11-617
no. 96-01

c. 3

Statistics Canada Statistique Canada

Methodology Branch

Business Survey Methods Division

Direction de la méthodologie

Division des méthodes d'enquêtes-entreprises

Canada

WORKING PAPER NO. BSMD-96-001E

METHODOLOGY BRANCH

CAHIER DE TRAVAIL NO. BSMD-96-001E

DIRECTION DE LA MÉTHODOLOGIE

A NONPARAMETRIC METHOD FOR BENCHMARKING SURVEY DATA VIA SIGNAL EXTRACTION

STATISTICS CANADA STATISTIQUE CANADA
APR AVR 23 1997
LIBRARY BIBLIOTHÈQUE

by

Z.-G. Chen, P.A. Cholette and E.B. Dagum

March 1996

A Nonparametric Method for Benchmarking Survey Data via Signal Extraction

by

Zhao-Guo Chen,[1] Pierre A. Cholette [1]

and

Estela Bee Dagum [2]

March 1996

Business Survey Methods Division

Statistics Canada

[1] Time Series Research and Analysis Center, 3H Coats Building, Statistics Canada, Ottawa, Ont. Canada K1A 0T6

[2] Faculty of Statistical Sciences, University of Bologna, Via delle Belle Arti 41, (40126) Bologna, Italy

Abstract

This paper compares two benchmarking methods, one based on regression modelling by Cholette and Dagum (1994) and the other based on ARIMA modelling signal extraction by Hillmer and Trabelsi (1987). The difficulties in the latter method of finding ARIMA models for the target series and the resulting implementation problems are discussed. A nonparametric approach to circumvent this problem is proposed, and more general models for estimating and benchmaking repeated survey data are introduced. Simulations show the proposed method and models are efficient and easy to implement.

Résumé

Le présent document compare deux méthodes d'étalonnage, l'une fondée sur la régression par Cholette et Dagum (1994), l'autre sur l'extraction de signal avec modèle ARIMA par Hillmer et Trabelsi (1987). Les difficultés liées à cette dernière, au niveau de la détermination du modèle ARIMA suivi par le signal et les problèmes d'application résultant, sont examinées. Une approche aparamétrique contournant le problème est proposée, et des modèles plus généraux pour l'estimation et l'étalonnage de données d'enquêtes périodiques sont présentés. Les simulations montrent que la méthode et les modèles proposés sont efficaces et facile d'application.

1 Introduction

Benchmarking is an important problem faced by statistical agencies. For a target socio-economic variable, say η, two sources of data from repeated survey with different frequencies and precisions may be available. For example, one set of data is its monthly observations

$$y_t = \eta_t + e_t, \quad t = 1, \ldots, n, \tag{1.1}$$

and the other set of data is annual:

$$z_i = \sum_{t \in i} \eta_t + \epsilon_i, \quad i = 1, \ldots, m. \tag{1.2}$$

where e_t and ϵ_i are the monthly and annual survey errors respectively, and $t \in i$ means month t is included in the time period covered by z_i. Usually, the "monthly" survey error is much larger than that of the "annual", so the "annual" data are referred to as benchmarks. Benchmarking is the process of estimating η_t from all "monthly" observations and "annual" benchmarks. When $\epsilon_i \equiv 0$, $\sum_{t \in i} \eta_t = z_i$ becomes a constraint and the benchmarks considered is called binding; and non-binding otherwise. Traditionary, benchmarks have been considered as binding (e.g. Denton, 1971).

Flow and stock series are already taken into account by the notation used in (1.2). For index series, (1.2) becomes (in case of monthly data)

$$z_i = \frac{1}{12} \sum_{t \in i} \eta_t + \epsilon_i. \tag{1.3}$$

The model defined by (1.1) and (1.2) (or (1.3)) can be written in matrix form as

$$\mathbf{y} = \boldsymbol{\eta} + \mathbf{e}, \tag{1.4}$$

$$\mathbf{z} = L\boldsymbol{\eta} + \boldsymbol{\epsilon}, \tag{1.5}$$

where $\mathbf{y} = (y_1 \cdots y_n)'$, $\mathbf{z} = (z_1 \cdots z_m)'$ and similarly for $\boldsymbol{\eta}$, $\mathbf{e}$ and $\boldsymbol{\epsilon}$ with suitable dimensions, and L is a $m \times n$ matrix. For example, corresponding to (1.3), the (i, t)th entry l_{it} of L is $l_{it} = 1/12$ if $t \in i$ and $l_{it} = 0$ otherwise.

In some situations, there are no benchmarks ((1.5) is absent), but with (1.4), some statistical information about series $\{\eta_t\}$ or $\{e_t\}$ is available, e.g. their covariance structures, then one can still obtain an estimate $\hat{\eta}_0$ of η from $\mathbf{y}$ alone, which is more precise than $\mathbf{y}$. Such a procedure is often called *signal extraction* (e.g. Whittle, 1963). However the availability of benchmarks makes it possible to further improve the estimate.

The benchmarking approaches most widely used by statistical agencies are of the Denton (1971) type, where $\boldsymbol{\eta}$ is estimated by minimizing the penalty function

$$P_A(\mathbf{y}, \boldsymbol{\eta}) = (\mathbf{y} - \boldsymbol{\eta})' A (\mathbf{y} - \boldsymbol{\eta}) \tag{1.6}$$

under the restriction (1.5) with $\boldsymbol{\epsilon} \equiv 0$ (binding benchmarks). By a suitable choice of the symmetric $n \times n$ matrix A, one may obtain a benchmarked value $\hat{\boldsymbol{\eta}}$ of $\boldsymbol{\eta}$ with some desired properties, such as continuity between consecutive years. This type of benchmarking is easy to apply, and no preliminary statistical information is required. But, because no information on the nature of the time series is used (when available), the estimation error, i.e. Var $(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}) = E\{(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})'\}$ is not necessarily minimized.

Cholette and Dagum (1994) combine (1.4) and (1.5) in the following *regression model*

$$\boldsymbol{\tau} = X\boldsymbol{\eta} + \mathbf{u}, \tag{1.7}$$

assuming that $\boldsymbol{\eta}$ is the vector of unknown parameters, where

$$\boldsymbol{\tau} = \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix}, \qquad X = \begin{pmatrix} I_n \\ L \end{pmatrix}, \qquad \mathbf{u} = \begin{pmatrix} \mathbf{e} \\ \boldsymbol{\epsilon} \end{pmatrix} \sim (0, \Sigma_u), \tag{1.8}$$

and $\mathbf{u} \sim (0, \Sigma_u)$ means that the random vector $\mathbf{u}$ has unspecified distribution with mean 0 and covariance matrix Σ_u. Assuming $\mathbf{e}$ and $\boldsymbol{\epsilon}$ are mutually uncorrelated, then Σ_u can be written as

$$\Sigma_u = \begin{pmatrix} \Sigma_e & 0 \\ 0 & \Sigma_\epsilon \end{pmatrix}. \tag{1.9}$$

Given Σ_u, the generalized least squares estimate of $\boldsymbol{\eta}$ is used as the benchmarked value:

$$\hat{\boldsymbol{\eta}}_R = (X'\Sigma_u^{-1}X)^{-1}X'\Sigma_u^{-1}\boldsymbol{\tau} = \mathbf{y} + \Sigma_e L'(L\Sigma_e L' + \Sigma_\epsilon)^{-1}(\mathbf{z} - L\mathbf{y}). \tag{1.10}$$

The covariance matrix of the estimation error, Var $(\hat{\boldsymbol{\eta}}_R - \boldsymbol{\eta})$, is

$$\begin{aligned} \Omega_R &= (X'\Sigma_u^{-1}X)^{-1} = (\Sigma_e^{-1} + L'\Sigma_\epsilon^{-1}L)^{-1} \\ &= \Sigma_e - \Sigma_e L'(L\Sigma_e L' + \Sigma_\epsilon)^{-1}L\Sigma_e, \end{aligned} \tag{1.11}$$

Under the criterion of minimizing the variance of the estimation error, the *regression benchmarking method* produces the *best linear unbiased estimates* (BLUE), provided that $\boldsymbol{\eta}$ is considered as a vector of fixed parameters.

Hillmer and Trabelsi (1987) consider that η_t follows an ARIMA model and e_t follows an ARMA model. Under the assumption of normality, they derive their

benchmarking formulae via signal extraction. Section 2 introduces their method and shows that, benchmarking via signal extraction produces estimates with lower variance and that the assumption of normality may be relaxed. Section 3 extends signal extraction methods and models and provides corresponding formulae. Section 4 proposes a nonparametric procedure to implement these formulae; Section 5 presents some simulation results which show the proposed approach is feasible, reliable and efficient. Section 6 summarizes the results.

2 Benchmarking via Signal Extraction

This section summarizes the signal extraction benchmarking method of Hillmer and Trabelsi (1987) and compares it with the regression method of Cholette and Dagum (1994).

