$\sqrt{m_n}(\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}_2) \stackrel{\mathcal{L}}{\longrightarrow} N(\mathbf{0}, \boldsymbol{\Sigma}_{22}^{(x)-1}\mathcal{E}[\mathbf{S}_2\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{S}_2']\boldsymbol{\Sigma}_{22}^{(x)-1})\ ,$$

where $\hat{\boldsymbol{\beta}} = (1, -\hat{\boldsymbol{\beta}}_2)$, $\mathbf{S}_2 = (\mathbf{0}, \mathbf{I}_{p-1})\mathbf{S}$, $\mathbf{S}$ is the limiting (normal) random matrix of $\sqrt{m_n}[\hat{\boldsymbol{\Sigma}}_x - \boldsymbol{\Sigma}_x]$ and $\boldsymbol{\Sigma}_{22}^{(x)} = (\mathbf{0}, \mathbf{I}_{p-1})\boldsymbol{\Sigma}_x(\mathbf{0}, \mathbf{I}_{p-1})'$ is the $(p-1) \times (p-1)$ right-lower-corner of the variance-covariance matrix $\boldsymbol{\Sigma}_x$. (We have assumed that rank$(\boldsymbol{\Sigma}_x) = p-1$ and then we take $|\boldsymbol{\Sigma}_{22}^{(x)}| \neq 0$.)

Also under a set of regularity conditions we have that the smallest eigenvalue λ_1 of (4.7),

$$\lambda_1 \longrightarrow 0 \text{ (in probability)} \tag{4.10}$$

as $n \to \infty$. Then we define the SILS (Separating Information Least Squares) method by solving

$$\hat{\mathbf{\Sigma}}_x \hat{\boldsymbol{\beta}}_{SILS} = \mathbf{0}\ . \tag{4.11}$$

When $p = 2, q = 1$, $\boldsymbol{\beta} = (1, -\beta_2)^{'}$, $\hat{\boldsymbol{\beta}}_{*,SIML} = (1, -\hat{\beta}_2)^{'}$, and $\mathbf{\Pi} = (\beta_2, 1)^{'}$ ($\mathbf{\Sigma}_v = \sigma_v^2 \mathbf{I}_2$) in Example 1, then the SILS estimation becomes

$$\hat{\beta}_2 = \frac{\sum_{k=1}^{m_n} z_{1k} z_{2k}}{\sum_{k=1}^{m_n} z_{2k}^2}\ , \tag{4.12}$$

which is the regression coefficient of the first transformed variable on the second transformed variable in $\mathbf{z}_k$ $(= (z_{1k}, z_{2k})^{'})$ $(k = 1, \cdots, m_n)$.

Some Remarks

We notice that the ML estimator of $\mathbf{\Sigma}_v$ is not consistent because $a_{k,n} = O(n^{-1})$ for a fixed k when n is large. It is a consequence of the errors-in-variables models and the problem of incidental parameters. In order to construct a consistent estimator we use the fact that for any positive integer l_n such that $l_n \to \infty, l_n/n \to 0$ $(n \to \infty)$, $a_{n+1-l_n,n} \to 4$ $(n \to \infty)$ and

$$\frac{1}{l_n} \sum_{k=n+1-l_n}^{n} a_{kn}^{-1} \mathbf{z}_k \mathbf{z}_k^{'} \xrightarrow{p} \frac{1}{4} \mathbf{\Sigma}_x + \mathbf{\Sigma}_v\ . \tag{4.13}$$

Then we can construct a consistent estimator of $\mathbf{\Sigma}_v$ as

$$\hat{\Sigma}_v = \frac{1}{l_n} \sum_{k=n+1-l_n}^{n} a_{kn}^{-1} \mathbf{z}_k \mathbf{z}_k^{'} - \frac{1}{4} \mathbf{\Sigma}_x\ . \tag{4.14}$$

Although we have developed the SIML estimation when $q = 1$, it is straightforward to extend the SIML procedure when we have several structural relationships among trend variables at the same time. The SIML estimation can be defined by the smaller q $(\leq p)$ roots and the corresponding q $(\leq p)$ vectors of the characteristic equation. It may correspond to the standard situation in the statistical multivariate analysis except the fact that the classical multivariate analysis was based on the case when the observations are realizations of independent random variables without seasonality as well as non-stationarity in time series data sets.

5. Discussions on seasonality

We return to the original setting with seasonality in Section 2 and consider the case when we have

$$\mathbf{y}_i = \mathbf{x}_i + \mathbf{s}_i + \mathbf{v}_i \ , \tag{5.1}$$

where $\mathbf{x}_i$ are a sequence of trend components and $\mathbf{s}_i$ are a sequence of seasonal components. When we transform the observed data by using the difference operator $\Delta = 1 - \mathcal{L}$ ($\mathcal{L}\mathbf{y}_i = \mathbf{y}_{i-1}$) and $\mathbf{P}_n$, we have the next result, which is a direct extension of Theorem 4.1.

Theorem 5.1 : We assume (5.1), (2.2) and (2.4) and also assume that $\mathbf{w}_i^{(x)} = (w_{ji}^{(x)}), \mathbf{w}_i^{(s)} = (w_{ji}^{(s)})$ and $\mathbf{v}_i = (v_{ji})$ are a sequence of independent random variables with $\mathcal{E}[w_{ig}^{(x)4}] < \infty, \mathcal{E}[w_{ig}^{(s)4}] < \infty$ and $\mathcal{E}[v_{ig}^4] < \infty$ $(i, j = 1, \cdots, n; g, h = 1, \cdots, p)$. Let $\hat{\mathbf{\Sigma}}_x$ be given by (4.1). Then (i) For $m_n = n^{\alpha}$ and $0 < \alpha < 1$, as $n \longrightarrow \infty$

$$\hat{\mathbf{\Sigma}}_x - \mathbf{\Sigma}_x \xrightarrow{p} \mathbf{O} \ . \tag{5.2}$$

(ii) For $m_n = n^\alpha$ and $0 < \alpha < 0.8$, as $n \longrightarrow \infty$

$$\sqrt{m_n}\left[\hat{\sigma}_{gh}^{(x)} - \sigma_{gh}^{(x)}\right] \xrightarrow{\mathcal{L}} N\left(0, \sigma_{gg}^{(x)}\sigma_{hh}^{(x)} + \left[\sigma_{gh}^{(x)}\right]^2\right) . \tag{5.3}$$

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 $\sigma_{gk}^{(x)}\sigma_{hl}^{(x)} + \sigma_{gl}^{(x)}\sigma_{hk}^{(x)}$ $(g, h, k, l = 1, \cdots, p)$.

Alternatively, it has been a common practice to use the seasonal difference of original time series since Box and Jenkins (1970) if we observe clear seasonal fluctuations. When we transform the observed data by using the seasonal difference operator $\Delta_s = 1 - \mathcal{L}^s$ ($\mathcal{L}^s \mathbf{y}_i = \mathbf{y}_{i-s}$) and $\mathbf{P}_n$, we have

$$\Delta_s \mathbf{y}_i = (1 + \mathcal{L} + \cdots + \mathcal{L}^{s-1})\Delta \mathbf{x}_i + (1 - \mathcal{L}^s)\mathbf{s}_i + (1 - \mathcal{L}^s)\mathbf{v}_i . \tag{5.4}$$

Then there can be alternative possibilities of transformation of $\mathbf{Y}_n$, but we may use $\mathbf{Z}_n^{(s)}$ $(= (\mathbf{z}_k^{(s)'}))$ by

$$\mathbf{Z}_n^{(s)} = \mathbf{P}_n \mathbf{C}_n^{(s)-1}\left(\mathbf{Y}_n - \bar{\mathbf{Y}}_0\right) , \tag{5.5}$$

where $\mathbf{C}_n^{(s)-1} = \mathbf{C}_N^{-1} \otimes \mathbf{I}_s$ and

$$\mathbf{C}_N^{-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}_{N \times N} , \tag{5.6}$$

where we have assumed that N, s and $n = Ns$ are positive integers. Then the analysis of the likelihood function in Section 3 can be extended to the above transformation with seasonality. We take the transformation matrix

$$\mathbf{B}_n^{(1)} = (b_{jk}^{(1)}) = \mathbf{P}_n \mathbf{C}_n^{(s)-1} \ , \tag{5.7}$$

Lemma A-1 in Appendix gives

$$\sum_{j=1}^{n} b_{kj}^{(1)} b_{k^{'},j}^{(1)} = \delta(k,k^{'}) 4 \sin^2 \left[\frac{\pi}{2} \frac{2k-1}{2n+1} s \right] + O(\frac{1}{n}) \ . \tag{5.8}$$

For any finite integer s (the seasonal lag), we take $k = k_n, k^{'} = k_n^{'}$ depending on n. If we take $k_n = [\frac{2n}{s}] + l_n$ (or $k_n^{'} = [\frac{2n}{s}] + l_n^{'}$) and $l_n, l_n^{'}$ being integers, (5.8) is $o(1)$ when $l_n/n \to 0$.

