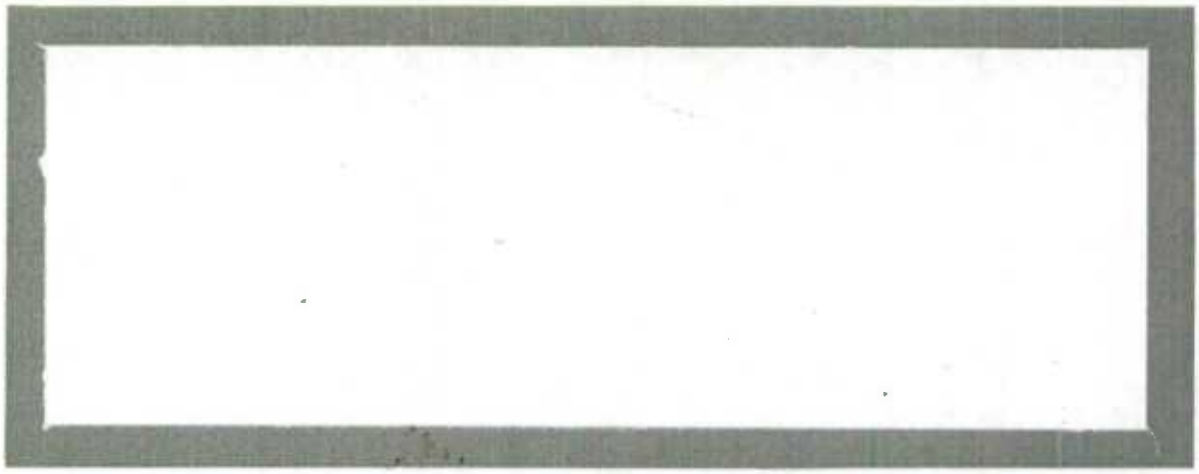
Statistics Canada
Statistique Canada



Methodology Branch

Business Survey Methods Division

Direction de la méthodologie

Division des méthodes d'enquêtes entreprises

Canada

WORKING PAPER NO. BSMD-93-009E

METHODOLOGY BRANCH

CAHIER DE TRAVAIL NO. BSMD-93-009E

DIRECTION DE LA MÉTHODOLOGIE

A NONPARAMETRIC TIME SERIES APPROACH FOR BENCHMARKING SURVEY DATA

by

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

July 1993

A Nonparametric Time Series Approach for Benchmarking Survey Data [1]

by

Zhao-Guo Chen, Pierre A. Cholette
and Estela Bee Dagum

July 1993

Time Series Research and Analysis Center
Business Survey Methods Division
Statistics Canada
3H, R.H.Coats bldg, Ottawa, Canada K1A 0T6

[1] For presentation to the annual meeting of American Statistical Association, Business and Economics Section, San Francisco, August 8-12, 1993

Abstract

This paper compares the errors of two benchmarking methods, one based on regression by Cholette and Dagum (1989, 1991) and the other based on signal extraction by Hillmer and Trabelsi (1987). The difficulties of finding an ARIMA model in the later method and the resulting implementation problems are discussed. A non-parametric approach to circumvent this problem is proposed. Simulations show the proposed approach is efficient, robust and easy to carry out.

Resume

Cet article compare les erreurs d'estimation de deux méthodes d'étalonnage, l'une fondée sur la régression de Cholette et Dagum (1989, 1991) et l'autre fondée sur l'extraction de signal de Hillmer et Trabelsi (1987). Les difficultés de trouver un modèle ARIMA dans la dernière méthode et les problèmes résultants de mise en oeuvre sont examinées. Une méthode non paramétrique, qui contourne ce problème, est proposée. Les simulations démontrent que l'approche proposée est efficiente, robuste et facile à appliquer.

1 Introduction

Benchmarking is an important problem faced by statistical agencies. For a target socio-economic variable η, two sources of data with different frequencies and precisions are available. For example, one set of data is monthly (or quarterly)

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

and the other one 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 in year i. Usually, e_t are much larger than ϵ_i. z_i are referred to as benchmarks. A benchmarking procedure estimates η_t, $t = 1, \ldots, n$ using these y_t and z_i. When all $\epsilon_i \equiv 0$, then the problem becomes estimating η_t from (1.1) under the restriction $\sum_{t \in i} \eta_t = z_i$ and the benchmarks are called binding; and non-binding otherwise.

For index and stock series, (1.2) becomes (in case of monthly data)

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

and (in case of benchmarks pertaining to December)

$$z_i = \eta_{12i} + \epsilon_k \tag{1.4}$$

respectively. In general, we may put (1.1) with anyone of (1.2) through (1.4) in the matrix form

$$\mathbf{y} = \boldsymbol{\eta} + \mathbf{e}, \tag{1.5}$$

$$\mathbf{z} = L\boldsymbol{\eta} + \boldsymbol{\epsilon}, \tag{1.6}$$

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. 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.

