Benchmarking by State Space Models
Author(s): J. Durbin and B. Quenneville
Source: *International Statistical Review / Revue Internationale de Statistique*, Vol. 65, No. 1 (Apr., 1997), pp. 23-48
Published by: International Statistical Institute (ISI)
Stable URL: https://www.jstor.org/stable/1403431
Accessed: 27-05-2019 01:45 UTC

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms

International Statistical Institute (ISI) is collaborating with JSTOR to digitize, preserve and extend access to *International Statistical Review / Revue Internationale de Statistique*

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

International Statistical Review (1997), **65**, 1, 23–48, Printed in Mexico
© International Statistical Institute

Benchmarking by State Space Models

J. Durbin[1]* and B. Quenneville[2]

[1]*London School of Economics and Political Science, Houghton St., London, WC2 2AE, UK*
[2]*Time Series Research and Analysis Centre, Statistics Canada, Ottawa, Ontario, Canada, K1A 0T6*

Summary

We have a monthly series of observations which are obtained from sample surveys and are therefore subject to survey errors. We also have a series of annual values, called benchmarks, which are either exact or are substantially more accurate than the survey observations; these can be either annual totals or accurate values of the underlying variable at a particular month. The benchmarking problem is the problem of adjusting the monthly series to be consistent with the annual values. We provide two solutions to this problem. The first of these is a two-stage method in which we first fit a state space model to the monthly data alone and then combine the results obtained at this stage with the benchmark data. In the second solution we construct a single series from the monthly and annual values together and fit a state space model to this series in a single stage. The treatment is extended to series which behave multiplicatively. The methods are illustrated by applying them to Canadian retail sales series.

Key words: Kalman Filter; Posterior mode estimation; Structural time series models; Trend; Seasonality; Trading day; Survey error; Smoothing.

1 Introduction

A common problem faced by official statistical agencies is the adjustment of monthly or quarterly time series which have been obtained from sample surveys to make them consistent with more accurate values obtained from other sources. These values can be aggregates or individual values at arbitrary points along the series. The sources can be censuses or more accurate sample surveys or administrative data or some combination of these. This adjustment process is called benchmarking and the more accurate values are called benchmarks. Typically, the benchmarks are either yearly totals or values observed at a particular time-point each year. For simplicity, we assume that the data are monthly and the benchmarks are annual, though the broader interpretation should be borne in mind.

Most previous work (see, for example, Cholette & Dagum (1994) and Mian & Laniel (1993) and the references therein) took the underlying target series as a series of constants to be estimated by some appropriate method such as maximum likelihood. Hillmer & Trabelsi (1987) introduced the idea of using a stochastic model for the target series. This takes account of autocorrelation of the underlying values thus leading to greater efficiency. We follow this second approach.

In this paper we solve the benchmarking problem by modelling the monthly or quarterly data by state space structural time series models. We give solutions both for the case where the monthly observations and the annual values behave additively, and also for the case, very common in practice for economic series, where the monthly observations behave multiplicatively but the annual values behave additively, for example yearly totals. It is frequently the case that the sample survey data are

*J. Durbin's work was done at the Time Series Research and Analysis Centre, Statistics Canada.

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

biased due to non-response and other factors. The existence of accurate annual values enables this bias to be estimated. We derive estimates for the bias in both the additive and multiplicative cases.

We present two methods for solving the benchmarking problem, first a two-stage method in which in the first stage we fit a structural time series model to the monthly values alone and then in the second stage we combine information from the first stage with information from the benchmarks. Secondly, we develop a single-stage method in which we combine the monthly and annual values into a single time series which is fitted to an appropriate state space model. We present the two-stage method first since the two-stage approach to the problem is already familiar to workers in the field through the paper of Hillmer & Trabelsi (1987) who used an ARIMA model to represent the monthly data at the first stage. Our view is that the state space approach is superior to the ARIMA approach for this type of problem because of its greater flexibility and comprehensiveness; for example, it permits us to allow for trading day variation in a straightforward way. Other differences from the Hillmer & Trabelsi (1987) treatment are that they assume that non-stationary series have zero means whereas we do not, we allow for heteroscedasticity in the survey errors, we estimate survey bias and finally we treat the multiplicative case.

We present the single-stage solution for a variety of reasons, the main one being that it is only by the single-stage approach that the structural time series model can be efficiently fitted to the data using all the information available including benchmarks. For example, we obtain estimates of the state vector based on all the data whereas in the two-stage method we can only estimate the state vector from the monthly data alone. If the number of benchmarks is large enough this can be significant for trend estimation, seasonal adjustment and other matters where analysis of the components is required. A further point is that from the standpoint of state space theory this is the natural way to solve the benchmarking problem. From this perspective the single-stage solution represents an advance in state space methodology which has substantial potential for further development, for example to the multivariate case. It has the additional advantage, characteristic of state space models, that it is relatively easy to update the system each time a new observation comes in. On the other hand, the two-stage method has the advantage that the structural time series model could be fitted at the first stage by existing software, so only the second stage would require new software.

In section 2, we develop a state space model for the monthly observations which includes trend, seasonal and trading-day components, and which also allows for survey errors. Section 3 describes the fitting of the model to the monthly observations and shows how the results of this stage can be combined with the benchmark data for the case where the monthly series behaves additively. In section 4 we consider the case where the monthly observations behave multiplicatively but the benchmarks are additive. After taking logarithms, the fitting of the model to the monthly observations takes place as in the previous section. The task of combining information from this stage with the benchmarks is a non-linear problem. We solve this by taking as our estimate of the underlying series the mode of its posterior density given the monthly data and the benchmarks. This means that we are taking as our estimates of the underlying quantities their most probable values given all the data. The mode is obtained by an iterative process.