For the estimation of the seasonal covariance matrix $\mathbf{\Sigma}_s = (\sigma_{gh}^{(s)})$ and $\hat{\mathbf{\Sigma}}_s = (\hat{\sigma}_{gh}^{(s)})$, instead of (4.1) we use

$$\hat{\mathbf{\Sigma}}_{s,SIML} = \frac{1}{m_n} \sum_{k \in I_n^{(s)}} \mathbf{z}_k^{(s)} \mathbf{z}_k^{(s)'} \ , \tag{5.9}$$

where s is the seasonal integer, $[x]$ is the largest integer being equal or less than x and $I_n^{(s)}$ is the set of integers such that $I_n^{(s)} = \{[2n/s] + 1, \cdots, [2n/s] + m_n]\}$ with $m_n = n^{\alpha}$ $(0 < \alpha < 1)$.
Alternatively, $I_n^{(s)}$ can be replaced by a symmetric region $I_n^{(s)} = \{[2n/s] - m_n/2, \cdots, [2n/s], \cdots, [2n/s] + m_n/2]\}$.
In this formulation $[2n/s]$ corresponds to the seasonal frequency in the frequency domain of the observed time series. For the quarterly and monthly data, we take $s = 4$ and $s = 12$, respectively. Then we have the next result.

Theorem 5.2 : We assume (5.1), (2.2) and (2.3) and also assume that $\mathbf{w}_i^{(x)} = (w_{ji}^{(x)})$, $\mathbf{w}_i^{(s)} = (w_{ji}^{(s)})$ and $\mathbf{v}_i = (v_{ji})$ are a sequence of independent random variables with $\mathcal{E}[w_{ig}^{(x)4}] < \infty$, $\mathcal{E}[w_{ig}^{(s)4}] < \infty$ and $\mathcal{E}[v_{ig}^{4}] < \infty$ $(i, j = 1, \cdots, n; g, h = 1, \cdots, p)$. Let $\hat{\mathbf{\Sigma}}_s$ be given by (5.9). Then (i) for $m_n = n^{\alpha}$ and $0 < \alpha < 1$, as $n \longrightarrow \infty$

$$\hat{\mathbf{\Sigma}}_s - \mathbf{\Sigma}_s \xrightarrow{p} \mathbf{O} \ . \tag{5.10}$$

(ii) For $m_n = n^{\alpha}$ and $0 < \alpha < 0.8$, as $n \longrightarrow \infty$

$$\sqrt{m_n}\left[\hat{\sigma}_{gh}^{(s)} - \sigma_{gh}^{(s)}\right] \xrightarrow{\mathcal{L}} N\left(0, \sigma_{gg}^{(s)}\sigma_{hh}^{(s)} + \left[\sigma_{gh}^{(s)}\right]^2\right) \ . \tag{5.11}$$

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

When we use (2.4) instead of (2.3) with (5.1) and (2.2) in Theorem 5.2, it is possible to obtain the similar results and

$$\hat{\mathbf{\Sigma}}_s - \mathbf{\Sigma}_s \xrightarrow{p} \mathbf{O} \ , \tag{5.12}$$

where $\hat{\mathbf{\Sigma}}_s$ is given by (5.9).

When we use (4.1) for the seasonally transformed data $\Delta_s \mathbf{y}_i$ $(i = 1, \cdots, n)$ in Theorem 5.2, however, its probability limit is given by

$$\hat{\mathbf{\Sigma}}_x \xrightarrow{p} s\mathbf{\Sigma}_x + \mathbf{\Sigma}_s \tag{5.13}$$

because the transformed trend component is given by

$$\Delta_s \mathbf{x}_i = (1 + \mathcal{L} + \cdots + \mathcal{L}^{s-1})\mathbf{w}_i^{(x)} \ . \tag{5.14}$$

The bias can be significant when $s > 1$, but there is a way to construct a consistent estimator.

Furthermore, in these cases it is possible to obtain the asymptotic distributions of the estimators for trends, seasonals and other quantities including the correlation coefficients.

6. Simulations and an Empirical Example

We have done several simulations. The data length is 80, the number of simulations is 3000, $\alpha = 0.6$, and $m_n = n^\alpha$ in each case. We have set three cases with the nonstationary trend with seasonality, whose typical simulation paths are given in Appendix. We have done a number of simulations including the traditional linear seasonal models. and we will report some results which nay be reasonable description of economic quarterly data ($s = 4$). Since we deal with the nonstationary seasonality, we need to control the parameter values carefully including the initial conditions. Figure-1 does not have any seasonality while Figure-2 and Figure-3 have non-linear seasonality and they are rather extreme cases in our simulations. In Simulations 1-3 we first generated the initial uniform random variables $s_{j,-3}, \cdots, s_{j,0}$, the sequence of i.i.d. random variable $sv_{j,i}$ for $j = 1, 2; i = 0, \cdots, n$. Then we set $\mathbf{s}_i = (s_{1i}, s_{2i})'$ such that $sw_{j,i} = sw_{j,j-1} + sv_{j,i}$ and $s_{j,i} = s_{j,0} \times sw_{j,i}$. We have summarized the four simulation results in Tables 6.1-6.4. In our tables *cor* $= 0.9$ means the true correlation coefficient among trend components and *cor* is the SIML estimate. (vol-1 is the correlation estimate based on the first differenced data and vol-4 is the correlation estimate based on the seasonal differenced data with $s = 4$.)

When we have the basic model with the trend and noise components without the seasonal and cycle components, the optimal choice of $m_n = n^\alpha$ would be $\alpha = 0.8$, but it seems that the choice of $\alpha = 0.6$ would be appropriate for the robustness of the results when we have extreme seasonality as well as non-stationary trends.

We have investigated the estimation of the correlation coefficient of the seasonal components and given Table 6-4 when the seasonals were generated by $\mathbf{s}_i = (s_{1i}, s_{2i})'$ and $\mathbf{w}_i^{(s)} = (w_{1i}^{(s)}, w_{2i}^{(s)})'$ such that $s_{ji} = -s_{j,i-1} - s_{j,i-2} - s_{j,i-3} + w_{ji}^{(s)}$ $(i = 1, \cdots, n; j = 1, 2)$ given the initial random variables, and we also have trend components and noise components (Simulation 4). The number of data is 400 and we took α and we have given a typical sample path as Figure-4 in Appendix.

We have found that even with the extreme cases given in our figures the macro-SIMLE method gives reasonable estimates while in more standard cases we have more favorable results for the use of the SIML estimation.

Table 6-1 : Simulation-1

($n = 80, \alpha = 0.6$, nsim=3000)

cor= 0.9	corr	vol-4	vol-1
mean	0.852	0.733	0.491
SD	0.088	0.076	0.095
cor= 0.0	corr	vol-4	vol-1
mean	0.007	0.003	0.001
SD	0.278	0.168	0.119

Table 6-2 : Simulation-2

($n = 80, \alpha = 0.6$, nsim=3000)

cor= 0.9	corr	vol-4	vol-1
mean	0.805	0.663	0.133
SD	0.118	0.088	0.295
cor= 0.0	corr	vol-4	vol-1
mean	-0.007	2.59E-03	0.005
SD	0.278	1.62E-01	0.287

Table 6-3 : Simulation-3

$(n = 80, \alpha = 0.6$, nsim=3000)

cor= 0.9	corr	vol-4	vol-1
mean	0.672	0.344	0.034
SD	0.196	0.185	0.191
cor= 0.0	corr	vol-4	vol-1
mean	0.002	0.002	0.002
SD	0.284	0.149	0.184

Table 6-4 : Simulation-4

$(n = 400, \alpha = 0.40$, nsim=1000)

cor= 0.8	corr	vol-4	vol-1
mean	0.759	0.550	0.753
SD	0.114	0.068	0.256
cor= 0.0	corr	vol-4	vol-1

Finally, we report an empirical estimate of the Japanese (real) GDP and fix-investment given as Figure 1-1 as a typical example. We have used the quarterly data which were taken from the official estimates from the Japanese Cabinet Office. When we take the first differences and the estimate of the correlation coefficient of the GDP-trend and investment-trend is 0.726176 while we take the seasonal difference and the estimate of the correlation coefficient of the GDP-trend and investment-trend is -0.12159. On the other hand, the SIML estimate of the correlation coefficient of the GDP-trend and investment-trend is 0.614224 (0.069623) while the SIML estimate of the correlation coefficient of the GDP-seasonal and investment-seasonal is 0.169324 (0.108598). We have used the symmetric region $I_n^*(s)$ and the parenthesis is the estimate of standard deviation calculated by the standard asymptotic formula in statistical multivariate analysis $(1-\hat{\rho}^2)/\sqrt{m_n}$. These estimates give some information on the statistical relationship between quarterly GDP and quarterly fixed-investment in Japan.