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

$$P_A(\mathbf{y}, \boldsymbol{\eta}) = (\mathbf{y} - \boldsymbol{\eta})' A (\mathbf{y} - \boldsymbol{\eta}) \tag{1.7}$$

under the restriction (1.6) with $\epsilon \equiv 0$ (binding benchmarks). By suitably choosing the symmetric $n \times n$ matrix A, one may obtain a benchmarked value $\hat{\eta}$ of η with some good properties, such as continuity between consecutive years. This type of benchmarking is easy to apply, and almost no preliminary statistical information is required. But, because no information on the natuture of the time series is used (when available), the estimation error, i.e., $E\{(\hat{\eta} - \eta)(\hat{\eta} - \eta)\}'$ is not minimized.

Cholette and Dagum (1989, 1991) combine (1.5) and (1.6) to the following regression model

$$\tau = X\eta + \mathbf{u}, \tag{1.8}$$

assuming that η is the vector of nukown parameters in this model. Where

$$\tau = \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix}, \qquad X = \begin{pmatrix} I_n \\ L \end{pmatrix}, \qquad \mathbf{u} = \begin{pmatrix} \mathbf{e} \\ \epsilon \end{pmatrix} \sim (\mathbf{0}, \Sigma_u), \tag{1.9}$$

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

$$\Sigma_u = \begin{pmatrix} \Sigma_e & \mathbf{0} \\ \mathbf{0} & \Sigma_\epsilon \end{pmatrix}. \tag{1.10}$$

In fact, Cholette and Dagum also allowed a bias in $\mathbf{e}$, however, since it can be removed preliminarily, we will not discuss it here.

Given Σ_u, we may use the generalized least squares estimate of η (the uniformly minimum variance unbiased estimate among all the linear unbiased estimates) as the benchmarked estimate:

$$\hat{\eta}_R = (X'\Sigma_u^{-1}X)^{-1}X'\Sigma_u^{-1}\tau = \mathbf{y} + \Sigma_e L'(L\Sigma_e L' + \Sigma_\epsilon)^{-1}(\mathbf{z} - L\mathbf{y}), \tag{1.11}$$

which has covariance matrix (since $\hat{\eta}_R$ is unbiased, so it is also the covarince matrix of the estimation error):

$$\Omega_R = (X'\Sigma_u^{-1}X)^{-1} = (\Sigma_e^{-1} + L'\Sigma_\epsilon^{-1}L)^{-1}. \tag{1.12}$$

The subscript R stands for "regression". This is a regression model, where the number of parameters (n) is so close to the number of observations ($n + m$). So, one should be careful when trying to apply a classical theorem of regression estimates on this model, and the residual vector

$$\hat{\mathbf{u}} = \tau - X\hat{\eta}_R \tag{1.13}$$

has properties different from those of the error vector $\mathbf{u}$.

Hillmer and Trabelsi (1987) considered that, η_t follow an ARIMA model, and then under the assumption of normal distribution, they derived their benchmarking formulae.

This paper will discuss the implementation of Hillmer and Trabelsi's formulae and proposed a nonparametric procedure which does not use ARIMA modelling for $\{\eta_t\}$. Simulation results are presented.

2 Benchmarking with Signal Extraction

2.1 The models and the formulae

Hillmer and Trabelsi (1987) assumed that η_t, the target series, follow an ARIMA model

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

while e_t, the survey error, follow 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 uncorrelated, then $y_t = \eta_t + e_t$ also follow an ARIMA model

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

In fact, a more important contribution of Hillmer and Trabelsi (1987) is deriving the following benchmarking formulae whether $\boldsymbol{\eta}$ and $\mathbf{e}$ are model based [say, as (2.1) - (2.3)] or not. That is under the normality assumption

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

knowing all the moments in (2.4), they obtained the conditional expectation of $\boldsymbol{\eta}$ given $\boldsymbol{\tau} = (\mathbf{y}'\ \mathbf{z}')'$ (which is the best unbiased estimate of $\boldsymbol{\eta}$):

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

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

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

where,

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

The covariance matrix of the benchmarked estimate $\hat{\boldsymbol{\eta}}$ of $\boldsymbol{\eta}$ is (since $\hat{\boldsymbol{\eta}}$ is unbiased, so it is also the the covariance matrix of the estimation error)

$$\Omega = (\Omega_0^{-1} + L'\Sigma_\epsilon^{-1}L)^{-1} = (\Sigma_e^{-1} + \Sigma_\eta^{-1} + L'\Sigma_\epsilon^{-1}L)^{-1}. \tag{2.9}$$

$\hat{\boldsymbol{\eta}}_0$ is the signal extraction, which is the best unbiased estimate of $\boldsymbol{\eta}$ given $\mathbf{y}$*only*. $\hat{\boldsymbol{\eta}}_C$ is the correction originating from the benchmarks $\mathbf{z}$.