Section 5 shows how to estimate survey bias. In section 6 we consider single-stage benchmarking. We begin by embedding the benchmarks within the monthly observations to form a single series. For the additive case we construct a state space model for this series which is then analysed by standard Kalman filtering and smoothing. Of course, the requisite model is somewhat more complicated than the model of section 2. For the multiplicative case, considered in section 7, we linearise the estimating equations for the unknowns concerned and put them into a form which can be solved by a Kalman filtering and smoothing operation analogous to that used for the additive case. This enables us to develop an iterative method for the estimation of the posterior mode of the target series.

In section 8 the methods we developed are applied to a single data set, the Canadian retail trade series from January 1980 to December 1989. There are seven benchmarks, four of which are twelve-month totals and three are individual values. The data are particularly interesting since

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

they demonstrate the need for bias estimation and for the use of multiplicative models for the benchmarking problem.

2 A State Space Model For Benchmarking

Suppose that we have a monthly series of univariate observations y_t, obtained from sample surveys, which are estimates of underlying true values η_t, and which satisfy the relation

$$y_t = \eta_t + k_t u_t, \qquad t = 1, \dots, n, \tag{2.1}$$

where $k_t u_t$ is the survey error. We assume that u_t is a unit-variance stationary ARMA(p,q) series and that values of k_t, p and q are available from survey experts. The inclusion of the factor k_t enables us to allow for heteroscedasticity in the survey errors; obviously, k_t is their standard deviation. Suppose that in addition we have a series of annual values $x_i, i = 1, \dots, l$ which are available from another source and which are regarded as more accurate than the y_t's. We assume that errors in the benchmarks are independent of errors in the monthly observations. Thus we are not taking account of any correlations that might exist between response errors in the annual values and the monthly series. It is important to note that the term annual values need not refer to calendar years. For example we may have a series of values which are totals of February to January values together with accurate values in particular months. This is the case for the example considered in section 8. Let $\eta = [\eta_1, \dots, \eta_n]'$ and $x = [x_1, \dots, x_l]'$. The x_i*'s* are assumed to satisfy the benchmarking relations

$$x = L\eta + e, \qquad e \sim N(0, \Sigma_e), \tag{2.2}$$

where L is a known matrix and where we assume that Σ_e is known either because it is known to be a zero matrix or because the x_i*'s* are obtained from accurate annual surveys, in which case survey experts can be expected to provide us with an estimate of Σ_e. The x_i*'s* are called *benchmarks* and when $e = 0$ the benchmarks are said to be *binding*.

We assume that the true values η_t are generated by the structural time series model

$$\eta_t = \mu_t + \gamma_t + \sum_{j=1}^{k} \delta_{jt} w_{jt} + \epsilon_t, \qquad \epsilon_t \sim N(0, \sigma_\epsilon^2), \qquad t = 1, \dots, n, \tag{2.3}$$

where μ_t is the trend, γ_t is the seasonal, $\sum_{j=1}^{k} \delta_{jt} w_{jt}$ is the trading day (plus leap year) adjustment and ϵ_t is an error term. A choice of models is available for μ_t and γ_t ; see, for example, Harvey (1989) section 2.3. In this paper we use the models

$$\mu_t = 2\mu_{t-1} - \mu_{t-2} + \xi_t, \qquad \xi_t \sim N(0, \sigma_\xi^2), \tag{2.4}$$

$$\gamma_t = -\sum_{j=1}^{11} \gamma_{t-j} + \omega_t, \qquad \omega_t \sim N(0, \sigma_\omega^2). \tag{2.5}$$

The idea behind (2.4) is that if $\xi_t = 0$, a straight line is followed by the trend exactly; inclusion of the error ξ_t allows the slope to change over time. The idea behind (2.5) is that if $\omega_t = 0$, the seasonal pattern is constant; inclusion of the error ω_t allows the pattern to vary. For the trading day coefficients we use the model

$$\delta_{jt} = \delta_{j,t-1} + \zeta_{jt}, \qquad \zeta_{jt} \sim N(0, \sigma_\zeta^2), \qquad j = 1, \dots, k. \tag{2.6}$$

This allows the coefficients to change over time, which is a desirable feature in many applications. The case where an adequate fit can be achieved by treating the coefficients as constant can be handled by putting $\zeta_{jt} = 0$; in this case the model-fitting process is simplified substantially. It should be noted that it is not essential that the coefficients are updated every month. For example, they could

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

be updated once a year each January by putting $\zeta_{jt} = 0$ for $t \neq 12i - 11$. More general models for time-varying trading day adjustments are considered by Dagum, Quenneville & Sutradhar (1992). The error series ϵ_t, ξ_t, ω_t, and the values of $\zeta_{1t}, \dots, \zeta_{kt}$ that are not identically zero are assumed to be white noise series independent of each other and of u_t.

By adding in the survey error we obtain the model for the observations y_t,

$$y_t = \mu_t + \gamma_t + \sum_{j=1}^{k} \delta_{jt} w_{jt} + \epsilon_t + k_t u_t, \qquad t = 1, \dots, n. \tag{2.7}$$

Let us consider the state space form of this model for the special case where the survey error term u_t follows the AR(1) model $u_t = \phi u_{t-1} + \chi_t$, $\chi_t \sim N(0, 1 - \phi^2)$, leaving until later the extension to a general ARMA(p,q) model. The observation equation (2.7) can be written in the form

$$y_t = \tilde{Z}_t \alpha_t$$

where

$$\tilde{Z}_t = [1 \ \ 0 \ \ 1 \ \ 0 \ \ \dots \ \ 0 \ \ w_{1t} \ \ \dots \ \ w_{kt} \ \ 1 \ \ k_t] \tag{2.8}$$

and the state vector α_t is

$$\alpha_t = \left[\mu_t,\ \mu_{t-1},\ \gamma_t,\ \gamma_{t-1}, \dots, \gamma_{t-10},\ \delta_{1t}, \dots, \delta_{kt},\ \epsilon_t,\ u_t\right]'. \tag{2.9}$$