7. Concluding Remarks

In this paper, we have proposed to use a new statistical method for estimating the statistical relationships in the non-stationary time series with trend, seasonality and noise. Instead of using the seasonally adjusted data published by the official statistics agencies, we propose to use the Separating Information Maximum Likelihood (SIML) estimator, which can be regarded as a modification of the classical Maximum Likelihood (ML) method in some sense. We have shown that the SIML estimator has reasonable asymptotic properties; it is consistent and it has the asymptotic normality when the sample size is large under reasonable conditions. The SIML estimator has reasonable finite sample properties and also it has the asymptotic robustness properties. Based on simulations the SIML estimator is so simple that it can be practically used for the multivariate non-stationary time series.

We also have suggested a number of possible applications in macroeconomic non-stationary time series since many important macro time series exhibit clear seasonality.

There are several possible extensions and directions. First, it may be natural to incorporate the stationary cycle components in the time series decompositions. Second, it may be straightforward to extend the cases when we have double unit roots in the trend variables. Third, we have done several data analysis of quarterly macro time series and have found that the SIML approach gives useful information, but is obvious to do more.

Finally, there are several important issues remained in the present work. It may be reasonable to have the cycle component to the time series decomposition in (2.1). Also there can be some extensions to the dynamic panel data analysis, which has many recent applications in econometrics. The results of our investigations on these issues will be reported in another occasion.

References

[1] Akaike, H. (1980), "Seasonal adjustment by a Bayesian modeling," *Journal of Time Series Analysis*, 1, 1-12.

[2] Anderson, T.W. (1984), "Estimating Linear Statistical Relationships," *Annals of Statistics*, 12, 1-45.

[3] Anderson, T.W. (2003), *An Introduction to Statistical Multivariate Analysis*, 3rd Edition, John-Wiley.

[4] Box, G.E.P. and G. Jenkins (1970), *Time Series Analysis : Forecasting and Control*, Holden-Day, San Francisco.

[5] Engle, R. (1974), "Band Spectrum Regression," *International Economic Review*, Vol.15, 1-11.

[6] Engle, R. and C.W.J. Granger (1987), "Co-integration and Error Correction," *Econometrica*, Vol.55, 251-276.

[7] Findley, D., B.C.Monsell, W.R. Bell, M.C. Otto and B.C. Chen (1998), "New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program," *Journal of Business and Economic Statistics*, 16, 127-176(with Discussions).

[8] Fuller, W. (1987), *Measurement Error Models*, John-Wiley.

[9] Johansen, S. (1995), *Likelihood Based Inference in Cointegrated Vector Autoregressive Models*, Oxford UP.

[10] Kitagawa, G. (2010), *Introduction to Time Series Modeling*, CRA Press.

[11] Kunitomo, N. and S. Sato (2008, 2013), "Separating Information Maximum Likelihood Estimation of Realized Volatility and Covariance with Micro-Market Noise," Discussion Paper CIRJE-F-581, Graduate School of Economics, University of Tokyo, (http://www.e.u-tokyo.ac.jp/cirje/research/dp/2008),

North American Journal of Economics and Finance, Elsevier, 26, 282-309.

[12] Kunitomo, N. and S. Sato (2011), "The SIML Estimation of Realized Volatility of Nikkei-225 Futures and Hedging Coefficient with Micro-Market Noise," *Mathematics and Computers in Simulation*, Vol.81, 1272-1289.

[13] Phillips, P. C. B. (1991), "Spectral regression fo cointegrated time series," In *Nonparametric and Semiparametric Methods in Econometrics and Statistics,Cambridge University Press*, 413-435.

APPENDIX : Mathematical Derivations of Theorems

In this Appendix, we give some details of the proofs in Sections 3 and 4. However, some of the proofs are are based on the modifications of the results by Kunitomo and Sato (2008, 2011, 2013) and they are similar to them, but there are some differences.

Proof of Theorem 3.1 : Because $|\mathbf{\Sigma}_v| \neq 0$ and $\text{rank}(\mathbf{\Sigma}_x) = p - q > 0$ $(p > q \geq 1)$, there exists a $p \times q$ matrix $\mathbf{B}$ such that B $= q$. We take $\mathbf{B}_* = \mathbf{\Sigma}_v^{-1}\mathbf{B}$ and a constant matrix $\mathbf{C}$ such that

(A.1) $$\mathbf{B}_*^{'}\mathbf{\Sigma}_v\mathbf{B}_* = \mathbf{C}\ .$$

The likelihood function is proportional to (-1/2) times

$$L_{1n} = n\log|\mathbf{\Sigma}_v| + \sum_{k=1}^{n}|a_{kn}\mathbf{I}_{p-q} + \mathbf{B}_*^{'}\mathbf{\Sigma}_v\mathbf{B}_*| + \sum_{k=1}^{n} a_{kn}^{-1}\mathbf{z}_k^{'}\mathbf{\Sigma}_v^{-1}\mathbf{z}_k$$

(A.2) $$- \sum_{k=1}^{n} a_{kn}^{-1}\mathbf{z}_k^{'}\mathbf{B}_*[a_{kn}\mathbf{I}_{p-q} + \mathbf{B}_*^{'}\mathbf{\Sigma}_v\mathbf{B}_*]^{-1}\mathbf{B}_*^{'}\mathbf{z}_k\ .$$

Then given the normalization (A.1) we can attain the maximum of the likelihood function when $\mathbf{\Sigma}_v = (1/n)\sum_{k=1}^{n} a_{kn}^{-1}\mathbf{z}_k\mathbf{z}_k^{'}$ and we take the larger $p - q$ characteristic roots of the associated characteristic equation because the last term of (A.2) can be rewritten as

(A.3) $$\text{tr}\mathbf{B}_*^{'}(\sum_{k=1}^{n} a_{kn}^{-1}\mathbf{z}_k\mathbf{z}_k^{'})(a_{kn}\mathbf{I}_{p-q} + \mathbf{C})^{-1}\mathbf{B}_*\ .$$

(**Q.E.D.**)

Proof of Theorem 4.1 :

(**Step 1**) : Let $\mathbf{z}_k^{(x)} = (z_{kj}^{(x)})$ and $Z_k^{(v)} = (z_{kj}^{(v)})$ $(k = 1, \cdots, n)$ be the

k-th vector elements of $n \times p$ matrices

(A.4) $$\mathbf{Z}_n^{(x)} = \mathbf{P}_n \mathbf{C}_n^{-1}(\mathbf{X}_n - \bar{\mathbf{Y}}_0)\ ,\ \mathbf{Z}_n^{(v)} = \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{V}_n,$$

respectively, where we denote $\mathbf{X}_n = (\mathbf{x}_k^{'}) = (x_{kg})$, $\mathbf{V}_n = (\mathbf{v}_k^{'}) = (v_{kg})$, $\mathbf{Z}_n = (\mathbf{z}_k^{'})$ $(= (z_{kg}))$ are $n \times p$ matrices with $z_{kg} = z_{kg}^{(x)} + z_{kg}^{(v)}$. We write z_{kg} as the $g-$th component of $\mathbf{z}_k$ $(k = 1, \cdots, n; g = 1, \cdots, p)$. By following the proof developed by Kunitomo and Sato (2013) for the case of fixed n, we use the decomposition of $z_{kg}^{(f)}$ $(f = x, v)$ for investigating the asymptotic distribution of $\sqrt{m_n}[\hat{\Sigma}_x - \Sigma_x] = (\sqrt{m_n}(\hat{\sigma}_{gh}^{(x)} - \sigma_{gh}^{(x)})_{gh})$ for $g, h = 1, \cdots, p$. We use the decomposition