The normality assumption (2.4) can be generalized to [refer to the statement under (1.9)]

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

then instead of the conditional expectation, using the projection on probability spaces, i.e., $\hat{\boldsymbol{\eta}}_0 = \hat{E}(\boldsymbol{\eta} \mid \mathbf{y})$ and $\hat{\boldsymbol{\eta}} = \hat{E}(\boldsymbol{\eta} \mid \boldsymbol{\tau})$ (the best unbiased *linear* estimate), the same formulae as (2.5) through (2.9) follow (see Appendix).

2.2 Properties and comparisons

Lemma 2.1 Suppose A and B are symmetric matrices of the same dimension, $A > 0$, $B \geq 0$, then

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

Proof There is a non-singular P such that $P'AP = I, P'BP = \Lambda$, where Λ is diagonal with all its elements non-negative, then the lemma follows easily. □

We now compare the signal extraction method and the regression method. Observing (1.11), (1.12) and (2.5) through (2.9), and in (1.11) denoting the last term

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

we have the following correspondence

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

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

In view of (1.12), (2.9) and Lemma 2.1, we have

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

The inequalities in (2.15) mean 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 that the variance of the error of the benchmarked estimate via the signal extraction method (Ω) is smaller than that of the sigal extraction (Ω_0). So, *the corrections originating from the benchmarks reduce the survey error or the signal extraction error in both methods.*

In view of (2.8), (2.9) with (1.12) and Lemma 2.1, we have

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

This means that, the variance of the error of the signal extraction 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 moments of both survey error and target series are known, then the signal extraction method is better than the regression method.* The reason is that, the regression method does not use Σ_η, the information about $\{\eta_t\}$.

Note that, both (1.12) and (2.9) can be rewritten to allow $\Sigma_\epsilon = 0$, i.e., binding benchmark. Using the formula in the partitioned inverse of a matrix, we have

$$\Omega_R = (\Sigma_e^{-1} + L'\Sigma_\epsilon^{-1}L)^{-1} = \Sigma_e - \Sigma_e L'(L\Sigma_e L' + \Sigma_\epsilon)^{-1}L\Sigma_e, \tag{2.17}$$

$$\Omega = (\Omega_0^{-1} + L'\Sigma_\epsilon^{-1}L)^{-1} = \Omega_0 - \Omega_0 L'(L\Omega_0 L + \Sigma_\epsilon)^{-1}L\Omega_0. \tag{2.18}$$

3 Implementation of benchmarking via model based signal extraction

Often, survey experts can offer a stationary model as (2.2) for survey error e_t and hence we may obtain the autocovariance function (ACVF) of $\{e_t\}$ ($Ee_t = 0$),

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

Then the Toeplitz matrix Σ_e is obtained by setting $\gamma_e(\mid i-j \mid)$ as its (i,j)th entry and the benchmarking procedure based on the regression model can be carried out without

difficulty. However, for the method via the signal extraction, we need $\boldsymbol{\mu} = E\boldsymbol{\eta} = E\mathbf{y}$ and Ω_0 (or equivalently, Σ_η), both of which are usually unknown.

Hillmer and Trabelsi (1987) suggested using the ARIMA models [see (2.1) - (2.3)] in dealing with $\boldsymbol{\mu}$ and Σ_η. First, consider $\boldsymbol{\mu}$. They suggested that $\boldsymbol{\mu} = \mu(1 \cdots 1)'$ if $\{\eta_t\}$ is stationary; and $\boldsymbol{\mu} = \mathbf{0}$ if nonstationary. However, from (2.5) one sees that a non-zero $\boldsymbol{\mu}$ may contribute a great deal to $\hat{\boldsymbol{\eta}}_0$. Suppose, say,

$$\phi_\eta(B) = \nabla^d \nabla_{12} \phi_\eta^*(B), \quad d > 0, \tag{3.2}$$

where $\nabla = 1 - B$, $\nabla_{12} = 1 - B^{12}$, $\phi_\eta^*(B)$ has its roots all outside the unit circle. When η_t satisfy (2.1), then $\eta_t - \tilde{\mu}_t$ also satisfy (2.1) if

$$\tilde{\mu}_t = \alpha_0 + \alpha_1 t + \ldots + \alpha_d t^d + \sum_{j=1}^{12} \beta_j \delta_{\{t,j\}}, \tag{3.3}$$

where α_i and β_j are constants and

$$\delta_{\{t,j\}} = \begin{cases} 1, & t \equiv j (\text{mod } 12), \\ 0, & \text{otherwise.} \end{cases} \tag{3.4}$$

So, one cannot in (2.5) arbitrarily set its mean, say, $\boldsymbol{\mu} = \mathbf{0}$.