2.1 The signal extraction method for benchmarking

Hillmer and Trabelsi (1987) assume that η_t, the target series, follows an ARIMA model

$$\varphi_\eta(B)(\eta_t - \mu) = \theta_\eta(B)b_t, \tag{2.1}$$

while e_t, the survey error, follows a stationary ARMA model (with zero mean)

$$\phi_e(B)e_t = \theta_e(B)c_t. \tag{2.2}$$

Assuming that the white noise shocks $\{b_t\}$ and $\{c_t\}$ are mutually uncorrelated, then $y_t = \eta_t + e_t$ also follows an ARIMA model

$$\varphi(B)(y_t - \mu) = \theta(B)a_t. \tag{2.3}$$

In our opinion, the more important contribution of Hillmer and Trabelsi is the following formulae which are valid whether or not $\boldsymbol{\eta}$ and $\mathbf{e}$ follow ARIMA models. Under the normality assumption

$$\boldsymbol{\eta} \sim N(\boldsymbol{\mu}, \Sigma_\eta), \qquad \mathbf{e} \sim N(\mathbf{0}, \Sigma_e), \tag{2.4}$$

they obtain the conditional expectation $\hat{\boldsymbol{\eta}}$ of $\boldsymbol{\eta}$ given $\boldsymbol{\tau} = (\mathbf{y}' \ \mathbf{z}')'$:

$$\hat{\boldsymbol{\eta}}_0 = E(\boldsymbol{\eta} \mid \mathbf{y}) = \Omega_0(\Sigma_e^{-1}\mathbf{y} + \Sigma_\eta^{-1}\boldsymbol{\mu}). \tag{2.5}$$

$$\hat{\boldsymbol{\eta}}_C = \Omega_0 L'(L\Omega_0 L' + \Sigma_\epsilon)^{-1}(\mathbf{z} - L\hat{\boldsymbol{\eta}}_0). \tag{2.6}$$

$$\hat{\boldsymbol{\eta}} = E(\boldsymbol{\eta} \mid \boldsymbol{\tau}) = \hat{\boldsymbol{\eta}}_0 + \hat{\boldsymbol{\eta}}_C, \tag{2.7}$$

where

$$\Omega_0 = \text{Var}\ (\hat{\boldsymbol{\eta}}_0 - \boldsymbol{\eta}) = (\Sigma_e^{-1} + \Sigma_\eta^{-1})^{-1}. \tag{2.8}$$

The covariance matrix of the error of the benchmarked estimate $\hat{\boldsymbol{\eta}}$ is

$$\begin{aligned}\Omega = \text{Var}\ (\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}) &= (\Omega_0^{-1} + L'\Sigma_\epsilon^{-1}L)^{-1} = (\Sigma_e^{-1} + \Sigma_\eta^{-1} + L'\Sigma_\epsilon^{-1}L)^{-1} \\ &= \Omega_0 - \Omega_0 L'(L\Omega_0 L' + \Sigma_\epsilon)^{-1}L\Omega_0. \end{aligned} \tag{2.9}$$

The *extracted signal* $\hat{\boldsymbol{\eta}}_0$ is the best unbiased estimate of $\boldsymbol{\eta}$ given $\mathbf{y}$ alone and the *benchmarked estimate* $\hat{\boldsymbol{\eta}}$ is the best unbiased estimate of $\boldsymbol{\eta}$ given $\mathbf{y}$ and $\mathbf{z}$. The correction originating from the benchmarks $\mathbf{z}$ is $\hat{\boldsymbol{\eta}}_C$.

The normality assumption (2.4) can be generalized to

$$\boldsymbol{\eta} \sim (\boldsymbol{\mu}, \Sigma_\eta), \quad \mathbf{e} \sim (0,\ \Sigma_e). \tag{2.10}$$

It is well-known that the projection of $\boldsymbol{\eta}$ on the linear probability space generated by $\mathbf{y}$ or by $\boldsymbol{\tau}$, denoted by $\hat{\boldsymbol{\eta}}_0 = \hat{E}(\boldsymbol{\eta} \mid \mathbf{y})$ or $\hat{\boldsymbol{\eta}} = \hat{E}(\boldsymbol{\eta} \mid \boldsymbol{\tau})$, are the BLUE of $\boldsymbol{\eta}$ from $\mathbf{y}$ or from $\boldsymbol{\tau}$ respectively. Appendix A shows that formulae (2.5) through (2.9) still hold when E is replaced by $\hat{E}$.

2.2 Comparisons with the regression method

To compare benchmarking via signal extraction with the regression method of Cholette and Dagum (1994) the following inequality will be used. Suppose A and B are symmetric matrices of the same dimension, $A > 0$ (i.e. positive definite), $B \geq 0$ (i.e. non-negative definite), then

$$(A + B)^{-1} \leq A^{-1}. \tag{2.11}$$

The signal extraction method is now compared to the regression method. Denoting the last term of (1.10) as

$$\hat{\boldsymbol{\eta}}_{RC} = \Sigma_e L'(L\Sigma_e L' + \Sigma_\epsilon)^{-1}(\mathbf{z} - L\mathbf{y}), \tag{2.12}$$

and examining (1.10), (1.11) and (2.5) through (2.9), the following correspondence is observed

$$\mathbf{y} \leftrightarrow \hat{\boldsymbol{\eta}}_0, \quad \hat{\boldsymbol{\eta}}_{RC} \leftrightarrow \hat{\boldsymbol{\eta}}_C, \quad \hat{\boldsymbol{\eta}}_R \leftrightarrow \hat{\boldsymbol{\eta}}; \tag{2.13}$$

$$\Sigma_e \leftrightarrow \Omega_0, \quad \Omega_R \leftrightarrow \Omega. \tag{2.14}$$

Obviously, the last expressions of (1.11) and of (2.9) yield

$$\Omega_R \leq \Sigma_e, \quad \Omega \leq \Omega_0. \tag{2.15}$$

The first inequality in (2.15) means that the variance of the error of the benchmarked estimate based on the regression method (Ω_R) is smaller than that of the survey error (Σ_e); and the second inequality means that the variance of the error of the benchmarked estimate via the signal extraction method (Ω) is smaller than that of the extracted signal (Ω_0). So, *benchmarking reduces the error in both methods.*

Applying inequality (2.11) to (2.8), (2.9) and (1.11) yields

$$\Omega_0 \leq \Sigma_e, \quad \Omega = \{(\Sigma_e^{-1} + L'\Sigma_\epsilon^{-1}L) + \Sigma_\eta^{-1}\}^{-1} \leq (\Sigma_e^{-1} + L'\Sigma_\epsilon^{-1}L)^{-1} = \Omega_R. \tag{2.16}$$

This means that, the variance of the error of the extracted signal is smaller than that of the survey error; and the variance of the error of the benchmarked estimate via the signal extraction method is smaller than that of the regression method. So, *if the first and the second order moments of both the survey error and the target series are known, then benchmarking via signal extraction is better than the regression method.* The reason is that the regression method does not use Σ_η, the information about $\{\eta_t\}$.

3 Implementation of Benchmarking via Signal Extraction

Survey experts can provide, sometimes, model (2.2) for the survey error e_t from which the autocovariance function (ACVF) of $\{e_t\}$ ($Ee_t = 0$)

$$\gamma_e(-k) = \gamma_e(k) = E(e_t e_{t+k}) \tag{3.1}$$

may be obtained (an algorithm is given by McLeod, 1975). In other cases, survey experts can provide directly $\gamma_e(k)$ without a model (e.g. Lee, 1990). In reality, $\gamma_e(k)$ is obtained first and then an ARMA model is fitted (e.g. Scott et al., 1977). In both cases, the Toeplitz matrix Σ_e is obtained by setting its (i, j)th entry equals to $\gamma_e(i-j)$. The regression benchmarking method can then be applied without difficulty. However, benchmarking via signal extraction requires the knowledge of $\boldsymbol{\mu} = E\boldsymbol{\eta} = E\mathbf{y}$ and of Σ_η which are usually unknown. These two difficulties in the implementation are now addressed in this section.

3.1 The difference stationary model

Hillmer and Trabelsi (1987) assume that η_t follows the ARIMA model (see (2.1) - (2.3)). This can be generalized by assuming a difference stationary (DS) model (Nelson and Plosser, 1982), that is

$$\nabla^d \nabla_s^{\tilde{d}} \eta_t = \zeta_t, \tag{3.2}$$

where s is the period of seasonality, say 12 for monthly data; $\nabla = 1 - B$, $\nabla_s = 1 - B^s$; ζ_t is a stationary series with mean zero not necessarily following an ARMA model.

First, let us consider Σ_η. The ACVF for a series following a nonstationary model like (3.2) is not well defined, it depends on the distribution of the initial values. For instance, if $\nabla y_t = x_t$, x_t is stationary, $y_t = y_1 + \sum_{j=2}^{t} x_j$, and the ACVF Cov (y_s, y_t), $s, t = 1, \ldots, n$, depends on Var y_1 and Cov (y_1, x_j) which can be defined in various ways. Cleveland and Tiao (1976) assume that the ARIMA series $\{\eta_t\}$ started from $-\infty$ which results in the variances of initial values $\boldsymbol{\eta}_i = (\eta_1 \ \ldots \ \eta_{d+\tilde{d}s})'$ converging to ∞, then a formula for Σ_η^{-1} is derived. Instead, we directly assume that

$$\text{Var } \boldsymbol{\eta}_i = h I_{d+\tilde{d}s}, \quad h \to \infty; \tag{3.3}$$

$$\text{Cov } (\boldsymbol{\eta}_i, \ \zeta_t) = 0, \quad t = d + \tilde{d}s + 1, \ldots, n. \tag{3.4}$$

which leads to the same results as those of Cleveland and Tiao.

Assumption (3.3) can be explained in a Bayesian framework: in the absence of prior knowledge about the distribution of the initial values of the series, their variance is set to ∞, and (3.3) is a standard representation for that. Another terminology for (3.3) is "assuming the initial values are diffuse". Assumption (3.4) means that the covariance structure of further development of the series does not depend on the initial state (the level and the variation of the initial values). This is again a commonly used assumption.