(A.5) $$\begin{aligned} &\sqrt{m_n}\left[\hat{\Sigma}_x - \Sigma_x\right] \\ = \ &\sqrt{m_n}\left[\frac{1}{m_n}\sum_{k=1}^{m_n} \mathbf{z}_k \mathbf{z}_k^{'} - \Sigma_x\right] \\ = \ &\sqrt{m_n}\left[\frac{1}{m_n}\sum_{k=1}^{m_n} \mathbf{z}_k^{(x)} \mathbf{z}_k^{(x)'} - \Sigma_x\right] + \frac{1}{\sqrt{m_n}}\sum_{k=1}^{m_n} \mathcal{E}[\mathbf{z}_k^{(v)} \mathbf{z}_k^{(v)'}] \\ &+ \frac{1}{\sqrt{m_n}}\sum_{k=1}^{m_n}\left[\mathbf{z}_k^{(v)} \mathbf{z}_k^{(v)'} - \mathcal{E}[\mathbf{z}_k^{(v)} \mathbf{z}_k^{(v)'}]\right] + \frac{1}{\sqrt{m_n}}\sum_{k=1}^{m_n}\left[\mathbf{z}_k^{(x)} \mathbf{z}_k^{(v)'} + \mathbf{z}_k^{(v)} \mathbf{z}_k^{(x)'}\right]\ . \end{aligned}$$

Then we can investigate the conditions that three terms except the first one of (A.2) are $o_p(1)$. When these conditions are satisfied, we could estimate the variance and covariance of the underlying processes consistently as if there were no noise terms because other terms can be ignored asymptotically as $n \to \infty$.

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

(A.6) $$(\mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{C}_n^{'-1} \mathbf{P}_n^{'})_{k,k'} = \delta(k, k^{'}) 4 \sin^2[\frac{\pi}{2n+1}(k - \frac{1}{2})]\ .$$

Then because we have $\sum_{j=1}^{n} b_{kj} b_{k'j} = \delta(k,k') a_{kn}$ and $\mathbf{\Sigma}_v$ is bounded, it is straightforward to find K_1 (a constant) such that

$$\text{(A.7)} \qquad \mathcal{E}[(z_{kg}^{(v)})]^2 = \mathcal{E}[\sum_{i=1}^{n} b_{ki} v_{ig} \sum_{j=1}^{n} b_{kj} v_{jg}] \leq K_1 \times a_{kn} \ .$$

Also Kunitomo and Sato (2013) have shown that

$$\text{(A.8)} \qquad \frac{1}{m_n} \sum_{k=1}^{m_n} a_{kn} = \frac{1}{m_n} 2 \sum_{k=1}^{m_n} \left[1 - \cos(\pi \frac{2k-1}{2n+1}) \right] = O(\frac{m_n^2}{n^2})$$

and the second term of (A.5) becomes

$$\text{(A.9)} \qquad \frac{1}{\sqrt{m_n}} \sum_{k=1}^{m_n} \mathcal{E}[z_{kn}^{(v)}]^2 \leq K_1 \frac{1}{\sqrt{m_n}} \sum_{k=1}^{m_n} a_{kn} = O(\frac{m_n^{5/2}}{n^2}) \ ,$$

which is $o(1)$ if we set α such that $0 < \alpha < 0.8$. For the fourth term of (A.2),

$$\begin{aligned} \mathcal{E} \left[\frac{1}{\sqrt{m_n}} \sum_{j=1}^{m_n} z_{kg}^{(x)} z_{kg}^{(v)} \right]^2 &= \frac{1}{m_n} \sum_{k,k'=1}^{m_n} \mathcal{E} \left[z_{kg}^{(x)} z_{k',g}^{(x)} z_{kg}^{(v)} z_{k',g}^{(v)} \right] \\ &= O(\frac{m_n^2}{n^2}) \ , \end{aligned}$$

where

$$s_{jk} = \cos[\frac{2\pi}{2n+1}(j-\frac{1}{2})(k-\frac{1}{2})]$$

for $j,k = 1,2,\cdots,n$. (See Lemma 1.3 of Kunitomo and Sato (2008a).) In the above evaluation we have used the relation

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

For the third term of (A.5), we need to consider the variance of

$$(z_{kg}^{(v)})^2 - \mathcal{E}[(z_{kg}^{(v)})^2] = \sum_{j,j'=1}^{n} b_{kj} b_{k,j'} \left[v_{jg} v_{j',g} - \mathcal{E}(v_{jg} v_{j',g}) \right] \ .$$

Then by using the assumption on the existence of the fourth order moments, we can find a positive constant K_2 such that

$$\begin{aligned}
&\mathcal{E}\left[\frac{1}{\sqrt{m_n}}\sum_{k=1}^{m_n}((z_{kg}^{(v)})^2-\mathcal{E}[(z_{kg}^{(v)})^2])\right]^2 \\
=\ & \frac{1}{m_n}\sum_{k_1,k_2=1}^{m_n}\mathcal{E}\left[\sum_{j_1,j_2,j_3,j_4=1}^{n} b_{k_1,j_1}b_{k_1,j_2}(v_{j_1,g}v_{j_2,g}-\mathcal{E}(v_{j_1,g}v_{j_2,g}))\right. \\
& \qquad \times b_{k_3,j_3}b_{k_4,j_4}(v_{j_3,g}v_{j_4,g}-\mathcal{E}(v_{j_3,g}v_{j_4,g}))] \\
\leq\ & K_2\frac{1}{m_n}[\sum_{k=1}^{m_n}a_{kn}]^2 \\
=\ & O(\frac{1}{m_n}\times(\frac{m_n^3}{n^2})^2)\ ,
\end{aligned}$$

which is $O(m_n^5/n^4)$. Thus the third term of (A.5) is negligible if we set α such that $0<\alpha<0.8$.

(**Step 2**) The second step is to give the asymptotic variance of the first term of (A.5), that is,

$$\text{(A.10)} \qquad \sqrt{m_n}\left[\frac{1}{m_n}\sum_{k=1}^{m_n}\mathbf{z}_k^{(x)}\mathbf{z}_k^{(x)'}-\mathbf{\Sigma}_x\right]$$

because it is of the order $O_p(1)$. We can write

$$\begin{aligned}
&\frac{1}{m_n}\sum_{k=1}^{m_n}\mathbf{z}_k^{(x)}\mathbf{z}_k^{(x)'} \\
=\ & \frac{1}{m_n}(\frac{2}{n+\frac{1}{2}})\sum_{k=1}^{m_n}[\sum_{i=1}^{n}\mathbf{r}_i\cos[\pi(\frac{2k-1}{2n+1})(i-\frac{1}{2})]\sum_{j=1}^{n}\mathbf{r}_j^{'}\cos[\pi(\frac{2k-1}{2n+1})(j-\frac{1}{2})]] \\
=\ & \sum_{i=1}^{n}c_{ii}\mathbf{r}_i\mathbf{r}_i^{'}+\sum_{i\neq j}c_{ij}\mathbf{r}_i\mathbf{r}_j\ ,
\end{aligned}$$

where $\mathbf{r}_i=\mathbf{x}_i-\mathbf{x}_{i-1}$ and

$$c_{ii} = (\frac{2n}{2n+1})\left[1+\frac{1}{m}\frac{\sin 2\pi m(\frac{i-1/2}{2n+1})}{\sin(\pi\frac{i-1/2}{2n+1})}\right]\ ,$$

$$c_{ij} \;=\; \frac{1}{2m}(\frac{2n}{2n+1})\left[\frac{\sin 2\pi m(\frac{i+j-1}{2n+1})}{\sin(\pi\frac{i+j-1}{2n+1})}+\frac{\sin 2\pi m(\frac{j-i}{2n+1})}{\sin(\pi\frac{j-i}{2n+1})}\right] \quad (i\neq j)\ .$$