Now turn to Ω_0 (or equivalently, Σ_η). Based on ARIMA modelling assumption, Hillmer and Trabelsi (1987) gave formula (4.1) in their paper which in "the high signal to noise ratio" situation may lead to getting an approximation of Ω_0 (see the example in Section 5 in the mentioned reference and also in Trabelsi and Hillmer, 1990). However, for their formula (4.1) to be applied, in general, the model (2.1) for $\{\eta_t\}$ should be known.

At a first look, it seems we may 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. Indeed, assuming that $\{\eta_t\}$ and $\{e_t\}$ are uncorrelated and the ARIMA models which they follow are known, it is relatively easy to derive the model of $\{y_t\}$, $y_t = \eta_t + e_t$. However, the problem facing us is the reverse: deriving the model for $\{\eta_t\}$ given those of $\{y_t\}$ and $\{e_t\}$. This is difficult as $\{y_t\}$ and $\{e_t\}$ are correlated.

In fact, we may avoid deriving a model for $\{\eta_t\}$, since in benchmarking formulae (2.5) through (2.8), only Σ_η is necessary. Assuming $\{\eta_t\}$ and $\{e_t\}$ are uncorrelated, $y_t = \eta_t + e_t$ implies

$$\Sigma_\eta = \Sigma_y - \Sigma_e. \tag{3.5}$$

Obtaining Σ_y and Σ_e from their ARIMA model and ARMA model respectively, Σ_η then follows immediately. However, the ARIMA assumption still causes some problems.

First, the ACVF for a series following a nonstationary ARIMA model is not well defined. Suppose $\nabla y_t = x_t$, x_t is stationary, i.e., $y_t = y_1 + \sum_{j=2}^{t} x_j$, then obviously, the ACVF of $y_t, t = 1, \ldots, n$, depends on Var y_1 and Cov (y_1, x_j), which are unknown and can not be estimated from data. Cleveland and Tiao (1976) suggested an approximation $\dot{\Sigma}_y^{-1}$ for Σ_y^{-1} [see formula (A.7) therein], in fact by letting Var $y_1 = \infty$, which results in a singular $\dot{\Sigma}_y^{-1}$, an hence $\dot{\Sigma}_y$ does not exist. Then how to get an approximation of Σ_η by (3.5)?

Furthermore, comparing (2.9) with (1.12), we see that the benchmarked estimate via the signal extraction has smaller variance than that based on the regression, because of a term Σ_η^{-1} being added to the operand of the inverse. Referring to Lemma 2.1 and its proof, we see that, the "larger" the Σ_η (the "smaller" the Σ_η^{-1}), the "closer" the Ω to Ω_R. The situation of "high signal to noise ratio", which Hillmer and Trabelsi discussed in their papers, is not the case for which the signal extraction method shows itself superior to the regression method. In fact, these authors deal with the situation by ignoring Σ_η^{-1} in the benchmarking process [see formulae (3.4) through (3.7) in Trabelsi and Hillmer 1990, compare it with our (2.12)], so it becomes the same as the regression method.

Given Σ_e, even assuming $\{y_t\}$ stationary, we still have a difficulty with modelling approaches (say, fit an ARMA model for $\{y_t\}$), since the parameter estimation procedure in modelling $\{y_t\}$ should guarantee that

$$\hat{\Sigma}_\eta = \hat{\Sigma}_y - \Sigma_e > 0, \tag{3.6}$$

where, $\hat{\Sigma}_y$ is the covariance matrix of $\mathbf{y}$ assuming y_t follow the estimated model. Our simulation in Section 5 shows that, when this property does not hold, the benchmark estimate becomes very unstable.

Further, in most practical situations, n, the number of observations is small. Say quarterly series of 5 to 10 years are very common, and the model identification procedure for such short data is often very unstable. Similarly, for monthly series.

For all above reasons, we here propose a *nonparametric* approach to the signal

extraction benchmarking. Our approach benifits from using the statatistical information given by data but wihout requiring explicit ARIMA models.