A slightly unusual feature of this formulation is that we have put the error term ϵ_t into the state vector. We do this in order to facilitate calculation of the variance matrix of errors of estimation of the η_t's. The state transition equation is

$$\alpha_t = T_t \alpha_{t-1} + R_t \nu_t, \qquad t = 1, \dots, n,$$

where initially T_t is a time-invariant block diagonal matrix with blocks

$$\begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}, \ \begin{bmatrix} \multicolumn{2}{c}{-1'_{11}} \\ \mathrm{I}_{10} & 0_{10,1} \end{bmatrix}, \ \mathrm{I}_k, \ 0; \ \phi \tag{2.10}$$

for the trend, seasonal, trading-day, observation error and survey error components respectively, R_t is a selection matrix composed of a subset of columns of the identity matrix and $\nu_t = [\xi_t,\ \omega_t,\ \zeta_{1t}, \dots, \zeta_{kt},\ \epsilon_t,\ \chi_t\ ]'$. Here, 1_c denotes a $c \times 1$ vector of ones, $0_{r,c}$ denotes an $r \times c$ matrix of zeros, and I_c denotes an identity matrix of order c. Later, when dealing with the single-stage case, we use a slightly different formulation in which T_t and R_t vary over time.

If u_t is an ARMA(p,q) series, let $r = \max(p, q + 1)$ and write its model in the form

$$u_t = \phi_1 u_{t-1} + \dots + \phi_r u_{t-r} + \chi_t + \theta_1 \chi_{t-1} + \dots + \theta_{r-1} \chi_{t-r+1}, \ \ \chi_t \sim N(0, \sigma_\chi^2),$$

where σ_χ^2 is chosen to ensure that $\mathrm{Var}(u_t) = 1$. Then replace k_t in (2.8) by the row vector $[k_t \ \ 0 \ \ \dots \ \ 0]'$, u_t in (2.9) by the vector

$$\begin{aligned} [u_t, \phi_2 u_{t-1} \ &+ \ \dots + \phi_r u_{t-r+1} + \theta_1 \chi_t + \dots + \theta_{r-1} \chi_{t-r+2}, \\ \phi_3 u_{t-1} \ &+ \ \dots + \phi_r u_{t-r+2} + \theta_2 \chi_t + \dots + \theta_{r-1} \chi_{t-r+3}, \ \dots, \ \phi_r u_{t-1} + \theta_{r-1} \chi_t \ ]', \end{aligned}$$

ϕ in (2.10) by the matrix

$$\left[\begin{array}{c:c} \phi_1 & \\ \vdots & \mathrm{I}_{r-1} \\ \phi_{r-1} & \\ \hdashline \phi_r & 0_{1,r-1} \end{array}\right]$$

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

and make a suitable adjustment of R_t. We wish to estimate the quantity

$$\eta_t = Z_t\alpha_t$$

using both monthly and benchmark data where Z_t is the same as $\tilde{Z}_t$ except that k_t in $\tilde{Z}_t$ is replaced by zero.

The estimate we shall use is the mean of the posterior distribution of η given all the data, that is,

$$\hat{\eta}_t = E\left(\eta_t \mid y, x\right), \qquad t = 1, \ldots, n,$$

where $y = [y_1, \ldots, y_n]'$. Let $\hat{\eta} = \left[\hat{\eta}_1, \ldots, \hat{\eta}_n\right]'$ and $z = \left[y', x'\right]'$. Then

$$\hat{\eta} = E\left(\eta \mid z\right). \tag{2.11}$$

By the properties of the multivariate normal distribution, $\hat{\eta}$ is a linear function of z, say $\hat{\eta} = Dz$. Denote its mean square error matrix (MSE) $E\left[\left(\hat{\eta} - \eta\right)\left(\hat{\eta} - \eta\right)'\right]$ by V.

We now show that the estimate $\hat{\eta}$ that is obtained by assuming normality has an optimum property even when the random variables in the model are non-normal. Let $\hat{\eta} - \eta = \kappa$. From (2.11), $E\left(\kappa \mid z\right) = 0$ so $E\left(z\kappa'\right) = E\left[zE\left(\kappa' \mid z\right)\right] = 0$. Now let $\eta^* = D^*z$ be any other linear function of z. Then $\eta^* - \eta = \left(D^* - D\right)z + \kappa$ so, since $E\left(z\kappa'\right) = 0$ and $E\left(\kappa\kappa'\right) = V$,

$$E\left[\left(\eta^* - \eta\right)\left(\eta^* - \eta\right)'\right] = \left(D^* - D\right)W\left(D^* - D\right)' + V \tag{2.12}$$

where $W = E\left(zz'\right)$.

Suppose now that the errors in the model, possibly including the survey errors, are non-normally distributed with zero means and the same variances and covariances as in the normal case and that we use the same estimate $\hat{\eta} = Dz$ to estimate η. Then because of the linearity of the model and of $\hat{\eta}$ and since $E\left(\hat{\eta} - \eta\right) = 0$ we must have MSE $\left(\hat{\eta}\right) = V$. Defining $\kappa = Dz - \eta$ we have $E\left(z\kappa'\right) = 0$ since it is exactly the same function of variances and covariances as in the normal case. It follows that (2.12) holds and consequently that MSE $\left(\eta^*\right)$ – MSE $\left(\hat{\eta}\right)$ is non-negative definite, so in the class of linear estimators of η, $\hat{\eta} = Dz$ has minimum mean square error.