Then Kunitomo and Sato (2008, 2011) have shown that

(A.11) $$\frac{\sqrt{m_n}}{n}\sum_{i=1}^{n}\left[\mathbf{r}_i\mathbf{r}_i^{'}-\mathbf{\Sigma}_x+(c_{ii}-1)\mathbf{r}_i\mathbf{r}_i^{'}\right]=o_p(1)$$

and by re-writing (A-7) as

(A.12) $$\frac{\sqrt{m_n}}{n}\sum_{i=1}^{n}\left[c_{ii}\ \mathbf{r}_i\mathbf{r}_i^{'}-\mathbf{\Sigma}_x\right]+\frac{\sqrt{m_n}}{n}\sum_{i\neq j}^{n}\left[c_{ij}\ \mathbf{r}_i\mathbf{r}_j^{'}\right]$$

we need to evaluate the asymptotic variance of its second term. Kunitomo and Sato (2008, 2011) have also shown that the variance of the limiting distribution of the (g,g)-the element of (A.10) is the limit of

(A.13) $$V_n(g,g)=2\sum_{i,j=1}^{n}\frac{m_n}{n}c_{ij}^2[\sigma_{gg}^{(x)}]^2\ ,$$

The resulting arguments of the derivations are the result of straightforward calculations and lengthy, but the final form becomes simple. Because of Lemma 3 of Kunitomo and Sato (2013) as

(A.14) $$\sum_{i,j=1}^{n}c_{ij}^2=\frac{4}{m_n}\left[\frac{n}{2}+\frac{1}{4}\right]^2\ ,$$

we have that as $n\rightarrow\infty$

(A.15) $$V_n(g,g)\longrightarrow V(g,g)=2\left[\sigma_{gg}^{(x)}\right]^2\ .$$

(**Step 3**) Finally, we need to give the proof of the asymptotic normality. Define the sequence of $\sigma-$fields $\mathcal{F}_{n,i}$ generated by the set of random variables $\{\mathbf{x}_j,\mathbf{v}_j;1\leq j\leq i\leq n\}$, As the proof of Theorem

3 of Kunitomo and Sato (2008), for $(g,g)-$the element we shall use a sequence of random variables

$$\text{(A.16)} \qquad U_n(g,g) = \sum_{j=2}^{n}[2\sum_{i=1}^{j-1}\sqrt{m_n}c_{ij}\frac{r_{gi}}{\sqrt{n}}]\frac{r_{gj}}{\sqrt{n}} \ ,$$

which is a discrete martingale. Since the log-returns $r_{gi} = x_{gi} - x_{g,i-1}$ $(g = 1,\cdots,p; i = 1,\cdots,n)$ are also (discrete) martingales, we set

$X_{nj}(g,g) = (2\sum_{i=1}^{j-1}\sqrt{m_n}c_{ij}\frac{r_{gi}}{\sqrt{n}})\frac{r_{gj}}{\sqrt{n}}$ $(j = 2,\cdots,n)$ and $V_{gg.n}^{*}(g,g) = \sum_{j=2}^{n}\mathcal{E}[X_{nj}^2|\mathcal{F}_{n,j-1}]$.

In order to prove

$$\text{(A.17)} \qquad U_n(g,g) = \sum_{i=1}^{n}X_{ni}(g,g) \xrightarrow{\mathcal{L}} N(0,V(g,g))$$

we need to show the conditions (i) $\sum_{i=1}^{n}\mathcal{E}[X_{ni}(g,g)^2|\mathcal{F}_{n,i-1}] \xrightarrow{p} V(g,g)$ and (ii) $\sum_{i=1}^{n}\mathcal{E}[X_{ni}(g,g)^2 I(|X_{ni}(g,g)| > \epsilon)|\mathcal{F}_{n,i-1}] \xrightarrow{p} 0$ (for any $\epsilon > 0$).

In the present situation, these conditions are satisfied, which have been basically given in the proof of Theorem 3 in Kunitomo and Sato (2008) as its special case.

For the covariance of the trend term $\sigma_{sf}^{(x)}$ $(s,f = 1,\cdots,p)$, we have the similar arguments and obtain the corresponding results.

(**Q.E.D.**)

Proof of Theorem 4.2 : By the proof of Theorem 4.1, we have found that the main order of the bias of the SIML estimator is $m_n^{-1}\sum_{k=1}^{m_n}a_{kn} = O(n^{2\alpha-2})$. Since the normalization of the SIML estimator is in the form of $\sqrt{m_n}[\hat{\sigma}_{gg}^{(x)} - \sigma_{gg}^{(x)}] = O_p(1)$, its variance is of the order $O(n^{-\alpha})$.

Hence when n is large we can approximate the mean squared error of $\hat{\sigma}_{gg}^{(x)}$ $(g = 1, \cdots, p)$ as

$$g_n(\alpha) = c_{1g}\frac{1}{n^{\alpha}} + c_{2g}n^{4\alpha-4} \ , \tag{A.18}$$

where c_{1g} and c_{2g} are some constants. The first term and the second term correspond to the order of the variance and the squared bias, respectively. By minimizing $g_n(\alpha)$ with respect to α, we obtain an optimal choice of m_n.

(**Q.E.D.**)

Proof of Theorem 4.3 : We consider the sample characteristic equation

$$\left[\hat{\mathbf{\Sigma}}_x - \lambda_1 \mathbf{\Sigma}_v\right]\hat{\boldsymbol{\beta}} = \mathbf{0} \ , \tag{A.19}$$

when λ_1 is the smallest root of the corresponding characteristic equation. By Theorem 4.1 we have

$$\hat{\mathbf{\Sigma}}_x \xrightarrow{p} \mathbf{\Sigma}_x \tag{A.20}$$

and we use

$$\hat{\boldsymbol{\beta}}^{'}\left[\hat{\mathbf{\Sigma}}_x - \lambda_1 \mathbf{\Sigma}_v\right]\hat{\boldsymbol{\beta}} = 0 \ . \tag{A.21}$$

Then we find $\lambda_1 \xrightarrow{p} 0$ as (4.10) because λ_1 is the minimum root of the characteristic equation and the rank of $\mathbf{\Sigma}_x$ is less than p. Since $\mathbf{\Sigma}_v$ is a nonsingular matrix, we have the consistency of (4.8). Furthermore, due to Part (ii) of Theorem 4.1 we write $\sqrt{m_n}[\hat{\mathbf{\Sigma}}_x - \mathbf{\Sigma}_x] \xrightarrow{\mathcal{L}} \mathbf{S}$, and then we have

$$\sqrt{m_n}\lambda_1 \xrightarrow{p} 0 \ . \tag{A.22}$$

Then we rewrite the sample characteristic equation

$$\text{(A.23)} \quad \left[(\mathbf{\Sigma}_x + \frac{1}{\sqrt{m_n}}\mathbf{S}) - \lambda_1 \mathbf{\Sigma}_v\right]\left[\boldsymbol{\beta} + \frac{1}{\sqrt{m_n}}\sqrt{m_n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})\right] = o_p(1) \ ,$$

which is asymptotically equivalent to

$$\text{(A.24)} \qquad \mathbf{\Sigma}_x \sqrt{m_n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) - \mathbf{S}\boldsymbol{\beta} = o_p(1) \ .$$

We use the representation

$$\text{(A.25)} \qquad \sqrt{m_n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \sqrt{m_n}\begin{pmatrix} 0 \\ -(\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}_2) \end{pmatrix} \ ,$$

where $\hat{\boldsymbol{\beta}}_2$ and $\boldsymbol{\beta}_2$ are is the $(p-1) \times 1$ vector of $\hat{\boldsymbol{\beta}}$ and $\boldsymbol{\beta}$. Then by multiplying the choice matrix $(\mathbf{0}, \mathbf{I}_{p-1})$ from the left, we have the asymptotic distribution of $\hat{\boldsymbol{\beta}}$.

(**Q.E.D.**)

For the proofs of Theorem 5.1 and Theorem 5.2, we give some preliminary lemmas, which are keys in our arguments.