4 A Nonparametric Procedure

4.1 The assumptions and the formulae

Following some works, for example, Pierce (1978), we assume that the target series $\{\eta_t\}$ has a deterministic mean series $\{\mu_t\}$ (which is the sum of a smooth trend and a stable seasonality) and a stochastic stationary component $\{\zeta_t\}$ with mean zero, i.e.

$$\eta_t = \mu_t + \zeta_t. \tag{4.1}$$

Then, $\{\eta_t\}$ has the same ACVF as $\{\zeta_t\}$:

$$\gamma_\eta(\pm k) = \gamma_\zeta(\pm k) = E(\zeta_t \zeta_{t+k}). \tag{4.2}$$

Again $\{e_t\}$ is a stationary series, with mean zero and known ACVF (3.1) (a special case would be e_t following an ARMA model). $\{e_t\}$ and $\{\zeta_t\}$ are uncorrelated. Then, (1.1) can be written as

$$y_t = \mu_t + w_t, \tag{4.3}$$

where

$$w_t = \zeta_t + e_t \tag{4.4}$$

are also stationary, with mean zero, and their ACVF have the relationship

$$\gamma_y(\pm k) = \gamma_w(\pm k) = E(w_t w_{t+k}) = \gamma_e(\pm k) + \gamma_\zeta(\pm k). \tag{4.5}$$

Assuming both $\{e_t\}$ and $\{\zeta_t\}$ are regular and without deterministic component (see, say, Doob 1953), so is $\{w_t\}$. Denote the spectral densities of these series by $f_e(\lambda)$, $f_\zeta(\lambda)$ and $f_w(\lambda)$ respectively, then (4.5) leads to

$$f_w(\lambda) = f_e(\lambda) + f_\zeta(\lambda), \quad -\pi < \lambda \leq \pi. \tag{4.6}$$

If μ_t are known, then from y_t (the data) $w_t = y_t - \mu_t$ are obtained, then as (2.5), the estimate of $\boldsymbol{\eta} = (\eta_1 \cdots \eta_n)'$ by $\mathbf{w} = (w_1 \cdots w_n)'$ is

$$\hat{\boldsymbol{\zeta}}_0 = E(\boldsymbol{\zeta} \mid \mathbf{w}) = (\Sigma_e^{-1} + \Sigma_\zeta^{-1})^{-1} \Sigma_e^{-1} \mathbf{w}. \tag{4.7}$$

Notice that, $\Sigma_\zeta = \Sigma_\eta$ is a Toeplitz matrix with $\gamma_\zeta(\mid i-j \mid)$ as its (i,j)th entry. As (2.8),

$$\text{Var}\ \hat{\boldsymbol{\zeta}}_0 = \Omega_0 = (\Sigma_e^{-1} + \Sigma_\zeta^{-1})^{-1}. \tag{4.8}$$

Let $\hat{\boldsymbol{\eta}}_0 = \boldsymbol{\mu} + \hat{\boldsymbol{\zeta}}_0$, then $\hat{\boldsymbol{\eta}}$ follows from (2.6) and (2.7).

From (4.5),

$$\Sigma_w = \Sigma_e + \Sigma_\zeta, \tag{4.9}$$

and then we may easily put (4.7) and (4.8) as

$$\hat{\boldsymbol{\zeta}}_0 = \Sigma_\zeta \Sigma_w^{-1} \mathbf{w}, \tag{4.10}$$

$$\Omega_0 = \Sigma_\zeta \Sigma_w^{-1} \Sigma_e. \tag{4.11}$$

which require much less matrix inverse than (4.7) and (4.8).

4.2 Estimation

We assume Σ_e is known. Suppose that we may get $\{\mu_t\}$ such that $\{w_t\}$, $w_t = y_t - \mu_t$, is stationary, then from data $\mathbf{w}$ we may obtain the estimate $\hat{\Sigma}_w$ and $\hat{\Sigma}_\zeta$ for Σ_w and Σ_ζ respectively which obey relation (4.9). Replacing the true covariance matrices in (4.10) by their estimates, we have

$$\hat{\boldsymbol{\zeta}}_0 = \hat{\Sigma}_\zeta \hat{\Sigma}_w^{-1} \mathbf{w}, \quad \hat{\boldsymbol{\eta}}_0 = \boldsymbol{\mu} + \hat{\boldsymbol{\zeta}}_0. \tag{4.12}$$

For simplicity we use the same notation $\hat{\boldsymbol{\zeta}}_0$. Then $\hat{\boldsymbol{\eta}}$, as a benchmarked estimate of $\boldsymbol{\eta}$, follows from (2.6) and (2.7) by using (4.11) and (4.12) [also, in (4.11), Σ_ζ, Σ_w and hence Ω_0 are replaced by their estimated values respectively].

Now, the problem becomes: (1) How to get μ_t, and (2) how to estimate Σ_ζ (or equivalently, Σ_w).

The first problem may be solved by using the least squares estimation (LSE) to model (4.3), regarding μ_t as a spline plus a stable seasonality like (3.3). Our simulation shows that, although we can only obtain an estimate $\hat{\boldsymbol{\mu}}$ of $\boldsymbol{\mu}$, but using $\hat{\mu}_t$ and $\hat{w}_t = y_t - \hat{\mu}_t$ instead of μ_t and w_t does not effect the final result very much. The intuitive reason would be that, the signal extraction (so does the benchmarked estimate) is sensitive to the "short memory" information, but not sensitive to the

"long memory" imformation that mostly has been taken away by LSE. That means the proposed procedure is robust to the choice of mean in practice, provided that the residuals can be regarded as a stationary series.