3 Two-Stage Benchmarking

Two different approaches can be used to calculate $\hat{\eta}$ given by (2.11). In the first approach the problem is solved in two stages. At the first stage we compute $\tilde{\eta} = E\left(\eta \mid y\right)$ by applying standard Kalman filtering and smoothing (KFS) to the monthly series alone, ignoring the benchmark values. Then at the second stage we combine the information about η in $\tilde{\eta}$ with that in the benchmark values x to obtain the final estimate $\hat{\eta}$. In the second approach we form a single series, $y_1, \ldots, y_{12}, x_1, y_{13}, \ldots, y_{24}, x_2, y_{25}, \ldots,$ in the case of annual benchmarks, and apply a specially designed KFS to this series. In this section we shall present the two-stage solution. The single-stage solution will be described in Section 6.

For initialising the filter, we first note that if, for example, annual values are available for the first two years it is better to fix the initial value of α_0 and estimate it by maximum likelihood or generalised least squares (GLS) rather than use a diffuse prior. This is because, with a diffuse prior, we are in effect treating the first τ values of y_t as fixed, so since our state vectors are usually rather long, the first two annual values cannot be used to adjust the values of y_t in the first two years. Here, τ is the smallest value of t such that $a_{t+1} = E\left(\alpha_{t+1} \mid y_1, \ldots, y_t\right)$ has a finite MSE. In consequence, the benchmark for the first year, and perhaps also the second year, cannot be satisfied by the adjusted monthly values. However, if α_0 is fixed and unknown and is estimated by maximum likelihood or GLS there is no analogous difficulty about using the first two annual values. Harvey (1989), section 3.4.4, describes methods of estimating a fixed α_0 by Rosenberg's algorithm or by GLS.

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

Now $\eta_t = Z_t \alpha_t$ so $\tilde{\eta}_t = Z_t \tilde{\alpha}_t$ where $\tilde{\alpha}_t = E(\alpha_t \mid y)$. Let $\Omega = E\left[(\tilde{\eta} - \eta)(\tilde{\eta} - \eta)'\right] = [\omega_{st}]$ where $s, t = 1, \ldots, n$. Then $\omega_{st} = Z_s E\left[(\tilde{\alpha}_s - \alpha_s)(\tilde{\alpha}_t - \alpha_t)'\right] Z_t'$. Our final estimate $\hat{\eta}$ of η based on both monthly and benchmark values is given by

$$\hat{\eta} = \tilde{\eta} + \Omega L' \left[L \Omega L' + \Sigma_e\right]^{-1} (x - L\tilde{\eta}). \tag{3.1}$$

Formula (3.1) is due to Hillmer & Trabelsi (1987), equation (2.16). However, their proof is complicated due to the fact that they do not relate the result directly to standard regression theory. Moreover, they make the unnecessary and, in our case, invalid assumption that a nonstationary series has zero mean. We therefore give an alternative brief proof that was communicated to us independently by Dr. Z.G. Chen and Dr. W.R. Bell.

Since $x = L\eta + e$, $E(x \mid y) = L\tilde{\eta}$. Consequently, $\hat{\eta} = E(\eta \mid y, x) = E(\eta \mid y, x - L\tilde{\eta})$. Now $E(x - L\tilde{\eta}) = 0$ and y and $x - L\tilde{\eta}$ are orthogonal. Hence, by standard regression theory,

$$E(\eta \mid y, x - L\tilde{\eta}) = E(\eta \mid y) + \operatorname{Cov}(\eta, x - L\tilde{\eta}) \left[\operatorname{Var}(x - L\tilde{\eta})\right]^{-1} (x - L\tilde{\eta}). \tag{3.2}$$

Now $E(\eta \mid y) = \tilde{\eta}$, $\operatorname{Cov}(\eta, x - L\tilde{\eta}) = E\left[\eta (x - L\tilde{\eta})'\right] = E\left[\eta (\eta - \tilde{\eta})' L'\right] = \Omega L'$ and $\operatorname{Var}(x - L\tilde{\eta}) = \operatorname{Var}\left[L(\eta - \tilde{\eta}) + e\right] = L\Omega L' + \Sigma_e$. Substituting in (3.2) gives (3.1).

We now give the MSE matrix for $\hat{\eta}$. Let $B = \Omega L' \left[L\Omega L' + \Sigma_e\right]^{-1}$. Then

$$\hat{\eta} - \eta = \tilde{\eta} - \eta + B(x - L\tilde{\eta})$$

so

$$\begin{aligned} V &= E\left[(\hat{\eta} - \eta)(\hat{\eta} - \eta)'\right] \\ &= E\left[(\tilde{\eta} - \eta)(\tilde{\eta} - \eta)'\right] + B \operatorname{Var}(x - L\tilde{\eta}) B' - 2E\left[\eta (x - L\tilde{\eta})' B'\right] \\ &= \Omega - \Omega L' \left[L\Omega L' + \Sigma_e\right]^{-1} L\Omega. \end{aligned} \tag{3.3}$$

Assuming that Σ_e is known, we see that implementation of (3.1) requires only a knowledge of $\tilde{\eta}$ and Ω. Using results of de Jong (1989, 1991) we find that we can compute $\tilde{\eta}_t$ and ω_{st} by the following recursions,

$$\tilde{\eta}_t = Z_t a_t + Z_t P_t r_{t-1} \tag{3.4}$$

where

$$r_{t-1} = \tilde{Z}_t' v_t / F_t + L_t' r_t, \qquad t = n, n-1, \ldots, 1 \tag{3.5}$$

with $r_n = 0$, and

$$\omega_{tt} = Z_t (P_t - P_t D_{t-1} P_t) Z_t' \tag{3.6}$$

$$\omega_{st} = Z_s P_s L_s' \ldots L_{t-1}' (\mathrm{I}_m - D_{t-1} P_t) Z_t', \qquad s < t, \tag{3.7}$$

where

$$D_{t-1} = \tilde{Z}_t' \tilde{Z}_t / F_t + L_t' D_t L_t, \qquad t = n, n-1, \ldots, 1 \tag{3.8}$$

with $D_n = 0$. Here a_t, P_t, v_t, F_t and L_t have been computed during the Kalman filtering process and are defined as follows,

$$\begin{aligned} a_t &= E(\alpha_t \mid y_1, \ldots, y_{t-1}), \\ P_t &= E\left[(a_t - \alpha_t)(a_t - \alpha_t)'\right], \\ v_t &= y_t - \tilde{Z}_t a_t, \\ F_t &= \operatorname{Var}(v_t), \\ L_t &= T_{t+1} - K_t \tilde{Z}_t, \\ K_t &= T_{t+1} P_t \tilde{Z}_t' / F_t. \end{aligned}$$