Denoting $\tilde{n} = n - d - \tilde{d}s$, from (3.2), the $\tilde{n} \times 1$ vector $\boldsymbol{\zeta} = (\zeta_{d+\tilde{d}s+1} \cdots \zeta_n)'$, may be written

$$\boldsymbol{\zeta} = D\boldsymbol{\eta}, \tag{3.5}$$

where D is $\tilde{n} \times n$ matrix. For example, $d = 1$ and $\tilde{d} = 1$, then its first diagonal has all elements equal to 1; those in the second diagonal (on the right), equal to -1; the $(s+1)$th and the $(s+2)$th diagonals, equal to -1 and 1 respectively; all other elements are zero.

Since $\{\zeta_t\}$ is stationary, the $\tilde{n} \times \tilde{n}$ Toeplitz matrix Σ_ζ is well defined. Appendix B shows that Σ_η^{-1} is then $D'\Sigma_\zeta^{-1}D$. A nonparametric method is proposed in the next section to estimate Σ_ζ.

Now turn to the problem of $\boldsymbol{\mu}$. For η_t satisfying (3.2), say, with $\tilde{d} = 1$, let

$$\mu_t = \alpha_0 + \alpha_1 t + \ldots + \alpha_d t^d + \sum_{j=1}^{s} \beta_j \delta_{\{t,j\}}, \tag{3.6}$$

with α_i and β_j constants and

$$\delta_{\{t,j\}} = \begin{cases} 1, & t \equiv j (\text{mod } s), \\ 0, & \text{otherwise}, \end{cases} \tag{3.7}$$

then also $\eta_t - \mu_t$ satisfies (3.2), because μ_t as in (3.6) is eliminated by $\nabla^d \nabla_s$. In other words, the mean of η_t has the form as in (3.6) (but the specified value is unknown). Appendix B proves that, under model assumptions (3.2), (3.3) and (3.4), $\boldsymbol{\mu}$ having elements as in (3.6) cancels in (2.5), i.e.

$$\hat{\boldsymbol{\eta}}_0 = \Omega_0 \Sigma_e^{-1} \mathbf{y} \tag{3.8}$$

only depends on the data, and that (2.8) becomes

$$\Omega_0 = (\Sigma_e^{-1} + D'\Sigma_\zeta^{-1}D)^{-1}. \tag{3.9}$$

These are answers to the questions about the difficulties of implementation for the DS model.

3.2 The trend stationary model

An alternative to the DS model is the *trend stationary* (TS) model (Nelson and Plosser, 1982), i.e.

$$\eta_t = \mu_t + \zeta_t. \tag{3.10}$$

We here assume that the target series $\{\eta_t\}$ has a deterministic component $\{\mu_t\}$ (the mean of the series), and a stationary stochastic component $\{\zeta_t\}$ with mean zero and autocorrelated. This model has been used in earlier works, e.g. Pierce (1978), Jones (1980).

Usually, the deterministic component μ_t is assumed to follow a linear model

$$\mu_t = \sum_{i=1}^{k} x_{ti}\beta_i = \mathbf{x}_t'\boldsymbol{\beta}. \tag{3.11}$$

which can represent a polynomial trend plus a stable seasonality as in (3.6), an abrupt changes caused by interventions (e.g. Box and Tiao, 1975), outliers (e.g. Tsay, 1986), trading-day variations (e.g. Bell and Hillmer, 1983; Dagum, Quenneville and Sutradhar, 1992) and so on.

Assuming the model for η_t is given by (3.10) and (3.11), we define

$$w_t = \zeta_t + e_t, \tag{3.12}$$

where $\{w_t\}$ is stationary, because $\{\zeta_t\}$ and $\{e_t\}$ are uncorrelated with each other and both are stationary. Then (1.1) and (3.10) lead to

$$y_t = \mu_t + w_t, \tag{3.13}$$

which is again a TS model with μ_t specified by (3.11) and can be estimated without difficulty.

For the moment, assume that μ_t is known, then $\mathbf{w} = (w_1 \ldots w_n)' = \mathbf{y} - \boldsymbol{\mu}$ can be considered as data; and $\boldsymbol{\zeta} = (\zeta_1 \ldots \zeta_n)'$, as the signal. Since $\{\zeta_t\}$ is stationary (then, its ACVF is well defined) and $E\boldsymbol{\zeta} = 0$, the benchmarking formula (2.5) through (2.9) are applicable. The formulae of signal extraction (2.5) and (2.8) become

$$\hat{\boldsymbol{\zeta}}_0 = \hat{E}(\boldsymbol{\zeta} \mid \mathbf{w}) = (\Sigma_e^{-1} + \Sigma_\zeta^{-1})^{-1}\Sigma_e^{-1}\mathbf{w}, \tag{3.14}$$

and

$$\mathrm{Var}\,(\hat{\boldsymbol{\zeta}}_0 - \boldsymbol{\zeta}_0) = \Omega_0 = (\Sigma_e^{-1} + \Sigma_\zeta^{-1})^{-1}. \tag{3.15}$$

Letting $\hat{\boldsymbol{\eta}}_0 = \boldsymbol{\mu} + \hat{\boldsymbol{\zeta}}_0$, $\hat{\boldsymbol{\eta}}$ follows from (2.6) and (2.7). Since $\{\eta_t\}$ and $\{e_t\}$ are mutually uncorrelated with each other, (3.10) implies that $\{\eta_t\}$ and $\{\zeta_t\}$, and $\{y_t\}$ and $\{w_t\}$, have the same ACVF, i.e.

$$\gamma_\eta(\pm k) = \gamma_\zeta(\pm k) = E(\zeta_t\zeta_{t+k}), \tag{3.16}$$

$$\gamma_y(\pm k) = \gamma_w(\pm k) = E(w_t w_{t+k}) = \gamma_e(\pm k) + \gamma_\zeta(\pm k), \tag{3.17}$$

which implies $\Sigma_w = \Sigma_e + \Sigma_\zeta$. Using identity (A.6) in Appendix A, (3.14) and (3.15) can be written as

$$\hat{\boldsymbol{\zeta}}_0 = \Sigma_\zeta\Sigma_w^{-1}\mathbf{w}, \tag{3.18}$$

$$\Omega_0 = \Sigma_\zeta\Sigma_w^{-1}\Sigma_e, \tag{3.19}$$

which require less matrix inversion and again Σ_ζ and Σ_w can be estimated by the method proposed in the next section.

4 A Nonparametric Procedure

This section focuses on the nonparametric estimation of Σ_η. In view of (1.1), it seems possible to get a model for $\{\eta_t\}$ from the model of $\{e_t\}$, which is known, and a model of $\{y_t\}$ fitted to the data; from the model of $\{\eta_t\}$, an estimate of Σ_η may be obtained. Indeed, assuming that $\{\eta_t\}$ and $\{e_t\}$ are uncorrelated with each other and that their ARIMA models are known or can be adequately fitted, it is relatively easy to derive the model of $\{y_t\}$, $y_t = \eta_t + e_t$. However, what we face is the reverse, the so-called modelling errors-in-variables problem: to obtain a model for $\{\eta_t\}$ given data of $\{y_t\}$ and the model of $\{e_t\}$. This is difficult even in the case that both series follow stationary ARMA models, because $\{y_t\}$ is correlated with $\{e_t\}$. The most recent results only deal with the case where $\{e_t\}$ follows a MA model and where the orders of the ARMA model of η_t are known (Chanda, 1995).

In fact, we may avoid deriving a model for $\{\eta_t\}$, because in the benchmarking formulae (2.5) through (2.8), only Σ_η is necessary. This shortcut is also observed by Jones (1980), who points out that, with repeated survey data, the number of observations is often quite low, and fit of an ARMA model will lead to unstable parameter estimates. We will develop a nonparametric method to estimate Σ_η.

4.1 Estimation procedure for the TS model

Let us consider the TS model and assume $\boldsymbol{\mu}$ is known for the moment. As $\{\zeta_t\}$ and $\{e_t\}$ are mutually uncorrelated with each other, (3.12) implies

$$\Sigma_\zeta = \Sigma_w - \Sigma_e. \tag{4.1}$$

All these theoretical covariance matrices are positive definite.

Since Σ_e is known, an estimate $\hat{\Sigma}_w$ of Σ_w can be obtained from $\mathbf{w} = \mathbf{y} - \boldsymbol{\mu}$, and an estimate $\hat{\Sigma}_\zeta$ of Σ_ζ follows. However, this solution does not guarantee that matrix $\hat{\Sigma}_\zeta$ is positive definite:

$$\hat{\Sigma}_\zeta = \hat{\Sigma}_w - \Sigma_e > 0. \tag{4.2}$$

Our simulation in Section 5 shows that, without this property, the benchmarked estimates become very unstable. A method for producing positive definite estimate of Σ_ζ is now proposed.

We first present the theoretical underpinning of the method assuming $\boldsymbol{\mu}$ is known and then state the steps of the benchmarking procedure.