Turning to the second problem. The most commonly used nonparametric estimate for $\gamma_w(k)$ is

$$\hat{\gamma}_w(k) = \frac{1}{n}\sum_{t=1}^{n-k} w_t w_{t+k}, \quad k = 0, \ldots, n-1. \tag{4.13}$$

Its direct transformation to the frequency domain

$$\hat{f}_w(\lambda) = \frac{1}{\pi}\{\frac{\hat{\gamma}_w(0)}{2} + \sum_{k=1}^{n-1}\hat{\gamma}_w(k)\cos k\lambda\}, \ -\pi < \lambda \leq \pi \tag{4.14}$$

is the periodogram which is not a good estimate of $f_w(\lambda)$. To get a better estimate of $f_w(\lambda)$, we smooth the periodogram or correspondingly we use

$$\hat{\gamma}_w^{(M)}(k) = \begin{cases} \hat{\gamma}_w(k)u(k/M), & k = 0, \ldots, M, \\ 0, & M < k \leq n. \end{cases} \tag{4.15}$$

as the estimates for $\gamma_w(k)$, $k = 1, \ldots, n-1$. Where $u(s)$ is a commonly used lag window function, such as Tukey-Hanning window, Parzen window and so on (see, say, Priestley 1981).

However, the same problem as in Section 3 [see(3.6)] remains (the definition for Toeplitz matrix $\hat{\Sigma}_\zeta^{(M)}$ and $\hat{\Sigma}_w^{(M)}$ is self-clear), for,

$$\hat{\Sigma}_\zeta^{(M)} = \hat{\Sigma}_w^{(M)} - \Sigma_e > 0 \tag{4.16}$$

maybe does not hold. For that we introduce the following revision procedure.

As $\gamma_e(k)$, or equivalently $f_e(\lambda)$, is given and in view of (4.6), $f_\zeta(\lambda) = f_w(\lambda) - f_e(\lambda) \geq 0$ always hold, so

$$\tilde{f}_\zeta(\lambda) = \max\{\hat{f}_w(\lambda) - f_e(\lambda), \ 0\} \tag{4.17}$$

is a more reasonable estimate of $f_\zeta(\lambda)$ than $\hat{f}_\zeta(\lambda) = \hat{f}_w(\lambda) - f_e(\lambda)$. Corresponding to $\tilde{f}_\zeta(\lambda)$, the estimate of $\gamma_\zeta(k)$ is

$$\tilde{\gamma}_\zeta(k) = 2\int_0^\pi \tilde{f}_\zeta(\lambda)\cos k\lambda d\lambda, \ k = 0, \ldots, n-1. \tag{4,18}$$

Since $\tilde{f}_\zeta(\lambda)$ is again very ragged, we prefer a smoothed estimate $\tilde{f}_\zeta^{(M)}(k)$ which corresponds to

$$\tilde{\gamma}_\zeta^{(M)}(k) = \begin{cases} \tilde{\gamma}_\zeta(k)u(k/M), & k = 0, \ldots, M, \\ 0, & M < k \leq n. \end{cases} \tag{4.19}$$

and the estimate of $\gamma_w(k)$ is given by

$$\tilde{\gamma}_w^{(M)}(k) = \gamma_e(k) + \tilde{\gamma}_\zeta^{(M)}(k), \quad k = 0, \ldots, n. \tag{4.20}$$

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.

Make Toeplitz matrix $\tilde{\Sigma}_\zeta^{(M)}$ and $\tilde{\Sigma}_w^{(M)}$ from $\tilde{\gamma}_\zeta^{(M)}(k)$ and $\tilde{\gamma}_w^{(M)}(k)$ respectively. In view of (4.17), $\tilde{f}_\zeta(k) \geq 0$, and usually $\tilde{f}_\zeta(\lambda) > 0$ holds in a subinterval of $(-\pi, \pi]$, then it is easy to show that $\tilde{\Sigma}_\zeta^{(M)} > 0$, and hence $\tilde{\Sigma}_w^{(M)} = \Sigma_e + \tilde{\Sigma}_\zeta^{(M)} > 0$.

4.3 The procedure

Step 1: Carry out a LSE on model (4.3) to obtain μ_t assuming an explicit model, e.g., a spline function plus a stable seasonality [see (3.3)] (The choice of degrees and knots for the spline is not so important). Denote the estimate of μ_t still by μ_t (rigorously, should be denoted by $\hat{\mu}_t$).