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

4 Two-Stage Benchmarking when the Monthly Model is Multiplicative

In many practical situations, social and economic variables behave multiplicatively, and when such variables are measured by sample surveys the survey errors also usually operate multiplicatively. Thus what we actually observe in such cases is

$$Y_t = N_t U_t, \qquad t = 1, \dots, n, \tag{4.1}$$

where N_t is the underlying true value that we wish to estimate and U_t is the survey error. Let $y_t = \log Y_t$, $\eta_t = \log N_t$ and $k_t u_t = \log U_t$, where we make the same assumptions about u_t as in section 2 but k_t is now the coefficient of variation (cv). The reason for this is that if U is an error with mean μ and if $U = \mu + e$ then $\log U = \log \mu + \log(1 + e/\mu) \approx \log \mu + e/\mu$ for e "small" so $\mathrm{Var}(\log \mathrm{U}) \approx \mathrm{Var(e)}/\mu^2 = (\mathrm{cv})^2$. Thus if we put $\log U = ku$ where $\mathrm{Var(u)} = 1$ then $k = \mathrm{cv}$. Equation (4.1) transforms into

$$y_t = \eta_t + k_t u_t,$$

which is the same as (2.1). We assume that η_t and y_t satisfy the same models as in (2.3) and (2.7).

The benchmarking relations are, however, expressed in terms of the true values N_t and not their logs η_t. Thus the benchmark values we observe are given by

$$x = LN + e, \tag{4.2}$$

where $x = [x_1, \dots, x_l]'$ as before, $N = [N_1, \dots, N_n]'$ and L and e are the same as in (2.2). The problem we shall consider is to estimate $N_1, \dots, N_n$ given both the observations Y_t and the benchmarks x_i. The difficulty is that the model is now nonlinear so the solution of section 3 does not apply. Moreover, because the model is nonlinear we cannot estimate N_t straightforwardly by its posterior mean $E\,(N_t \mid y, x)$ since an analytical solution is intractable. What we suggest is that we estimate η_t or N_t by the *modes* of their posterior densities given y and x instead of their *means*. We call these estimates *posterior mode estimates* (PME's). Such estimates have been used by Fahrmeir (1992), Durbin & Koopman (1992) and Durbin & Cordero (1993) for estimating the state vector in state space models for non-Gaussian observations. Of course, when the posterior density is symmetric the mode is equal to the mean, so since calculation of the mode is relatively straightforward, it gives an easy way of computing the mean. If the posterior density is not symmetric, use of the mode can be justified on the intuitive ground that it is the estimate of the unknown which is most probable given the data. We assume that the density is unimodal.

We have the choice between taking $\hat{\eta}_t$ as the mode of its density and then taking $\hat{N}_t = \exp\left(\hat{\eta}_t\right)$ or taking $\hat{N}_t$ as the mode of its own density. However, in the usual case in which standard errors are small relative to means the difference between these estimates will be small.

Taking $p\,(\eta \mid y)$ as the prior and $p\,(x \mid \eta, y)$ as the likelihood, the posterior density of η is

$$p\,(\eta \mid y, x) \propto p\,(\eta \mid y)\, p\,(x \mid \eta, y)\,,$$

so the log of the conditional density of η given y and x is

$$\log p\,(\eta \mid y, x) = \text{constant} - \frac{1}{2}\left[(\eta - \tilde{\eta})'\,\Omega^{-1}\,(\eta - \tilde{\eta}) + (x - LN)'\,\Sigma_e^{-1}\,(x - LN)\right].$$

To obtain the PME of η we differentiate this with respect to η_t and equate to zero for $t = 1, \dots, n$. This gives

$$\begin{aligned} \frac{\partial \log p}{\partial \eta_t} &= t^{th}\text{element of } \Omega^{-1}\,(\tilde{\eta} - \eta) + \exp(\eta_t) \times t^{th}\text{element of } L'\Sigma_e^{-1}\,(x - LN) \\ &= t^{th}\text{element of } \left[\Omega^{-1}\,(\tilde{\eta} - \eta) + KL'\Sigma_e^{-1}\,(x - LN)\right] \end{aligned}$$

where K is a diagonal matrix with elements $\exp(\eta_1), \dots, \exp(\eta_n)$ on the main diagonal. The PME

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

$\hat{\eta}$ is therefore the solution of the equation

$$\Omega^{-1}(\tilde{\eta}-\eta)+KL'\Sigma_e^{-1}(x-LN)=0. \tag{4.3}$$

This should be compared with the equation that is obtained by applying the same technique to the additive model considered in Section 3, namely

$$\Omega^{-1}(\tilde{\eta}-\eta)+L'\Sigma_e^{-1}(x-L\eta)=0. \tag{4.4}$$

Since in the Gaussian case modes and means are the same, the solution of (4.4) is given by (3.1). We see that whereas equation (4.4) is linear, equation (4.3) is nonlinear in η. We shall solve (4.3) by iteration by linearising it and then putting the linearised equation into a form analogous to (4.4) so that we can use the theory for the linear case given in Section 3 at each step of the iteration.