Assuming both $\{e_t\}$ and $\{\zeta_t\}$ have no deterministic component in their Wold decomposition (see, e.g. Doob, 1953), so does $\{w_t\}$, then the spectral densities of

these series exist. Denote them by $f_e(\lambda)$, $f_\zeta(\lambda)$ and $f_w(\lambda)$ respectively, then (3.17) leads to

$$f_w(\lambda) = f_e(\lambda) + f_\zeta(\lambda), \quad -\pi < \lambda \leq \pi. \tag{4.3}$$

The most common estimator for $\gamma_w(-k) = \gamma_w(k)$ is

$$\hat{\gamma}_w(k) = \frac{1}{n}\sum_{t=1}^{n-k} w_t w_{t+k}, \quad k = 0, \ldots, n-1. \tag{4.4}$$

Its direct transformation to the frequency domain

$$\hat{f}_w(\lambda) = \frac{1}{\pi}\{\frac{\hat{\gamma}_w(0)}{2} + \sum_{k=1}^{n-1}\hat{\gamma}_w(k)\cos k\lambda\}, \quad -\pi < \lambda \leq \pi \tag{4.5}$$

is the periodogram. As an estimate of the spectral density $f_w(\lambda)$ of $\{w_t\}$, $\hat{f}_w(\lambda)$ is ragged and not consistent. A consistent estimate $\hat{f}_w^{(M)}(\lambda)$ is obtained by smoothing the periodogram, i.e. by correspondingly using

$$\hat{\gamma}_w^{(M)}(k) = \begin{cases} \hat{\gamma}_w(k)u(k/M), & k = 0, \ldots, M, \\ 0, & M < k \leq n. \end{cases} \tag{4.6}$$

in (4.4), where $u(s)$ is a commonly used lag window function, such as Tukey-Hanning window, Parzen window (e.g. Priestley 1981). The Toeplitz matrix $\hat{\Sigma}_w^{(M)}$ is formed from $\hat{\gamma}_w^{(M)}(k)$. However $\hat{\Sigma}_\zeta^{(M)} = \hat{\Sigma}_w^{(M)} - \Sigma_e > 0$ may not hold, so we introduce the following revised procedure.

As $\gamma_e(k)$, or equivalently $f_e(\lambda)$, is given and in view of (4.3), we may take $\hat{f}_\zeta(\lambda) = \hat{f}_w(\lambda) - f_e(\lambda)$ as an estimate of $f_\zeta(\lambda)$. Because a spectral density is always non-negative,

$$\tilde{f}_\zeta(\lambda) = \max\{\hat{f}_w(\lambda) - f_e(\lambda),\ 0\} \tag{4.7}$$

is a more reasonable estimate of $f_\zeta(\lambda)$ than $\hat{f}_\zeta(\lambda)$. The corresponding estimate of $\gamma_\zeta(k)$ is

$$\tilde{\gamma}_\zeta(k) = 2\int_0^\pi \tilde{f}_\zeta(\lambda)\cos k\lambda d\lambda, \ k = 0, \ldots, n-1. \tag{4.8}$$

Like $\hat{f}_\zeta(\lambda)$, $\tilde{f}_\zeta(\lambda)$ is again ragged and not consistent, a better estimate $\tilde{f}_\zeta^{(M)}(k)$ is produced by using

$$\tilde{\gamma}_\zeta^{(M)}(k) = \begin{cases} \tilde{\gamma}_\zeta(k)u(k/M), & k = 0, \ldots, M, \\ 0, & M < k \leq n. \end{cases} \tag{4.9}$$

in (4.5). The estimate of $\gamma_\zeta(k)$ is set to

$$\tilde{\gamma}_w^{(M)}(k) = \gamma_e(k) + \tilde{\gamma}_\zeta^{(M)}(k), \quad k = 0, \ldots, n. \tag{4.10}$$

We call $\tilde{\gamma}_\zeta^{(M)}(k)$ and $\tilde{\gamma}_w^{(M)}(k)$ the *revised window estimates* of $\gamma_\zeta(k)$ and $\gamma_w(k)$ respectively. The Toeplitz matrix $\tilde{\Sigma}_\zeta^{(M)}$ and $\tilde{\Sigma}_w^{(M)}$ obtained from $\tilde{\gamma}_\zeta^{(M)}(k)$ and $\tilde{\gamma}_w^{(M)}(k)$ are used as the estimates of Σ_ζ and Σ_w.

It is easy to show that, if $\tilde{f}_\zeta(\lambda) \geq 0$, and $\tilde{f}_\zeta(\lambda) > 0$ holds in a subinterval of $(-\pi, \pi]$, then $\tilde{\Sigma}_\zeta^{(M)} > 0$ (and hence $\tilde{\Sigma}_w^{(M)} = \Sigma_e + \tilde{\Sigma}_\zeta^{(M)} > 0$), so (4.2) is satisfied with this estimator. In view of (4.7), $\tilde{f}_\zeta(\lambda) \geq 0$ always holds. Theoretically, it can be proved that with probability as close to 1 as we wish, the strict inequality $\tilde{f}_\zeta(\lambda) > 0$ holds in the mentioned manner provided that n is large enough; in fact, in our simulation, it always holds for n as small as 28.

The proposed benchmarking procedure is as follows.

1 Apply OLS to the model (3.11) and (3.13) to obtain an estimate $\hat{\boldsymbol{\beta}}$ and then $\hat{\mu}_t$. (We will comment on this estimator later.)

2 Set $w_t = y_t - \hat{\mu}_t$ (the rigorous notation should be $\hat{w}_t = y_t - \hat{\mu}_t$), obtain $\hat{\gamma}_w(k)$, $\hat{f}_w(\lambda)$ and $\tilde{f}_\zeta(\lambda)$ from (4.4), (4.5) and (4.7). Calculate $\tilde{\gamma}_\zeta(k)$ from (4.8) using a numerical method, and then $\tilde{\gamma}_\zeta^{(M)}(k)$ and $\tilde{\gamma}_w^{(M)}(k)$ from (4.9) and (4.10).

3 Calculate $\hat{\boldsymbol{\zeta}}_0$ and Ω_0 from (3.18) and (3.19), with Σ_ζ and Σ_w replaced by $\tilde{\Sigma}_\zeta^{(M)}$ and $\tilde{\Sigma}_w^{(M)}$ respectively. Set $\hat{\boldsymbol{\eta}}_0 = \hat{\boldsymbol{\mu}} + \hat{\boldsymbol{\zeta}}_0$, and obtain $\hat{\boldsymbol{\eta}}_C$ and $\hat{\boldsymbol{\eta}}$ by (2.6) and (2.7).

For obtaining an estimate $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ from model (3.13) and (3.11) (and hence an estimate $\hat{\mu}_t = \mathbf{x}_t'\hat{\boldsymbol{\beta}}$ of μ_t follows), Jones (1980) suggests using the GLS estimator which is theoretically more efficient than the OLS estimator. However the unknown Σ_y^{-1} required in the GLS estimation procedure complicates the problem. In fact, the OLS estimate is simpler and consistent, moreover, it can be shown that, under certain conditions, for $\hat{\mu}_t$ obtained through the OLS avenue, the sample covariance of $\hat{w}_t = y_t - \hat{\mu}_t$ and that of $w_t = y_t - \mu_t$ are asymptotically the same (Chen and Ni, 1989, Theorem 4.1). That means if n is not too small, we can use $\hat{\mathbf{w}} = (\hat{w}_1 \ \ldots \ \hat{w}_n)'$ to replace the unknown $\mathbf{w}$ in estimating Σ_w and Σ_ζ without affecting the results very much. Our simulation in Section 5 confirms this conclusion even for n as small as 28 (compare the rows labelled "linear" with rows labelled "known" in Table 5.3).

4.2 Estimation procedure for the DS model

Now we consider the DS model where $\{\eta_t\}$ follows (3.2) with $d = 1$ and $\tilde{d} = 1$ for simplicity of notation. Denoting

$$w_t = \nabla\nabla_s\, y_t, \quad \dot{e}_t = \nabla\nabla_s\, e_t, \tag{4.11}$$

then the $\tilde{n} \times 1$ vector $\mathbf{w} = (w_{s+2} \ \dots \ w_n)'$ is obtained from $\mathbf{y} = (y_1, \dots, y_n)'$. From (1.1) and (3.2) we have

$$w_t = \zeta_t + \dot{e}_t, \tag{4.12}$$

where $\{\dot{e}_t\}$ is still stationary, but overdifferenced with known spectral density:

$$f_{\dot{e}}(\lambda) = \mid (1 - e^{i\lambda})(1 - e^{is\lambda}) \mid^2 f_e(\lambda). \tag{4.13}$$

From formulae as (4.4) through (4.10) with Σ_e and f_e replaced by $\Sigma_{\dot{e}}$ and $f_{\dot{e}}$, we obtain $\tilde{\gamma}_\zeta^{(M)}(k)$ and then $\tilde{\Sigma}_\zeta^{(M)}$.

From (3.9) with Σ_ζ replaced by $\hat{\Sigma}_\zeta^{(M)}$, an estimate of Ω_0 is obtained; then $\hat{\boldsymbol{\eta}}_0$ follows from (3.8); and $\hat{\boldsymbol{\eta}}$, from (2.6) and (2.7).

5 Simulation

5.1 TS models and the data

The simulation involves benchmarking quarterly series to annual benchmarks. The following two models are used to create ζ_t and e_t :

$$\begin{cases} (1 - 0.7B)\zeta_t = (1 - 0.4B)b_t, & \sigma_b = 5.0 \\ (1 - 0.5B)e_t = c_t, & \sigma_c = 2.5, \end{cases} \tag{5.1}$$

and

$$\begin{cases} (1 - 0.9B)\zeta_t = (1 - 0.6B^4)b_t, & \sigma_b = 5.0, \\ (1 - 0.7B + 0.49B^2)e_t = c_t, & \sigma_c = 4.0, \end{cases} \tag{5.2}$$

where $\{b_t\}$ and $\{c_t\}$ are independent Gaussian white noise with mean zero and standard deviations (SD) as indicated. We generate b_t and c_t by SAS, then ζ_t and e_t are obtained from the corresponding difference equations by recursion. For every set of ζ_t and e_t, the first 100 values were abandoned to eliminate the effect of initial values.