Step 2: Let $w_t = y_t - \mu_t$ (rigorously, should be denoted by $\hat{w}_t$), use (4.13), (4.14) and (4.17) to obtain $\hat{\gamma}_w(k), \hat{f}_w(\lambda)$ and $\tilde{f}_\zeta(\lambda)$ consecutively. Then calculate $\tilde{\gamma}_\zeta(k)$ following (4.18) by using a numerical method, and finally (4.19) and (4.20) are used to obtain $\tilde{\gamma}_\zeta^{(M)}(k)$ and $\tilde{\gamma}_w^{(M)}(k)$.

Step 3: Calculate $\hat{\boldsymbol{\zeta}}_0$ and Ω_0 by using (4.10) and (4.11), where Σ_ζ, Σ_w are replaced by $\tilde{\Sigma}_\zeta^{(M)}$ and $\tilde{\Sigma}_w^{(M)}$ respectively. Let $\hat{\boldsymbol{\eta}}_0 = \boldsymbol{\mu} + \hat{\boldsymbol{\zeta}}_0$, obtain $\hat{\boldsymbol{\eta}}_C$ and $\hat{\boldsymbol{\eta}}$ by (2.6) and (2.7) consecutively.

Note: In the above procedure, Σ_e, Σ_ϵ are known (usually given by survey experts), L is given by the model assumption and $\mathbf{y}$ and $\mathbf{z}$ are data.

5 Simulation

5.1 The 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 recursively by the corresponding difference equations. 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 can not 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(3.3) and (3.4)]

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

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), 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 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 $\hat{\boldsymbol{\eta}}_0$ and the benchmarked value $\hat{\boldsymbol{\eta}}$. For example, RMSE (all benchmarked) 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 $\hat{\boldsymbol{\eta}}^{(j)}$ and $\hat{\boldsymbol{\eta}}_0^{(j)}$ respectively, which means that, after benchmarking the error should be smaller than before. We say, a replication is valid if (5.5) holds. The reality is that, if a method is bad, the correction term $\hat{\boldsymbol{\eta}}_C$ may even explode (beyond hundreds, thousands in our simulation), so the high percentage of the valid replication means the method is stable. The first two columns in the Table 5.1 and 5.2 have similar definitions as (5.4) but 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), given the true value of $\Sigma_\eta = \Sigma_\zeta$, so does of $\Sigma_y = \Sigma_w$. 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 RMSE of benchmarked estimate is 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 (4.13) 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 (4.15) for $\gamma_w(k)$ with $M = [n/3], [2n/3]$ and $n-1$ respectively; 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 positive definite as no revision procedure has been taken.

Method 7 uses (4.18) to estimate $\gamma_\zeta(k)$; method 8, 9 and 10 use (4.19) 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 (4.20). 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 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 is unknown

We now turn to a situation closer to the reality: where μ_t are given by (5.3) but we do not know the parameters, so we get the LSE of the parameters from "data" 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 (4.10), method 2 is now not as good as in the situation of "mean is known", but 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 where we do not know the type of the function which μ_t follow. The trend could be linear, quadratic or other kind of smooth curves. For testing our approach, we designed the following simulation: μ_t are generatead 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 misspecification 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 misspecification, parameter estimation error only). In other words, the right model specification of μ_t is not so crucial.

TABLE 5.1 Simulation results for model 1

Method	RMSE (valid sig. extr.)	RMSE (valid benchmarked)	no. of valid rep.	RMSE (all sig.extr.)	RMSE (all benchmarked)
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 Simulation results for model 2

Method	RMSE (valid sig. extr.)	RMSE (valid benchmarked)	no. of valid rep.	RMSE (all sig.extr.)	RMSE (all benchmarked)
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 RMSE of $\hat{\eta}$ for different models and situations

			model 1			model 2		
mean	S/N	meth.	RMSE	%	no.	RMSE	%	no.
k n o w n	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
l i n e a r	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
q u a d .	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 two models is as following:

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

6 Conclusions

In this paper we analysed the properties of the regression and the signal extraction methods for benchmarking survey data 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 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 latter 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 techniques. In this approach we assume that $\{\eta_t\}$ is the sum of a deterministic component $\{\mu_t\}$ which is a smooth trend plus a stable seasonality and a stationary stochastic component $\{\zeta_t\}$ which accounts for the presence of both stochastic trend-cycle and seasonality.

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

In the simulation, $\{\mu_t\}$ is created following a linear or a quadratic trend plus a stable seasonality, and $\{\zeta_t\}$ and $\{e_t\}$ following ARMA models. Two models (1 and 2) are chosen to distinguish two different situations: the spectral densities $f_\eta(\lambda)$ and $f_e(\lambda)$, of $\{\zeta_t\}$ and $\{e_t\}$ respectively, have similar or very different patterns. Furthermore the effects of the levels of the signal to noise ratio are considered. The results show:

(1) A smooth spectral estimate corresponding to an estimate of the covariance matrix Σ_η is needed and this estimate of Σ_η must be positive definite for the NP method to produce estimates close to the "true" values of η_t.