Suppose that $\bar{\eta}=[\bar{\eta}_1,\ldots,\bar{\eta}_n]'$ is a trial value of η which we want to improve at the next iterative step. Expanding $\exp(\eta_t)$ about $\bar{\eta}_t$ we have

$$\exp(\eta_t)=\exp(\bar{\eta}_t)+(\eta_t-\bar{\eta}_t)\exp(\bar{\eta}_t)=\exp(\bar{\eta}_t)(1-\bar{\eta}_t+\eta_t)$$

to a first approximation, on ignoring higher-order terms in $(\eta_t-\bar{\eta}_t)$. Writing $L=[l_1\ l_2\ \ldots\ l_n]$ we have

$$\begin{aligned} x-LN &= x-\sum_{t=1}^{n} l_t\exp(\eta_t) \\ &= x-\sum_{t=1}^{n} l_t\exp(\bar{\eta}_t)(1-\bar{\eta}_t+\eta_t) \\ &= x-L\bar{K}(1_n-\bar{\eta})-L\bar{K}\eta \\ &= \bar{x}-\bar{L}\eta \end{aligned}$$

where $\bar{K}$ is a diagonal matrix with elements $\exp(\bar{\eta}_1),\ldots,\exp(\bar{\eta}_n)$ on the main diagonal, $\bar{L}=L\bar{K}$ and $\bar{x}=x-\bar{L}(1_n-\bar{\eta})$, to a first approximation. We see that an appropriate linearised form of (4.3) is

$$\Omega^{-1}(\tilde{\eta}-\eta)+\bar{L}'\Sigma_e^{-1}\left(\bar{x}-\bar{L}\eta\right)=0,$$

which has the same form as (4.4) and therefore by (3.1) has the solution

$$\bar{\bar{\eta}}=\tilde{\eta}+\Omega\bar{L}'\left[\bar{L}\Omega\bar{L}'+\Sigma_e\right]^{-1}\left(\bar{x}-\bar{L}\tilde{\eta}\right) \tag{4.5}$$

The estimate $\bar{\bar{\eta}}$ is then taken as an improved approximation to $\hat{\eta}$ and is used as the starting value for the next iteration. The iterative procedure is initiated with $\bar{\eta}=\tilde{\eta}$ and is continued until adequate convergence is achieved. When the benchmark is binding we put $\Sigma_e=0$ in (4.5).

Since the final estimate $\hat{\eta}$ is not linear in y and x, it does not have the minimum mean square error linear property discussed at the end of Section 2. An approximation to the MSE matrix of $\hat{\eta}$ can be obtained by taking L equal to the final value of $\bar{L}$ in formula (3.3). The resulting estimate of N_t is $\exp\left(\hat{\eta}_t\right)$ with approximate MSE, using the standard form for the covariance of the lognormal distribution,

$$\begin{aligned} E\left(\hat{N}_t-N_t\right)\left(\hat{N}_s-N_s\right) &= \left\{\exp\left[E\left(\hat{\eta}_t-\eta_t\right)\left(\hat{\eta}_s-\eta_s\right)\right]-1\right\}\times \\ &\exp\left\{\hat{\eta}_t+\hat{\eta}_s+\frac{1}{2}\left[E\left(\hat{\eta}_t-\eta_t\right)^2+E\left(\hat{\eta}_s-\eta_s\right)^2\right]\right\} \end{aligned} \tag{4.6}$$

It is well known that if $X\sim N\left(\mu,\sigma^2\right)$ and $Y=\exp(X)$ then

$$E(Y)=\exp\left(\mu+\frac{1}{2}\sigma^2\right).$$

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

Now the constant which is a minimum mean square error estimator of Y is $E\,(Y)$. Also, $E\,(\eta \mid y, x) = \hat{\eta}_t$ approximately and we have an estimate $\hat{v}_t$ of $E\left(\hat{\eta}_t - \eta_t\right)^2$ obtained from (3.3) with $L = \bar{L}$. This suggests that $\exp\left(\hat{\eta}_t + \frac{1}{2}\hat{v}_t\right)$ is a better estimate of N_t than $\exp\left(\hat{\eta}_t\right)$. However, if the estimate $\exp\left(\hat{\eta}_t + \frac{1}{2}\hat{v}_t\right)$ is used instead of $\exp\left(\hat{\eta}_t\right)$ the matrix L will need to be modified. This can best be seen by considering the case in which the benchmark is binding. Let $\hat{N} = \left[\exp\left(\hat{\eta}_1\right), \ldots, \exp\left(\hat{\eta}_n\right)\right]'$ and $N^* = \left[\exp\left(\hat{\eta}_1 + \frac{1}{2}\hat{v}_1\right), \ldots, \exp\left(\hat{\eta}_n + \frac{1}{2}\hat{v}_n\right)\right]'$. The foregoing theory has provided us with an estimate $\hat{N}$ such that the benchmark relation $x = L\hat{N}$ is satisfied exactly. If, however, we wish to use N^*, we want the relation $x = LN^*$ to be satisfied exactly. This can be written

$$\begin{aligned} x &= \sum_{t=1}^{n} l_t \exp\left(\hat{\eta}_t + \frac{1}{2}\hat{v}_t\right) \\ &= \sum_{t=1}^{n} \exp\left(\frac{1}{2}\hat{v}_t\right) l_t \times \exp\left(\hat{\eta}_t\right) \\ &= \hat{L}\hat{N} \end{aligned}$$

where $\hat{L} = \left[\exp\left(\frac{1}{2}\hat{v}_1\right) l_1 \ \ \exp\left(\frac{1}{2}\hat{v}_2\right) l_2 \ \ \cdots \ \ \exp\left(\frac{1}{2}\hat{v}_n\right) l_n\right]$. Thus the result we desire can be achieved by replacing L in the above theory by $\hat{L}$. The same applies to the non-binding benchmark $x = LN + e$ where $e \neq 0$. The disadvantage of this approach lies in the computation of $\hat{v}_t$ which is known only at the end. However, initial values can be used and updated at each step in the iterative process described above.