Lemma A-1 : Let

$$\text{(A.26)} \qquad \mathbf{B}_n^{(1)} = (b_{jk}^{(1)}) = \mathbf{P}_n \mathbf{C}_n^{(s)-1}$$

in (5.5). Then we have

$$\text{(A.27)} \qquad \sum_{j=1}^{n} b_{kj}^{(1)} b_{k',j}^{(1)} = \delta(k, k') 4 \sin^2\left[\frac{\pi}{2}\frac{2k-1}{2n+1}s\right] + O(\frac{1}{n}) \ .$$

Lemma A-2 : Let

$$\text{(A.28)} \qquad \mathbf{B}_n^{(2)} = (b_{jk}^{(2)}) = \mathbf{P}_n \mathbf{C}_n^{(s)-1} \mathbf{C}_n \ .$$

Then we have

$$\text{(A.29)} \qquad \sum_{j=1}^{n-s} b_{kj}^{(2)} b_{k',j}^{(2)} = \delta(k,k') \frac{\sin^2\left[\frac{\pi}{2}\frac{2k-1}{2n+1}s\right]}{\sin^2\left[\frac{\pi}{2}\frac{2k-1}{2n+1}\right]} + O(\frac{1}{n}) \ .$$

Lemma A-3 : Let $n = Ns, N$ and s be positive integers and

$$\text{(A.30)} \qquad \mathbf{B}_n^{(3)} = (b_{jk}^{(3)}) = \mathbf{P}_n \mathbf{C}_n^{2} \mathbf{C}_n^{(s)-1} \ .$$

Then we have

$$\text{(A.31)} \qquad \sum_{j=1}^{n-s} b_{kj}^{(2)} b_{k',j}^{(2)} = \delta(k,k') 4 \frac{\sin^4\left[\frac{\pi}{2}\frac{2k-1}{2n+1}s\right]}{\sin^2\left[\frac{\pi}{2}\frac{2k-1}{2n+1}\right]} + O(\frac{1}{n}) \ .$$

Proof of Lemma A-1 : The proof is the result of lengthy, but straightforward calculations of the trigonometric functions. We set

$$\text{(A.32)} \qquad b_{kj}^{(1)} = p_{kj} - p_{k,j+s} \ \ (1 \le j \le n-s) \ ,$$

which can be written as

$$\text{(A.33)} \qquad b_{kj}^{(1)} = \frac{1}{\sqrt{2n+1}} \{ [1 - e^{i\frac{2\pi}{2n+1}(k-\frac{1}{2})s}] e^{i\frac{2\pi}{2n+1}(k-\frac{1}{2})(j-\frac{1}{2})} + [1 - e^{-i\frac{2\pi}{2n+1}(k-\frac{1}{2})s}] e^{-i\frac{2\pi}{2n+1}(k-\frac{1}{2})(j-\frac{1}{2})} \} \ .$$

Then we evaluate each terms of

$$\begin{aligned} \sum_{j=1}^{n-s} b_{kj}^{(1)} b_{k'j}^{(1)} &= \frac{1}{2n+1} \sum_{j=1}^{n-s} \{ [A_{1j}(k) + A_{2j}(k)][A_{1j}(k') + A_{2j}(k')] \} \\ \text{(A.34)} \qquad &= \frac{1}{2n+1} \sum_{j=1}^{n-s} \{ A_{1j}(k) A_{1j}(k') + A_{2j}(k) A_{2j}(k') \\ &\qquad + A_{1j}(k) A_{2j}(k') + A_{2j}(k) A_{2j}(k') \} \ , \end{aligned}$$

where we denote

$$A_{1j}(k) = (1 - e^{i\theta_k^s}) e^{i\theta_{k,j}} \ , A_{2j}(k) = (1 - e^{-i\theta_k^s}) e^{-i\theta_{k,j}} \ ,$$

and

$$\theta_k^s = \frac{2\pi}{2n+1}(k - \frac{1}{2})s,\ \theta_{k,j} = \frac{2\pi}{2n+1}(k - \frac{1}{2})(j - \frac{1}{2})\ .$$

There are four terms in the summation of (A.34). For instance, the first term of (A.34) is given by

$$\sum_{j=1}^{n-s} A_{1j}(k)A_{1j}(k^{'}) \ = \ (1 - e^{i\theta_k^s})(1 - e^{i\theta_{k'}^s})\frac{1 - e^{i\frac{2\pi}{2n+1}(k+k^{'}-1)(n-s+1)}}{1 - e^{i\frac{2\pi}{2n+1}(k+k'-1)}} \times e^{i\frac{2\pi}{2n+1}(k+k^{'}-1)\frac{1}{2}}$$

and the third term of (A.34) is

$$\sum_{j=1}^{n-s} A_{1j}(k)A_{2j}(k^{'}) \ = \ (1 - e^{i\theta_k^s})(1 - e^{-i\theta_{k'}^s})\frac{1 - e^{i\frac{2\pi}{2n+1}(k-k^{'})(n-s+1)}}{1 - e^{i\frac{2\pi}{2n+1}(k-k')}} \times e^{i\frac{2\pi}{2n+1}(k+k^{'}-1)\frac{1}{2}}$$

when $k \neq k^{'}$. When $k = k^{'}$, the third term of (A.34) becomes

$$\text{(A.35)} \quad \begin{aligned} \sum_{j=1}^{n-s} A_{1j}(k)A_{2j}(k^{'}) \ &= \ (n-s)(1 - e^{i\theta_k^s})(1 - e^{-i\theta_k^s}) \\ &= \ (n-s)(-1)[e^{-i\theta_k^s/2} - e^{i\theta_k^s/2}]^2 \\ &= \ 4(n-s)\sin^2[\frac{\theta_k^s}{2}]\ . \end{aligned}$$

Then by using similar calculations of the second and fourth terms and by summarizing four terms of (A.35), we have the desired result. (**Q.E.D.**)

Proof of Lemma A-2 : The derivation of Lemma A-2 is similar to that of Lemma A-1. For $k = 1, \cdots, n; j = 1, \cdots, n-s+1$, we set

$$\text{(A.36)} \quad b_{kj}^{(2)} = p_{kj} + \cdots + p_{k,j+s-1}\ ,$$

which can be written as

$$(A.37) \qquad b^{(2)}_{kj} = \frac{1}{\sqrt{2n+1}}\{\frac{1-e^{i\frac{2\pi}{2n+1}(k-\frac{1}{2})s}}{1-e^{i\frac{2\pi}{2n+1}(k-\frac{1}{2})}}e^{i\frac{2\pi}{2n+1}(k-\frac{1}{2})(j-\frac{1}{2})} + \frac{1-e^{-i\frac{2\pi}{2n+1}(k-\frac{1}{2})s}}{1-e^{-i\frac{2\pi}{2n+1}(k-\frac{1}{2})}}e^{-i\frac{2\pi}{2n+1}(k-\frac{1}{2})(j-\frac{1}{2})}\} .$$

Then the rest of derivation is similar to that of Lemma A-1.
(**Q.E.D.**)

Proof of Lemma A-3 : The derivation of Lemma A-3 is similar to those of Lemmas A-1 and A-2. For $k=1,\cdots,n; j=1,\cdots,n-s+1$, we set

$$b^{(3)}_{kj} = [(p_{kj}-p_{k,j-1})-(p_{k,j+1}-p_{k,j})]+\cdots$$
$$(A.38) \qquad +[(p_{k,(N-1)s}-p_{k,(N-1)s-1})-(p_{k,Ns}-p_{k,Ns-1})],$$

which can be written as

$$(A.39) \qquad b^{(3)}_{kj} = \frac{1}{\sqrt{2n+1}}\{\frac{(1-e^{i\frac{2\pi}{2n+1}(k-\frac{1}{2})})^2}{1-e^{i\frac{2\pi}{2n+1}(k-\frac{1}{2})s}}e^{i\frac{2\pi}{2n+1}(k-\frac{1}{2})(j-\frac{1}{2})} + \frac{(1-e^{-i\frac{2\pi}{2n+1}(k-\frac{1}{2})})^2}{1-e^{-i\frac{2\pi}{2n+1}(k-\frac{1}{2})s}}e^{-i\frac{2\pi}{2n+1}(k-\frac{1}{2})(j-\frac{1}{2})}\} .$$

Then the rest of derivation is similar to those of Lemmas A-1 and A-2.
(**Q.E.D.**)

Proof of Theorem 5.1 : The proof of Theorem 5.1 is similar to that of Theorem 4.1 except the fact we use a different transformation of seasonal effects. Let $n=sN$ and N is an integer. (In the general case when $n=sN+j$ $(1\le j<s)$ we need some arguments, but the effects of additional terms $n=sN+j$ $(1\le j<s)$ are small.)