(2) Using RMSE as a measure of efficiency, SE is always better than RM, NP is only slightly worse than SE.

(3) 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 of these two patterns is big, then the gains are larger.

(4) If the signal to noise ratio $S/N = \gamma_\zeta(0)/\gamma_e(0)$ is large, the efficiency gains are small and conversely.

(5) The only parametric model we used is to estimate μ_t. The NP method is robust to model misspecification and estimation error of the parameters.

Appendix: Proof of (2.5)-(2.9)

The projection of random vector $\boldsymbol{\eta}$ on the probability space spanded by $\mathbf{y}$ is

$$\hat{\boldsymbol{\eta}}_0 = \hat{E}(\boldsymbol{\eta} \mid \mathbf{y}) = E\boldsymbol{\eta} + \text{Cov}\,(\boldsymbol{\eta}, \mathbf{y})(\text{Var}\,\mathbf{y})^{-1}(\mathbf{y} - E\mathbf{y}), \qquad (A.1)$$

and the variance matrix of the error using $\hat{\boldsymbol{\eta}}_0$ to estimate $\boldsymbol{\eta}$ is

$$\Omega_0 = E[(\hat{\boldsymbol{\eta}}_0 - \boldsymbol{\eta})(\hat{\boldsymbol{\eta}}_0 - \boldsymbol{\eta})'] = \text{Var}\,\boldsymbol{\eta} - \text{Cov}\,(\boldsymbol{\eta}, \mathbf{y})(\text{Var}\,\mathbf{y})\text{Cov}\,(\mathbf{y}, \boldsymbol{\eta}). \qquad (A.2)$$

Noticing

$$\text{Cov}\,(\boldsymbol{\eta}, \mathbf{y}) = \Sigma_\eta, \quad \Sigma_y = \text{Var}\,\mathbf{y} = \Sigma_\eta + \Sigma_e, \quad E\mathbf{y} = E\boldsymbol{\eta} = \boldsymbol{\mu}, \qquad (A.3)$$

(2.5) and (2.8) can be derived easily.

We are now going to show

$$\hat{\boldsymbol{\eta}} = \hat{E}(\boldsymbol{\eta} \mid \boldsymbol{\tau}) = \Omega(X'\Sigma_u^{-1}\boldsymbol{\tau} + \Sigma_\eta^{-1}\boldsymbol{\mu}), \qquad (A.4)$$

then the same derivation gives (2.6) and (2.7) as Hillmer and Trabelsi (1987) did.

From (1.8) through (1.10) and (A.3),

$$\text{Cov}\,(\boldsymbol{\eta}, \boldsymbol{\tau}) = \Sigma_\eta X', \quad \text{Var}\,\boldsymbol{\tau} = X\Sigma_\eta X' + \Sigma_u, \quad E\boldsymbol{\tau} = X\boldsymbol{\mu}, \qquad (A.5)$$

then by the similar formula as (A.1), we have

$$\begin{aligned}\hat{\boldsymbol{\eta}} &= \boldsymbol{\mu} + \Sigma_\eta X'(X\Sigma_\eta X' + \Sigma_u)^{-1}(\boldsymbol{\tau} - X\boldsymbol{\mu}) \\ &= \boldsymbol{\mu} + \Sigma_\mu X'[\Sigma_u^{-1} - \Sigma_u^{-1}X(\Sigma_\eta^{-1} + X'\Sigma_u^{-1}X)^{-1}X'\Sigma_u](\boldsymbol{\tau} - X\boldsymbol{\mu}). \qquad (A.6)\end{aligned}$$

The second equality in (A.6) is due to the partitioned inverse of a matrix. Then (A.4) follows by some rearrangements and basic operations of metrices. Using similar formula as (2.2) to express $E[(\hat{\eta} - \eta)(\hat{\eta} - \eta)']$ it gives Ω.

Acknowledgement

We thank our colleagues Kim Chiu, Benoit Quenneville and Paul Wong for their help in the simulation.

References

Cholette, P.A. and Dagum, E.B. (1989). Benchmarking socio-economic time series data: a unified approach. Working paper TSRA-89-006E, Statistics Canada.

Cholette, P.A. and Dagum, E.B. (1991). Benchmarking time series with autocorrelated sampling errors. Working paper TSRA-91-005E, Statistics Canada.

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

Denton, F.T. (1971). Adjustment of monthly or quarterly series to annual totals: an approach based on quadratic minimization. *JASA*, **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. *JASA*, **82**, 1064-1071.

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.

Trabelsi, A. and Hillmer, S.C. (1990). Bench-marking time series with reliable benchmarks. *Appl. Statist.*, **39**, 367-379.