For completeness we now consider the estimation of N_t by its own posterior mode. Taking $p\,(\eta \mid y)$ as the density of η given y, the density of N given y is

$$p\,(\log N \mid y)\,\frac{\partial\,(\eta_1, \ldots, \eta_n)}{\partial\,(N_1, \ldots, N_n)}.$$

Thus the log of the conditional density of N given y and x is

$$\begin{aligned} \log p\,(N \mid y, x) = \text{constant} - \sum_{t=1}^{n} \log N_t \ - \ & \frac{1}{2}\left[\left(\log N - \log \tilde{N}\right)' \Omega^{-1} \left(\log N - \log \tilde{N}\right)\right. \\ & \left. + (x - LN)'\,\Sigma_e^{-1}\,(x - LN)\right] \end{aligned} \qquad (4.7)$$

where $\log N = \left[\log N_1, \ldots, \log N_n\right]'$ and $\log \tilde{N} = \left[\log \tilde{N}_1, \ldots, \log \tilde{N}_n\right]'$ with $\tilde{N}_t = \exp(\tilde{\eta}_t)$ for $t = 1, \ldots, n$. To obtain the PME of N we differentiate this with respect to N_t and equate to zero for $t = 1, \ldots, n$ giving

$$\begin{aligned} \frac{\partial \log p}{\partial N_t} &= \frac{1}{N_t}\left[-1 + t^{th}\text{element of } \Omega^{-1}\left(\log \tilde{N} - \log N\right)\right] \\ & \quad + t^{th}\text{element of } L'\Sigma_e^{-1}\,(x - LN) = 0, \end{aligned}$$

which is equivalent to

$$-1 + t^{th}\text{element of } \Omega^{-1}\,(\tilde{\eta} - \eta) + \exp\,(\eta_t) \times t^{th}\text{element of } L'\Sigma_e^{-1}\,(x - LN) = 0.$$

The PME $\hat{N}^*$ of N is therefore $\hat{N}^* = \left[\hat{N}_1^*, \ldots, \hat{N}_n^*\right]'$ where $\hat{N}_t^* = \exp\left(\hat{\eta}_t^*\right)$ and where $\hat{\eta}^*$ is the solution of the equation

$$-1_n + \Omega^{-1}\,(\tilde{\eta} - \eta) + KL'\Sigma_e^{-1}\,(x - LN) = 0, \qquad (4.8)$$

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

where K is the same as in (4.3). As for (4.3) the linearised form of this is

$$-1_n + \Omega^{-1}(\tilde{\eta} - \eta) + \bar{L}'\Sigma_e^{-1}\left(\bar{x} - \bar{L}\eta\right) = 0,$$

the solution of which is

$$\hat{\eta}^{**} = \left[\Omega^{-1} + \bar{L}'\Sigma_e^{-1}\bar{L}\right]^{-1}\left[-1_n + \Omega^{-1}\tilde{\eta} + \bar{L}'\Sigma_e^{-1}\bar{x}\right]$$

We know from the solution (4.5) to the equation just above it that

$$\begin{aligned}\left[\Omega^{-1} + \bar{L}'\Sigma_e^{-1}\bar{L}\right]^{-1}\left[\Omega^{-1}\tilde{\eta} + \bar{L}'\Sigma_e^{-1}\bar{x}\right] &= \tilde{\eta} + \Omega\bar{L}'\left[\bar{L}\Omega\bar{L}' + \Sigma_e\right]^{-1}\left(\bar{x} - \bar{L}\tilde{\eta}\right) \\ &= \bar{\bar{\eta}}.\end{aligned}$$

If follows that

$$\begin{aligned}\hat{\eta}^{**} &= \tilde{\eta} - \Omega 1_n + \Omega\bar{L}'\left[\bar{L}\Omega\bar{L}' + \Sigma_e\right]^{-1}\left(\bar{x} - \bar{L}\tilde{\eta} + \bar{L}\Omega 1_n\right) \\ &= \bar{\bar{\eta}} - \left\{\Omega - \Omega\bar{L}'\left[\bar{L}\Omega\bar{L}' + \Sigma_e\right]^{-1}\bar{L}\Omega\right\}1_n, \qquad (4.9)\end{aligned}$$

where we have used the well known matrix inversion lemma

$$\left[\Omega^{-1} + \bar{L}'\Sigma_e^{-1}\bar{L}\right]^{-1} = \Omega - \Omega\bar{L}'\left[\bar{L}\Omega\bar{L}' + \Sigma_e\right]^{-1}\bar{L}\Omega.$$

Starting with a trial value of η, (4.9) can be used to iterate to the solution $\hat{\eta}^*$ of (4.8) and hence to $\hat{N}^*$, in the same way as with (4.5).

5 Estimation of Survey Bias

It is frequently the case that estimates obtained from surveys are known to be biased due to non-response and other factors. The existence of benchmarks free from bias enables us to estimate the bias in the survey observations. Assume that in the additive case considered in Sections 2 and 3, the observations y_t contain a constant additive bias b and in the multiplicative case considered in Section 4, the observations Y_t contain a constant multiplicative bias $B = \exp(b)$. Then in the additive case directly and in the multiplicative case after taking logs we have the relation

$$y_t = \eta_t + b + k_t u_t,$$

where the survey error $k_t u_t$ is assumed to have zero mean. Let $\breve{\eta}_t = \eta_t + b$. Then in both the additive and multiplicative cases, KFS applied to the y_t series provides us with the estimate $\tilde{\eta} = E(\breve{\eta} \mid y)$ where $\breve{\eta} = \left[\breve{\eta}_1, \ldots, \breve{\eta}_n\right]'$.