In the following, model 1 will refer to (5.1) and model 2 to (5.2). For model 1, both spectral densities $f_\zeta(\lambda)$ of $\{\zeta_t\}$ and $f_e(\lambda)$ of $\{e_t\}$ have very similar patterns (a peak at the origin then damping out on both sides). Consequently, any linear filter cannot extract ζ_t from w_t efficiently, so we do not expect a significant improvement over the regression method for this model.

For model 2, $f_\zeta(\lambda)$ has a sharp peak at the origin and has troughs at $\pm\pi$ and $\pm\pi/2$, while $f_e(\lambda)$ has peaks at $\pm\pi/3$. Consequently, the signal extraction procedure should bring in a remarkable improvement over the regression method for this model. Our simulation results will show such a difference for these two models.

The following model is used to create μ_t (refer(3.6) and (3.7))

$$\mu_t = 100 + t - 10\delta_{\{t,1\}} + 5\delta_{\{t,3\}} + 5\delta_{\{t,4\}}, \tag{5.3}$$

that is a linear trend plus a seasonal variation $\{-10, 0, 5, 5\}$.

Every data set is created for 7 years ($n = 28$). Then we have "data" $y_t = \mu_t + \zeta_t + e_t$ ($t = 1, \ldots, 28$). We keep all three underlying components of y_t in our records, so we know $\eta_t = \mu_t + \zeta_t$, and then the "data" $z_l = \eta_{4l-3} + \eta_{4l-2} + \eta_{4l-1} + \eta_{4l}$ ($l = 1, \ldots, 7$) (binding benchmarks), are created. Using only $z_1, \ldots, z_6$ as annual benchmarks (leaving the last year without benchmark, this could happen in practice because of reporting delay), we carried out benchmarking, 50 replications were done for each case (data are created by the same models with indicated parameters and using the same indicated method).

5.2 The mean in TS models is known

At first, we assume that the μ_t are known, so we know the true values of w_t which can be obtained from "data" y_t by $w_t = y_t - \mu_t$ ($t = 1, \ldots, 28$). Table 5.1 and 5.2 show *the "root mean square error"* (denoted by RMSE) both for the extracted signal $\hat{\eta}_0$ and the benchmarked value $\hat{\eta}$. For example, for $\hat{\eta}$, the RMSE (if over all 50 replications) is

$$\{\frac{1}{50 \times 28} \sum_{j=1}^{50} \sum_{t=1}^{28} (\hat{\eta}_t^{(j)} - \eta_t^{(j)})^2\}^{1/2}, \tag{5.4}$$

where j denotes the jth replication.

For a method, in jth replication, we would expect

$$\{\frac{1}{28} \sum_{t=1}^{28} (\hat{\eta}_t^{(j)} - \eta_t^{(j)})^2\}^{1/2} < \{\frac{1}{28} \sum_{t=1}^{28} (\hat{\eta}_{0t}^{(j)} - \eta_t^{(j)})^2\}^{1/2}. \tag{5.5}$$

Temporarily, both sides of (5.5) are called the standard deviation (SD) of $\hat{\eta}^{(j)}$ and $\hat{\eta}_0^{(j)}$ respectively, which means that, after benchmarking the error should be smaller than before. We say jth replication is valid if (5.5) holds. The reality is that, if a method is bad, the correction term $\hat{\eta}_C^{(j)}$ may even explode (beyond hundreds, thousands in our simulation), so a higher number (close or equal to 50) of valid replication may indicate the stability of a method. These numbers are in the columns "valid no." of Table 5.1 and 5.2. The RMSEs in "valid $\hat{\eta}_0$" and "valid $\hat{\eta}$" columns of these tables have similar definitions as (5.4) but the average is only over the valid replications. For all these methods, the true value of Σ_e is always given.

Method 1 refers to the regression method (1.11), according to (2.13), we put RMSE of $\mathbf{y}$ in the place of $\hat{\boldsymbol{\eta}}_0$ in Tables 5.1 and 5.2. In every case and all 50 replications, benchmarking reduces the SD, i.e., the SD of $\hat{\boldsymbol{\eta}}_R^{(j)}$ is always less than the SD of $\mathbf{y}^{(j)}$.

Method 2 refers to using Hillmer and Trabelsi's formulae (2.5) through (2.8) and assumes that $\Sigma_\eta = \Sigma_\zeta$ is known (so does $\Sigma_y = \Sigma_w$). We see that the signal extraction reduces RMSE of $\mathbf{y}$, on average, by $(2.8426 - 2.6137)/2.8426 = 8.1\%$ for model 1 and 25.1% for model 2 respectively. All 50 replications are valid. Comparing with method 1, the RMSEs of benchmarked values are reduced by 7.6% for model 1 and 29.2% for model 2 respectively.

Methods 3 through 10 are variants of method 2 where Σ_η and the corresponding Σ_w are replaced by different estimates. Method 3 uses (4.4) for $\gamma_w(k)$ and $\hat{\gamma}_\zeta(k) = \hat{\gamma}_w(k) - \gamma_e(k)$ for $\gamma_\zeta(k)$; method 4, 5 and 6 use Tukey-Hanning window in (4.6) for $\gamma_w(k)$ with $M = [n/3], [2n/3]$ and $n-1$ respectively ($[x]$ denotes the integer part of x); and, $\hat{\gamma}_\zeta^{(M)}(k) = \hat{\gamma}_w^{(M)}(k) - \gamma_e(k)$ for $\gamma_\zeta(k)$. We see that, none of methods 3 to 6 are satisfactory, only method 4 (corresponding to the smoothest spectral estimate) performs the best among them. In all these methods, the estimates of Σ_ζ may be not necessary positive definite as no revision procedure has been made.

Method 7 uses (4.8) to estimate $\gamma_\zeta(k)$; method 8, 9 and 10 use (4.9) to estimate $\gamma_\zeta(k)$, again with Tukey-Hanning window and $M = [n/3], [2n/3]$ and $n-1$ respectively. Then the corresponding estimates for $\gamma_w(k)$ are obtained as (4.10). From these results, we can see that methods 8 and 9 are quite satisfactory, and method 8 ($M = [n/3]$), the one corresponding to the smoothest spectral estimate is almost as good as method 2.

Defining the *signal to noise ratio* by

$$S/N = \gamma_\zeta(0)/\gamma_e(0), \tag{5.6}$$

then for model 1, $S/N = 29.41/8.33 = 3.53$; for model 2, $S/N = 75.53/27.02 = 2.79$. Now, in both models, we keep all the same parameters, except for σ_c. For model 1, σ_c is changed to 4.0 and 1.0; for model 2, σ_c is changed to 6.0 and 1.5, respectively. The corresponding S/N are listed under Table 5.3 (A larger σ_c makes S/N lower, a smaller σ_c makes S/N higher). By the same way as before, for each value of σ_c, we obtained a set of tables as Table 5.1 or 5.2, (not shown for space reason) again method 8 was the best among the methods 3 through 10.

In Table 5.3 (the rows of mean is "known"), we summarize the most important results: only the RMSEs of $\hat{\boldsymbol{\eta}}$ of method 8 are listed to compare with method 1 (the regression method) and method 2 (knowing the true value of Σ_η). The cells in

columns with heading "no." give the number of valid replications among 50 for each case. The "%" column shows the percentage reduction of the RMSE of the indicated methods to compare to method 1. For example, for model 1, low S/N, method 8, 47 replications are valid; the RMSE of is 2.7953, which reduces 3.2960 by 15.2%. The number 3.0213 in the brackets is the RMSE of all 50 replications, that means, in this case, even those 3 invalid replications are included for comparison, method 8 is still better than method 1.

From Table 5.3, we can see that the signal extraction benchmarking method performs more efficiently as the S/N goes down; for high S/N cases, using signal extraction does not bring in very much benefits (for model 1, it makes no difference with regression method).

5.3 The mean in TS models is unknown

We now turn to a situation closer to reality: where $\mu_t$is given by (5.3) but we do not know the parameters, so we get the LSE of the parameters from "data" $\mathbf{y}$. Using these estimates to replace the true values of the parameters in (5.3), $\hat{\mu}_t$ and $\hat{w}_t = y_t - \hat{\mu}_t$ are obtained. Then following Step 2 and Step 3 in Section 4, we carry out the same simulation as before, the main results are listed in Table 5.3 (the rows of mean is "linear").

Notice that, method 1 does not need any assumption about μ_t, so the result is the same as in the "known" part, which we do not repeat.

In this situation, although method 2 uses the true second moments Σ_ζ, Σ_w and hence Ω_0, but since $\hat{\mathbf{w}}$ replaces the true value of $\mathbf{w}$ in (3.19), method 2 is now not as good as in the situation of "mean is known", but only slightly worse.

For method 8, $\mathbf{w}$ and the second moments in the benchmarking formulae are all estimated, but the results show that this method is only marginally worse than method 2.

The most realistic situation is that we do not know the type of the function which μ_t follow. Assuming μ_t is a piece-wise smooth trend plus seasonality, trading-day effect and so on, we may at first look at that if there are abrupt changes in the trend and identify the changing points; then use a linear or quadratic function, or more generally, a spline to model the trend in each interval between consecutive changing points. For testing this approach, we designed the following simulation: μ_t are generated as (5.3) *plus an extra quadratic term* $0.04t^2$. The choice of the coefficient 0.04 makes this term contributing a variation as large as the linear term t in the range

of t from 1 to 28. Data were generated with this new model of the mean, but we still specified the mean μ_t as (5.3) and then applied the LSE. This mis-specification almost did not affect the results as shown in the "quad." part of Table 5.3 compared to those of the "linear" part (no model mis-specification, parameter estimation error only). In other words, the model specification of smooth trend is not so crucial. The reason is that, the signal extraction benchmarking method uses the correlation structures of the target series and the survey error series which are the features of "short-memory" of the series, while the selection of a smooth function mainly affects the features of "long-memory" of the series.