We set $\mathbf{z}_k^{(x)} = (z_{kg}^{(x)})$,$\mathbf{z}_k^{(v)} = (z_{kg}^{(v)})$ and $\mathbf{z}_k^{(s)} = (z_{kg}^{(g)})$, $(k = 1, \cdots, n; g = 1, \cdots, p)$ be the k-th vector elements of $n \times p$ matrix such that

$$\mathbf{Z}_n^{(x)} = \mathbf{P}_n \mathbf{C}_n^{-1} (\mathbf{X}_n - \bar{\mathbf{Y}}_0) \ , \ \mathbf{Z}_n^{(v)} = \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{V}_n \ , \ \mathbf{Z}_n^{(s)} = \mathbf{P}_n \mathbf{C}_n^{-1} \mathbf{S}_n \ ,$$

where $\mathbf{S}_n = (\mathbf{s}_i^{'}) = (s_{ig})$, $\mathbf{V}_n = (\mathbf{v}_i^{'})$ $(= (v_{ig}))$ and $\mathbf{Z}_n = (\mathbf{z}_k^{'})$ $(= (z_{kg}))$ are $n \times p$ matrices with $z_{kg} = z_{kg}^{(x)} + z_{kg}^{(v)} + z_{kg}^{(s)}$. Then we can write

$$\mathbf{Z}_n^{(s)} = \mathbf{B}_n^{(3)} [\mathbf{C}_n \mathbf{C}_n^{(s)-1} \mathbf{S}_n]$$

and we use the fact that $(1 - \mathcal{L})^{-1} (1 - \mathcal{L}^s) \mathbf{s}_i = \mathbf{w}_i^{(s)}$ and $\mathbf{w}_i^{(s)}$ are the sequence of i.i.d. random variables for $i = s, s+1, \cdots, n$ in (2.3), where we have set $\mathbf{B}_n^{(3)}$ in (A.30).

Then we have several additional terms in the decomposition of $\mathbf{z}_k$ $(k = 1, \cdots, m_n)$ as

$$\frac{1}{\sqrt{m_n}} \sum_{k=1}^{m_n} \mathcal{E}(\mathbf{z}_k^{(s)} \mathbf{z}_k^{(s)'}), \frac{1}{\sqrt{m_n}} \sum_{k=1}^{m_n} \left[\mathbf{z}_k^{(s)} \mathbf{z}_k^{(s)'} - \mathcal{E}(\mathbf{z}_k^{(s)} \mathbf{z}_k^{(s)'}) \right],$$

and

$$\frac{1}{\sqrt{m_n}} \sum_{k=1}^{m_n} (\mathbf{z}_k^{(x)} \mathbf{z}_k^{(s)'} + \mathbf{z}_k^{(s)} \mathbf{z}_k^{(x)'}), \frac{1}{\sqrt{m_n}} \sum_{k=1}^{m_n} (\mathbf{z}_k^{(v)} \mathbf{z}_k^{(s)'} + \mathbf{z}_k^{(s)} \mathbf{z}_k^{(v)'}) \ .$$

We need to show that these terms are stochastically negligible. The resulting evaluations are rather straightforward, but quite tedious. We illustrate a typical argument such that for any constant $p \times 1$ vector $\mathbf{a}$ and $\mathbf{b}$ we have

$$\begin{aligned} \mathcal{E} \left[\frac{1}{\sqrt{m_n}} \sum_{k=1}^{m_n} \mathbf{a}^{'} \mathbf{z}_k^{(x)} \mathbf{z}_k^{(s)'} \mathbf{b} \right]^2 &\leq \frac{1}{m_n} \mathcal{E} \left[(\sum_{k=1}^{m_n} \mathbf{a}^{'} \mathbf{z}_k^{(x)})^2 (\sum_{k=1}^{m_n} \mathbf{z}_k^{(s)'} \mathbf{b})^2 \right] \\ &\leq \frac{1}{m_n} \mathcal{E} \left[(\sum_{k=1}^{m_n} \mathbf{a}^{'} \mathbf{z}_k^{(x)})^2 \right] \mathcal{E} \left[(\sum_{k=1}^{m_n} \mathbf{z}_k^{(s)'} \mathbf{b})^2 \right] \ . \end{aligned}$$

Then by using Lemma A-3 it is possible to see the fact that this term and other extra terms due to the seasonality are of the smaller order $o_p(1)$ than constants. Since the evaluation of each terms are quite similar to the proof of Theorem 4.1, we omit some details.
(**Q.E.D.**)

Proof of Theorem 5.2 : The proof of Theorem 5.3 is similar to that of Theorem 4.1 except the fact we use a different transformation of seasonal effects. Let $n = sN$ and N is an integer. (In the general case when $n = sN + j$ $(1 \le j < s)$ we need some arguments, but the effects of additional terms $n = sN + j$ $(1 \le j < s)$ are small.)

Let $\mathbf{z}_k^{(x)} = (z_{kg}^{(x)}), Z_k^{(v)} = (z_{kg}^{(v)})$ and $\mathbf{z}_k^{(s)} = (z_{kg}^{(s)})$ $(k = 1, \cdots, n; g = 1, \cdots, p)$ be the k-th vector elements of $n \times p$ matrices such that

$$\mathbf{Z}_n^{(x)} = \mathbf{P}_n \mathbf{C}_n^{(s)-1}(\mathbf{X}_n - \bar{\mathbf{Y}}_0) \ , \ \mathbf{Z}_n^{(v)} = \mathbf{P}_n \mathbf{C}_n^{(s)-1}\mathbf{V}_n \ , \mathbf{Z}_n^{(s)} = \mathbf{P}_n \mathbf{C}_n^{(s)-1}\mathbf{S}_n \ ,$$

respectively, and $\mathbf{X}_n = (\mathbf{x}_k^{'}) = (x_{kg})$, $\mathbf{V}_n = (\mathbf{v}_k^{'}) = (v_{kg})$, $\mathbf{S}_n = (\mathbf{s}_k^{'})$ $(= (s_{kg}))$ $\mathbf{Z}_n = (\mathbf{z}_k^{'})$ $(= (z_{kg}))$ are $n \times p$ matrices with $z_{kg} = z_{kg}^{(x)} + z_{kg}^{(v)} + z_{kg}^{(s)}$. (We have written z_{kg} as the $g-$th component of $\mathbf{z}_k$.) Then we can write

$$\mathbf{Z}_n^{(x)} = \mathbf{B}_n^{(2)} \mathbf{C}_n^{-1}(\mathbf{X}_n - \bar{\mathbf{Y}}_n) \ , \ \mathbf{Z}_n^{(v)} = \mathbf{B}_n^{(1)} \mathbf{V}_n \ ,$$

and we use the fact that $(1 - \mathcal{L})\mathbf{x}_i = \mathbf{w}_i^{(x)}$ and $\mathbf{w}_i^{(x)}$ are the sequence of i.i.d. random variables for $i = 2, \cdots, n$ in (2.2), where we have set $\mathbf{B}_n^{(1)}$ and $\mathbf{B}_n^{(2)}$ in (A.26) and (A.28).

Next, we extend the decomposition in the present case as

$$\sqrt{m_n}\left[\hat{\Sigma}_s - \Sigma_s\right]$$

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

In order to evaluate many terms, we use the relations of Lemma A-1 and Lemma A-2. For instance, we can find a positive constant $K_1^{'}$ such that

$$\mathcal{E}[(z_{ks}^{(v)})]^2 \leq K_1^{'} \times a_{kn}^{(s)} , \tag{A.40}$$

where

$$a_{kn}^{(s)} = 4\sin^2[\frac{\pi}{2n+1}(k-\frac{1}{2})s] .$$

Also we find that

$$\frac{1}{m_n}\sum_{k\in I_n(s)} a_{kn}^{(s)} = \frac{1}{m_n}2\sum_{k\in I_n(s)}\left[1-\cos(\pi\frac{2k-1}{2n+1})s\right] = O(\frac{m_n^2}{n^2}) \tag{A.41}$$

and then the second term of the decomposition becomes

$$\frac{1}{\sqrt{m_n}}\sum_{k\in I_n(s)} \mathcal{E}[z_{ks}^{(v)}]^2 \leq K_1^{'}\frac{1}{\sqrt{m_n}}\sum_{k\in I_n(s)} a_{kn}^{(s)} = O(\frac{m_n^{5/2}}{n^2}) . \tag{A.42}$$

This term is $o(1)$ if $0 < \alpha < 0.8$. The remaining arguments of the proof are quite similar to that of Theorem 4.1 and

$$\mathcal{E}\left[\frac{1}{\sqrt{m_n}}\sum_{j=1}^{m_n}((z_{kg}^{(2)})^2-\mathcal{E}[(z_{kg}^{(2)})^2])\right]^2 \leq K_2^{'}\frac{1}{m_n}[\sum_{k=1}^{m_n} a_{kn}]^2$$
$$= O(\frac{1}{m_n}\times(\frac{m_n^3}{n^2})^2) ,$$

where $K_2^{'}$ is a positive constant. Since the rest of arguments are quite similar to the proofs of Theorem 4.1 and Theorem 5.1, we omit some details.