5.4 DS models

The target series $\{\eta_t\}$ is given by (3.2) with $d = 1$ and $s = 4$, $\{\zeta_t\}$ in (3.2) and the series of the survey error $\{e_t\}$ follows model

$$\begin{cases} \zeta_t = (1 - 0.5B)(1 - 0.6B^4)b_t, \\ (1 - 0.7B + 0.49B^2)e_t = c_t, \end{cases} \tag{5.7}$$

which will refer next as model 3. The properties and how to generate b_t and c_t, and then ζ_t and e_t, are stated in subsection 5.1. Thus, η_t may be obtained from ζ_t by recursion using (3.2) given 0 as initial values of η_t, and then we may have "data" $y_t = \eta_t + e_t$, but y_t has mean 0. Since "data" should represent more general and more practical situations, after the recursion step from ζ_t to η_t, we add μ_t as (5.3) to create final η_t (so, η_t have mean like (5.3)), but this time, we use 400 to replace 100 and use $2t$ to replace t in (5.3), since now $\{\eta_t\}$ is not stationary, η_t often runs wildly, this replacement keeps the data positive as they are in practice. Adding μ_t as (5.3) is also for confirming that (3.8), which excludes the involvement of non-zero $\boldsymbol{\mu}$ like (3.6), is a valid formula. The main purpose of the simulation aims to verify that the diffuse assumption for initial values which leads to $\Sigma_\eta^{-1} = D'\Sigma_\zeta^{-1}D$ (see (B.9) in Appendix) is reasonable.

In (5.7), we keep $\sigma_b = 5.0$ which gives $\gamma_\zeta(0) = 42.50$ (the variance of ζ_t), but take $\sigma_c = 5.0$, 3.0 and 1.0 consequently, then correspondingly, the signal to noise ratios S/N are 1.01, 2.80 and 25.16. Again $n = 28$ (7 years), $\mathbf{z}$ (benchmarks) is generated from η_t as in subsection 5.1 and we leave the last year without benchmarks. For each level of S/N, 50 replications have been done and the RMSEs of $\hat{\boldsymbol{\eta}}_0$ and $\hat{\boldsymbol{\eta}}$ defined as (5.4) for different methods are listed in Table 5.4.

We know already that for estimating Σ_ζ, the revised procedure is absolutely necessary and a good choice of M for smoothing is between $n/3$ and $2n/3$, so, we no longer investigate what happens if the spectral window estimation without revision and/or without proper smoothing, $M = [n/2]$ and Parzen window are chosen for smoothing. The results given in NP (stands for nonparametric) rows of Table 5.4 are obtained by using the estimation procedure described in subsection 4.3 for Σ_ζ. The numbers in the TRUE rows of Table 5.4 are results by using the true values of $\gamma_\zeta(k)$ for Σ_ζ. In the "$\hat{\gamma}_\zeta^{(M)}$" rows, the results are obtained by using smoothed (the same M and the same window) $\hat{\gamma}_\zeta(k) = (1/n)\sum_{t=1}^{n-k}\zeta_t\zeta_{t+k}$ (cf. (4.6)) to estimate $\gamma_\zeta(k)$ for Σ_ζ. In practice this is also impossible since ζ_t is unknown. Our intention of doing that is to give another standard for monitoring our NP procedure. We were expecting that using "$\hat{\gamma}_\zeta^{(M)}$" would be better than NP but less efficient than TRUE. To our surprise using $\hat{\gamma}_\zeta^{(M)}(k)$ is better than using the true $\gamma_\zeta(k)$ (perhaps this is due to adaptiveness). This outcome indicates further potential power of nonparametric approaches: more efficient estimates rather than one given by this paper may be developed in the future.

The numbers in REG row are the benchmarking results by using the regression method. As before, the percentages in other rows indicate the RMSE reduction to compare to the regression method. From Table 5.4, we have similar conclusions as in the case of TS models: benchmarking via signal extraction performs efficiently than benchmarking via regression when the S/N goes down.

For DS models, we always come across to the situation that the patterns of spectral densities of $\{\zeta_t\}$ and of $\{\dot{e}_t\}$ are quite different, since $\{\dot{e}_t\}$ is overdifferenced (see (4.11)) with spectral density (4.13), but $\{\zeta_t\}$ is not.

"$\gamma_e(0)^{1/2}$" row of Table 5.4 shows the theoretical standard deviations of e_t, while the numbers in " $\hat{\gamma}_e(0)^{1/2}$" row are the sample standard deviation of $e_t = y_t - \eta_t$, i.e., the RMSEs of y_t (in (5.4), $\hat{\eta}_t^{(j)}$ replaced by $y_t^{(j)}$). In comparing $\hat{\gamma}_e(0)^{1/2}$ with any other of the remaining rows, we see that either signal extraction or benchmarking dramatically reduces the survey error when S/N is low.

Using

$$\frac{1}{50 \times 28}\sum_{j=1}^{50}\sum_{t=1}^{28}(\hat{\eta}_t^{(j)} - \eta_t^{(j)}), \tag{5.8}$$

and a similar formula for $\hat{\eta}_{0t}^{(j)}$ to indicate the bias of an estimation method, our simulation shows that, smoothing by signal extraction does not reduce the bias in all cases, but benchmarking does. Table 5.5 shows the results. In this table, the last row

is the sample bias of $e_t = y_t - \eta_t$. We have similar results for TS models but are not shown in this paper.

TABLE 5.1 RMSRs of $\hat{\eta}_0$ and $\hat{\eta}$ for model 1

Method	valid $\hat{\eta}_0$	valid $\hat{\eta}$	valid no.	all $\hat{\eta}_0$	all $\hat{\eta}$
1	2.8426	2.0600	50	2.8426	2.0600
2	2.6137	1.9082	50	2.6137	1.9082
3	2.8599	2.0835	40	2.8941	10.8079
4	2.6668	2.0052	48	2.6984	5.0063
5	2.6709	1.9812	41	2.7717	2.5804
6	2.7906	2.0920	46	2.8129	3.6217
7	2.6913	2.0844	46	2.7084	409.9062
8	2.6389	1.9423	49	2.6366	2.0743
9	2.6482	1.9961	48	2.6646	2.7197
10	2.6664	2.0206	47	2.6819	3.5931

TABLE 5.2 RMSEs of $\hat{\eta}_0$ and $\hat{\eta}$ for model 2

Method	valid $\hat{\eta}_0$	valid $\hat{\eta}$	valid no.	all $\hat{\eta}_0$	all $\hat{\eta}$
1	5.2062	4.1746	50	5.2062	4.1746
2	3.9011	2.9569	50	3.9011	2.9569
3	4.6293	3.6423	30	4.8124	43.3872
4	4.2195	3.3336	44	4.3006	17.4169
5	4.3204	3.3244	35	4.5012	45.9034
6	4.4382	3.4146	32	4.6059	2793.1038
7	4.3216	3.3495	48	4.3359	3.8291
8	4.1426	3.1640	50	4.1426	3.1640
9	4.2148	3.2077	49	4.2262	3.4747
10	4.2654	3.2590	49	4.2778	3.6134

TABLE 5.3 Some RMSEs of $\hat{\eta}$ for models 1 and 2

			model 1			model 2		
mean	S/N	meth.	RMSE	%	no.	RMSE	%	no.
known	low	1	3.2960	0	50	6.2616	0	50
		2	2.6899	18.5	50	3.6226	42.1	50
		8	2.7953 (3.0213)	15.2	47	4.0626 (4.0861)	35.1	49
	mid	1	2.0600	0	50	4.1746	0	50
		2	1.9082	7.6	50	2.9569	29.2	50
		8	1.9423 (2.0743)	5.7	49	3.1640	24.2	50
	high	1	0.8240	0	50	1.5655	0	50
		2	0.8207	0.4	50	1.4588	7.5	50
		8	0.8253	-0.2	50	1.4797	5.5	50
linear	low	2	2.7734	15.9	50	4.0791	34.5	50
		8	2.9034 (5.2093)	11.9	43	4.3497 (5.0732)	30.5	49
	mid	2	1.9279	6.4	50	3.1912	23.6	50
		8	1.9904 (2.1800)	3.4	49	3.3085 (3.3964)	20.7	49
	high	2	0.8212	0.3	50	1.4921	4.7	50
		8	0.8203	0.5	50	1.5033	4.0	50
quad.	low	2	2.7981	15.3	50	4.0888	34.7	50
		8	2.9554 (3.5026)	10.3	46	4.3944 (5.5082)	29.8	49
	mid	2	1.9325	6.3	50	3.1924	23.5	50
		8	1.9878 (2.0617)	3.5	49	3.3274 (3.9527)	20.5	49
	high	2	0.8212	0.3	50	1.4913	4.7	50
		8	0.8232	0.0	50	1.5059	3.8	50

Note: S/N, the signal to noise ratio, for the two models are as follows

	model 1	model 2
low	1.38	1.24
mid	3.53	2.79
high	22.11	19.83

TABLE 5.4 RMSEs of $\hat{\eta}_0$ and $\hat{\eta}$ for model 3

S/N		low (1.01)	mid.(2.80)	high (25.16)
$\hat{\eta}_0$	TRUE	3.9496	2.8154	1.1941
	$\hat{\gamma}_\zeta^{(M)}$	3.9041	2.7453	1.1675
	NP	4.3616	2.9216	1.2002
$\hat{\eta}$	REG	5.1347	3.0808	1.0268
	TRUE	3.0370 (40.9%)	2.2015 (28.5%))	0.9529 (7.2%)
	$\hat{\gamma}_\zeta^{(M)}$	3.0159 (41.3%)	2.1480 (30.3%))	0.9345 (9.0%)
	NP	3.4170 (33.5%)	2.3464 (23.8%))	0.9670 (5.8%)
$\hat{\gamma}_e(0)^{1/2}$		6.34	3.80	1.27
$\gamma_e(0)^{1/2}$		6.50	3.90	1.30

TABLE 5.5 Bias of $\hat{\eta}_0$ and $\hat{\eta}$ for model 3

S/N		low (1.01)	mid.(2.80)	high (25.16)
$\hat{\eta}_0$	TRUE	-0.2975	-0.1819	-0.0623
	$\hat{\gamma}_\zeta^{(M)}$	-0.3002	-0.1829	-0.0624
	NP	-0.3054	-0.1862	-0.0633
$\hat{\eta}$	REG	-0.0420	-0.0252	-0.0084
	TRUE	-0.0265	-0.0164	-0.0070
	$\hat{\gamma}_\zeta^{(M)}$	-0.0318	-0.0189	-0.0071
	NP	-0.0764	-0.0478	-0.0102
e_t		-0.3126	-0.1876	-0.0625

6 Conclusions

In this paper we analyzed the properties of the regression and the signal extraction methods for benchmarking repeated survey data (the signal extraction may also be used for the purpose of smoothing only) and proved that the corrections originating from the benchmarks reduce the variance of the survey error and the error of the signal extraction, respectively. Furthermore, when the first and the second order moments of the survey error $\{e_t\}$ and the target series $\{\eta_t\}$ are known, the variance of the error of the benchmarked estimate via signal extraction is smaller than that of using the

regression method because the regression method does not use the information on the nature of $\{\eta_t\}$ given by its covariance matrix Σ_η.

We also discussed the limitations of the ARIMA model based approach in implementing the signal extraction benchmarking. In order to overcome some difficulties associated with the ARIMA modelling of $\{\eta_t\}$, we proposed a nonparametric approach based on spectral estimation techniques and derived the formulae for two types of models, the TS models and the DS models. The DS models are more general than ARIMA models. A TS model for $\{\eta_t\}$ is assumed to have a deterministic component $\{\mu_t\}$ which reflects the trend (allowed to have abrupt changes), stable seasonality, outliers, trading-day effects and so on. The stochastic component $\{\zeta_t\}$ of a TS model is assumed to be stationary. For DS models, we assume that the differenced series $\{\zeta_t\}$ of $\{\eta_t\}$ is stationary and allow a hidden polynomial plus stable seasonality to exist in $\{\eta_t\}$. For both TS and DS models, the stationarity of $\{\zeta_t\}$ is in wide sense, no distribution specification, e.g. normality and no parametric model assumption such as ARMA are required.

We carried out a simulation to compare the efficiency of the regression method (RM), the signal extraction (SE) assuming Σ_η is known (not available in real life cases) and the signal extraction method where Σ_η is estimated by the nonparametric (NP) approach proposed in this study.

In the simulation, for a TS model $\{\mu_t\}$, is given as a linear or a quadratic trend plus a stable seasonality, and $\{\zeta_t\}$ and $\{e_t\}$ are created following ARMA models. For a DS model, $\{\eta_t\}$ is obtained by recursion from a stationary $\{\zeta_t\}$ then adding a linear trend plus a stable seasonality. For all these models, the effects of the levels of the signal to noise ratio are considered. For the TS type, two models (1 and 2) are chosen to distinguish two different situations: the spectral densities $f_\zeta(\lambda)$ and $f_e(\lambda)$, of $\{\zeta_t\}$ and $\{e_t\}$ respectively, have similar or very different patterns. The results show:

(1) A smooth spectral estimate, from which we may obtain an estimate of the covariance matrix Σ_ζ, is needed, and this estimate of Σ_ζ must be positive definite for finally obtaining an estimate of $\boldsymbol{\eta}$ close to that by using the true value of Σ_ζ. The proposed NP approach satisfies these requirements.

(2) Using the RMSE as a measure of efficiency, SE is always better than RM, NP is close to SE and better than RM.

(3) If the signal to noise ratio $S/N = \gamma_\zeta(0)/\gamma_e(0)$ is small, the gains in efficiency of SE and NP to compare with RM are large. When the S/N is very high, there is almost no lost in efficiency by using simpler RM rather than NP in practice.

(4) For TS models, if the patterns of $f_\eta(\lambda)$ and $f_e(\lambda)$ are similar, the gains in

efficiency of SE and NP with respect to RM are small; and if the difference between these two patterns is big, then the gains are larger.

(5) For TS models, the only parametric model used is to estimate μ_t. The NP method is robust to smooth trend mis-specifications and to the estimation error of the parameters.

Appendix A. A derivation of the signal extraction benchmarking formulae via projection

Lemma Denote the projection of $\boldsymbol{\eta}$ on the linear probability space generated by $(\mathbf{y}'\ \mathbf{z}')'$ by $\hat{E}(\boldsymbol{\eta} \mid \mathbf{y}, \mathbf{z})$. Similar notation for $\hat{E}(\boldsymbol{\eta} \mid \mathbf{y})$ and so on. Then

$$\hat{E}(\boldsymbol{\eta} \mid \mathbf{y}, \mathbf{z}) = \hat{E}(\boldsymbol{\eta} \mid \mathbf{y}) + \hat{E}[\boldsymbol{\eta} \mid \mathbf{z} - \hat{E}(\mathbf{z} \mid \mathbf{y})] - E\boldsymbol{\eta} \tag{A.1}$$

Proof For simplicity of notation, we will denote $\Sigma_y = \text{Var } \mathbf{y}$, $\Sigma_{zy} = \text{Cov } (\mathbf{z}, \mathbf{y})$ and so on. The well-known general formula of projection (see e.g. Whittle, 1963, Section 4.3; Brillinger, 1981, Section 8.2) $\mathbf{z}$ on $\mathbf{y}$, is

$$\hat{E}(\mathbf{z} \mid \mathbf{y}) = E\mathbf{z} + \Sigma_{zy}\Sigma_y^{-1}(\mathbf{y} - E\mathbf{y}). \tag{A.2}$$

Denote

$$\mathbf{v} = \mathbf{z} - \hat{E}(\mathbf{z} \mid \mathbf{y}) = (\mathbf{z} - E\mathbf{z}) - \Sigma_{zy}\Sigma_y^{-1}(\mathbf{y} - E\mathbf{y}), \tag{A.3}$$

then

$$E\mathbf{v} = 0, \quad E[(\mathbf{y} - E\mathbf{y})\mathbf{v}'] = \Sigma_{yz} - \Sigma_y\Sigma_y^{-1}\Sigma_{yz} = 0.$$

Since $(\mathbf{y}'\ \mathbf{z}')'$ and $(\mathbf{y}'\ \mathbf{v}')'$ generate the same linear probability space, so

$$\begin{aligned} \hat{\boldsymbol{\eta}} &= \hat{E}(\boldsymbol{\eta} \mid \mathbf{y}, \mathbf{z}) = \hat{E}(\boldsymbol{\eta} \mid \mathbf{y}, \mathbf{v}) \\ &= E\boldsymbol{\eta} + (\Sigma_{\eta y}\ \Sigma_{\eta v}) \begin{pmatrix} \Sigma_y & 0 \\ 0 & \Sigma_v \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{y} - E\mathbf{y} \\ \mathbf{v} \end{pmatrix}, \end{aligned} \tag{A.4}$$

which is the right hand side of (A.1). The lemma is proved.

Now we prove the same formula as (2.5) through (2.9) for model (1.4) and (1.5) under the meaning of projection. The distribution assumption is (2.10). Again, formula (A.2) is repeatedly used in the following but $\mathbf{y}$ and $\mathbf{z}$ in (A.2) may be replaced by other vectors. Also, the general formula for the variance matrix of projection error, for example that of $\boldsymbol{\eta}$ on $\mathbf{y}$, is

$$\text{Var } [\hat{E}(\boldsymbol{\eta} \mid \mathbf{y}) - \boldsymbol{\eta}] = \Sigma_\eta - \Sigma_{\eta y}\Sigma_y^{-1}\Sigma_{y\eta}, \tag{A.5}$$

will also be used.

From (1.4), $\Sigma_{\eta y} = \Sigma_\eta = \Sigma_{y\eta}$ and $\Sigma_y = \Sigma_\eta + \Sigma_e$, so (A.5) gives

$$\text{Var}\ (\hat{\boldsymbol{\eta}}_0 - \boldsymbol{\eta}) = \Sigma_\eta[I_n - (\Sigma_\eta + \Sigma_e)^{-1}\Sigma_\eta] = (\Sigma_\eta^{-1} + \Sigma_e^{-1})^{-1} = \Omega_0,$$

that is (2.8). Here and in the following, the identities $I_n = (\Sigma_\eta + \Sigma_e)^{-1}\ (\Sigma_\eta + \Sigma_e)$ and

$$\Sigma_\eta(\Sigma_\eta + \Sigma_e)^{-1}\Sigma_e = \Sigma_e(\Sigma_\eta + \Sigma_e)^{-1}\Sigma_\eta = (\Sigma_\eta^{-1} + \Sigma_e^{-1})^{-1} = \Omega_0 \qquad (A.6)$$

are very useful.