(**Q.E.D.**)

Fig.1−1:Real GDP and Investment(red line)

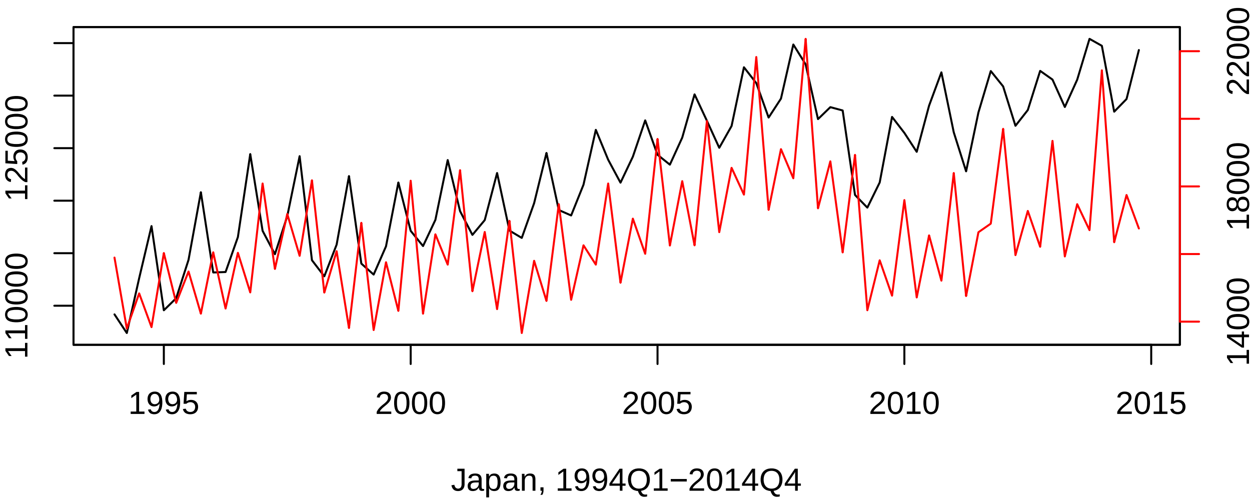

Fig.6−1:Trend+Noise

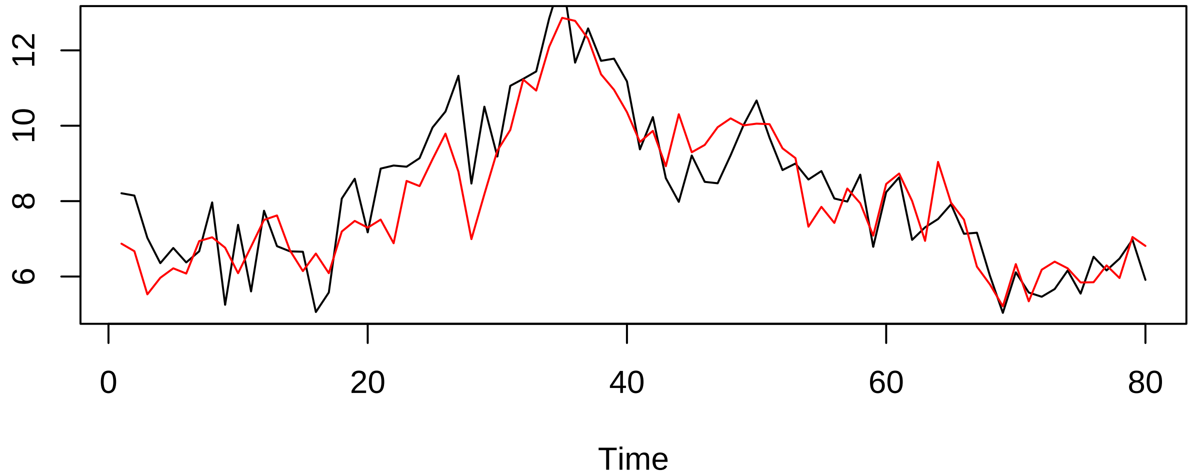

Fig.6−2:Trend+Seasonal+Noise

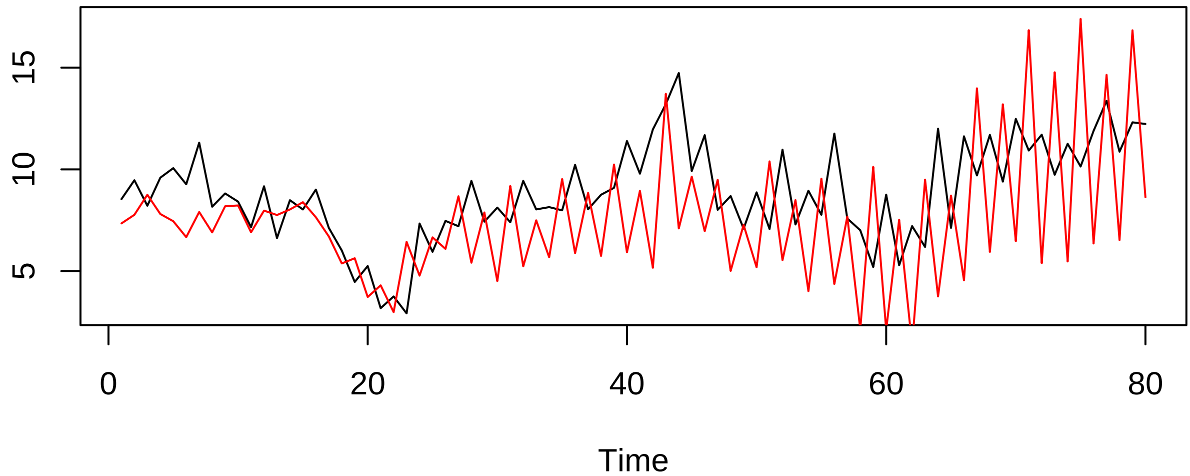

Fig.6−3:Trend+Seasonal(irregular case)+Noise

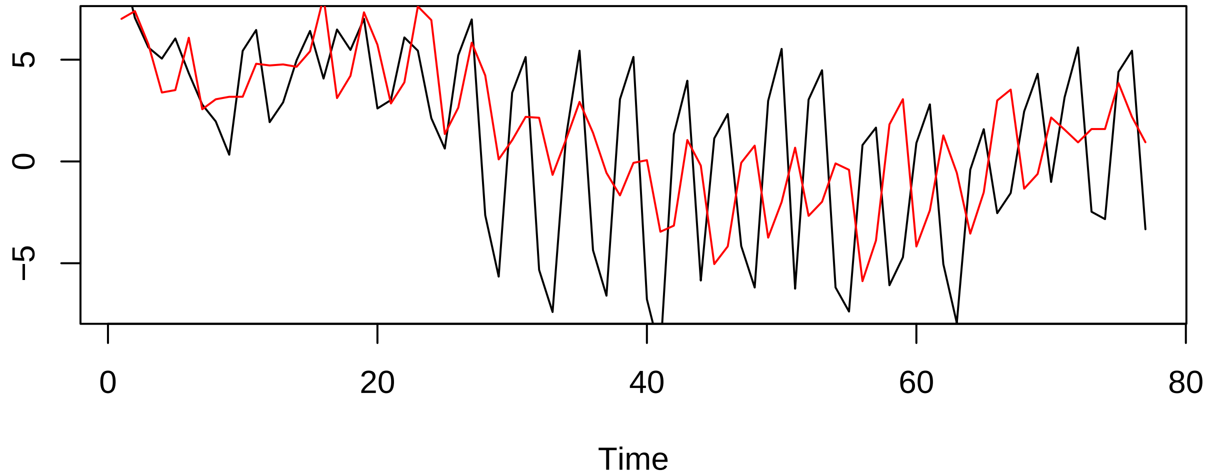

Fig.6−4:Trend+Seasonal+Noise(small)