Using the general formula (A.2) for $\hat{E}(\boldsymbol{\eta} \mid \mathbf{y})$, (1.4) and (2.4) lead to

$$\hat{\boldsymbol{\eta}}_0 = \boldsymbol{\mu} + \Sigma_\eta(\Sigma_\eta + \Sigma_e)^{-1}(\mathbf{y} - \boldsymbol{\mu}) = \Omega_0(\Sigma_e^{-1}\mathbf{y} + \Sigma_\eta^{-1}\boldsymbol{\mu}), \qquad (A.7)$$

that is (2.5). From (1.4), (1.5), (2.4) and (A.3),

$$\mathbf{v} = (\mathbf{z} - L\boldsymbol{\mu}) - L\Sigma_\eta(\Sigma_\eta + \Sigma_e)^{-1}(\mathbf{y} - \boldsymbol{\mu}) = \mathbf{z} - L\hat{\boldsymbol{\eta}}_0. \qquad (A.8)$$

The middle expression of (A.8) also leads to

$$\mathbf{v} = L(\boldsymbol{\eta} - \boldsymbol{\mu}) + \boldsymbol{\epsilon} - L\Sigma_\eta(\Sigma_\eta + \Sigma_e)^{-1}[(\boldsymbol{\eta} - \boldsymbol{\mu}) + \mathbf{e}], \qquad (A.9)$$

which gives

$$\Sigma_{\eta v} = \Sigma_\eta L' - \Sigma_\eta(\Sigma_\eta + \Sigma_e)^{-1}\Sigma_\eta L' = \Omega_0 L', \qquad (A.10)$$

$$\Sigma_v = L\Sigma_\eta L' + \Sigma_\epsilon - L\Sigma_\eta(\Sigma_\eta + \Sigma_e)^{-1}\Sigma_\eta L' = L\Omega_0 L' + \Sigma_\epsilon, \qquad (A.11)$$

and hence

$$\hat{E}(\boldsymbol{\eta} \mid \mathbf{v}) = E\boldsymbol{\eta} + \Omega_0 L'(L\Omega_0 L' + \Sigma_\epsilon)^{-1}(\mathbf{z} - L\hat{\boldsymbol{\eta}}_0). \qquad (A.12)$$

Using (A.1), (2.6) and (2.7) follow.

Using the general formula (A.5) for $\boldsymbol{\eta}$ on $(\mathbf{y}'\ \mathbf{z}')'$, and referring (A.4), then (A.10) and (A.11) gives

$$\text{Var}\ (\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}) = \Sigma_\eta - (\Sigma_{\eta y}\ \Sigma_{\eta v}) \begin{pmatrix} \Sigma_y & 0 \\ 0 & \Sigma_v \end{pmatrix}^{-1} \begin{pmatrix} \Sigma_{y\eta} \\ \Sigma_{yv} \end{pmatrix}$$

$$= \Sigma_\eta - \Sigma_\eta(\Sigma_\eta + \Sigma_e)^{-1}\Sigma_\eta - \Omega_0 L'(L\Omega_0 L' + \Sigma_\epsilon)L\Omega_0$$

and the last expression in (2.9) follows immediately by identities (A.6) and the one above (A.6).

Appendix B. The derivation of the special signal extraction formulae for the DS model

Put

$$J = \begin{pmatrix} I_{d+\tilde{d}s} & 0 & \cdots & \cdots & 0 \\ & & D & & \end{pmatrix}. \tag{B.1}$$

then

$$\zeta^* = \begin{pmatrix} \eta_i \\ \zeta \end{pmatrix} = \begin{pmatrix} \eta_i \\ D\eta \end{pmatrix} = J\eta, \tag{B.2}$$

and hence, $\Sigma_{\zeta^*} = J\Sigma_\eta J'$ (at this moment, in (3.3), h is still a fixed large number), or

$$\Sigma_\eta = J^{-1}\Sigma_{\zeta^*}(J')^{-1}, \quad \Sigma_\eta^{-1} = J'\Sigma_{\zeta^*}^{-1}J. \tag{B.3}$$

From (A.7) and referring (A.6), we have

$$\hat{\eta}_0 = (\Sigma_e^{-1} + \Sigma_\eta^{-1})^{-1}\Sigma_e^{-1}\mathbf{y} + [I_n - (I_n + \Sigma_e\Sigma_\eta^{-1})^{-1}]\mu. \tag{B.4}$$

Now, the $\mathbf{y}$ term already has the form of (3.8) and (3.9) except that Σ_η^{-1} has not been specified yet.

Using $I_n = (I_n + \Sigma_e\Sigma_\eta^{-1})^{-1}(I_n + \Sigma_e\Sigma_\eta^{-1})$, the μ term in (B.4) is

$$(I_n + \Sigma_e\Sigma_\eta^{-1})^{-1}\Sigma_e\Sigma_\eta^{-1}\mu = (I_n + \Sigma_e\Sigma_\eta^{-1})^{-1}\Sigma_e J'\Sigma_{\zeta^*}^{-1}J(E\eta). \tag{B.5}$$

From (3.3) and (3.4) and letting $h \to \infty$ yields

$$\Sigma_{\zeta^*}^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & \Sigma_\zeta^{-1} \end{pmatrix}. \tag{B.6}$$

On the other hand, notice that $\mu_t = E\eta_t$ is a function as (3.6) and $E\zeta_t = 0$, then $E(D\eta) = D\mu + E\zeta = 0$ and from (B.1),

$$J(E\eta) = E(J\eta) = (E\eta_i' \;\; 0 \;\; \ldots \;\; 0)'. \tag{B.7}$$

Then (B.6) and (B.7) gives

$$\Sigma_{\zeta^*}^{-1}J(E\eta) = 0. \tag{B.8}$$

i.e., (B.5) is zero.

From (B.1), (B.3) and (B.6),

$$\Sigma_\eta^{-1} = D'\Sigma_\zeta^{-1}D, \tag{B.9}$$

that completes the derivation of (3.8) and (3.9).

Acknowledgement

The authors thank K. Chiu, M. Morry, B. Quenneville, P. Wong and Prof. J.Durbin for stimulating discussion and support in the simulation.

References

Bell, W.R. and Hillmer, S.C. (1983). Modelling time series with calendar variation. *Journal of the American Statistical Association*, **78**, 526-534.

Box, G.E.P. and Tiao, G.C. (1975). Intervention analysis with application to economic and environmental problems. *Journal of the American Statistical Association*, **70**, 70-79.

Brillinger, D.R. (1981). *Time Series: Data analysis and theory.* Holden-Day, San Francisco.

Chanda, K.C. (1995). Large sample analysis of autoregressive moving-average models with errors in variables. *Journal of Time Series Analysis*, **16**, 1-16.

Chen, Z.G. and Ni, J.Y. (1989). Subset regression time series and its modelling procedures. *Journal of Multivariate Analysis*, **31**, 266-288.

Cholette, P.A. and Dagum, E.B. (1994). Benchmarking time series with autocorrelated survey errors. *International Statistical Review*, **62**, 365-377.

Cleveland, W.P. and Tiao, G.C. (1976). Decomposition of seasonal time series: a model for the census X-11 program. *Journal of American Statistical Association*, **71**, 581-587.

Dagum, E.B., Quenneville, B. and Sutradhar, B. (1992). Trading-day variations multiple regression models with random parameters. *International Statistical Review* **60**, 55-73.

Denton, F.T. (1971). Adjustment of monthly or quarterly series to annual totals: an approach based on quadratic minimization. *Journal of American Statistical Association*, **66**, 99-102.

Doob, J.L. (1953). *Stochastic Processes.* John Wiley & Sons, Inc. New York, London.

Hillmer, S.C. and Trabelsi, A. (1987). Benchmarking of economic time series. *Journal of the American Statistical Association*, **82**, 1064-1071.

Jones, R.G. (1980). Best linear unbiased estimators for repeated surveys. *Journal of Royal Statistical Society*, Series **B**, **42**, 221-226.

Lee, H. (1990). Estimation of panel correlation for the Canadian labour force survey. *Survey Methodology*, **16**, 283-292.

McLeod, A.I. (1975). Derivation of theoretical autocovariance function of autoregressive-moving average time series. *Applied Statistics*, **24**, 255-256.

Ca OOS

Nelson, C.R. and Plosser, C.I. (1982). Trends and random walks in macroeconomic time series. *Journal of Monetary Economics* **10**, 139-162.

Pierce, D.A. (1978). Seasonal adjustment where both deterministic and and stochastic seasonality are present. In *Seasonal Analysis of Economic Time Series.* (Arnold Zellner, Editor) 242-264.

Priestley, M.B. (1981). *Spectral Analysis and Time Series.* Vol.1, Academic Press. London, New York.

Scott, A.J., Smith, T.M.F. and Jones, R.G. (1977). The application of time series methods to the analysis of repeated surveys. *International Statistical Review*, **45**, 13-28.

Tsay, R.S. (1986). Time series model specification in the presence of outliers. *Journal of the American Statistical Association*, **81**, 132-141.

Whittle, P. (1963). *Prediction and Regulation by Linear Least-Square Methods.* University of Minnesota Press. Minneapolis, Minnesota.

STATISTICS CANADA LIBRARY
BIBLIOTHEQUE STATISTIQUE CANADA
1010236288