5.1 Estimation of Survey Bias Under the Additive Model

Assuming that the benchmarks x_i are bias-free, the benchmarking relations in the additive case are

$$x = L\eta + e = L\breve{\eta} - L1_n b + e.$$

The joint density of $\breve{\eta}$ and x given y is $p(\breve{\eta}, x \mid y) = p(\breve{\eta} \mid y)\, p(x \mid \eta, y)$ so

$$\begin{aligned}\log p(\breve{\eta}, x \mid y) &= \text{constant} - \frac{1}{2}\left[(\breve{\eta} - \tilde{\eta})'\,\Omega^{-1}(\breve{\eta} - \tilde{\eta})\right. \\ &\quad \left. + (x - L\breve{\eta} + L1_n b)'\,\Sigma_e^{-1}(x - L\breve{\eta} + L1_n b)\right]. \qquad (5.1)\end{aligned}$$

We can estimate b in terms of the unknown $\breve{\eta}$ by differentiating (5.1) with respect to b and equating

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

to zero, giving the estimate

$$\breve{b} = -\frac{1_n' L' \Sigma_e^{-1} (x - L\breve{\eta})}{1_n' L' \Sigma_e^{-1} L 1_n} \tag{5.2}$$

We now concentrate out b from (5.1) by replacing it by $\breve{b}$. The term $x - L\breve{\eta} + L1_n b$ becomes

$$\begin{aligned} x - L\breve{\eta} + L1_n\breve{b} &= x - L\breve{\eta} - \frac{L1_n 1_n' L' \Sigma_e^{-1} (x - L\breve{\eta})}{1_n' L' \Sigma_e^{-1} L 1_n} \\ &= \breve{x} - \breve{L}\breve{\eta} \end{aligned}$$

where

$$\breve{x} = \left[\mathrm{I}_l - \frac{L1_n 1_n' L' \Sigma_e^{-1}}{1_n' L' \Sigma_e^{-1} L 1_n} \right] x$$

and

$$\breve{L} = \left[\mathrm{I}_l - \frac{L1_n 1_n' L' \Sigma_e^{-1}}{1_n' L' \Sigma_e^{-1} L 1_n} \right] L.$$

The concentrated form of $\log p\,(\breve{\eta}, x \mid y)$ is therefore

$$\log p_c\,(\breve{\eta}, x \mid y) = \text{constant } - \frac{1}{2} \left[(\breve{\eta} - \tilde{\eta})' \, \Omega^{-1} \, (\breve{\eta} - \tilde{\eta}) + \left(\breve{x} - \breve{L}\breve{\eta} \right)' \Sigma_e^{-1} \left(\breve{x} - \breve{L}\breve{\eta} \right) \right].$$

But this is exactly the same as for the case $b = 0$ considered in Section 3 except that x is replaced by $\breve{x}$ and L is replaced by $\breve{L}$. It is straightforward to show, using the matrix inversion lemma, that the expression for $\hat{\eta}$ in (3.1) can be obtained by differentiating $\log p$ with respect to η and equating to zero, where

$$\log p = \text{constant } - \frac{1}{2} \left[(\eta - \tilde{\eta})' \, \Omega^{-1} \, (\eta - \tilde{\eta}) + (x - L\eta)' \, \Sigma_e^{-1} \, (x - L\eta) \right]. \tag{5.3}$$

In fact, (5.3) is the log posterior density of η given y and x in the unbiased case. Let $\breve{\eta}^* = E\,(\breve{\eta} \mid y, x)$. It follows from (3.1) and (3.3) that

$$\breve{\eta}^* = \tilde{\eta} + \Omega\breve{L}' \left[\breve{L}\Omega\breve{L}' + \Sigma_e \right]^{-1} \left(\breve{x} - \breve{L}\tilde{\eta} \right), \tag{5.4}$$

with MSE matrix,

$$\mathrm{MSE}(\breve{\eta}^*) = \Omega - \Omega\breve{L}' \left[\breve{L}\Omega\breve{L}' + \Sigma_e \right]^{-1} \breve{L}\Omega. \tag{5.5}$$

From (5.2) and (5.4) we obtain for the estimate of b,

$$\hat{b} = -\frac{1_n' L' \Sigma_e^{-1} (x - L\breve{\eta}^*)}{1_n' L' \Sigma_e^{-1} L 1_n} \tag{5.6}$$

To compute the MSE of $\hat{b}$ define $\breve{B} = \Omega\breve{L}' \left[\breve{L}\Omega\breve{L}' + \Sigma_e \right]^{-1}$. From (5.4) and the definition of $\breve{x}$ we find

$$x - L\breve{\eta}^* = \left\{ \mathrm{I}_l - L\breve{B} \left[\mathrm{I}_l - \frac{L1_n 1_n' L' \Sigma_e^{-1}}{1_n' L' \Sigma_e^{-1} L 1_n} \right] \right\} (x - L\tilde{\eta}).$$

Using $\mathrm{Var}(x - L\tilde{\eta}) = L\Omega L' + \Sigma_e$ we get

$$\mathrm{Var}(x - L\breve{\eta}^*) \;=\; \left\{ \mathrm{I}_l - L\breve{B} \left[\mathrm{I}_l - \frac{L1_n 1_n' L' \Sigma_e^{-1}}{1_n' L' \Sigma_e^{-1} L 1_n} \right] \right\}$$

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

$$\times \quad \left[L\Omega L' + \Sigma_e\right]$$
$$\times \quad \left\{ \mathrm{I}_l - L\breve{B}\left[\mathrm{I}_l - \frac{L1_n1_n'L'\Sigma_e^{-1}}{1_n'L'\Sigma_e^{-1}L1_n}\right]\right\} \tag{5.7}$$

It follows that

$$\mathrm{Var}(\hat{b}) = \frac{1_n'L'\Sigma_e^{-1}\mathrm{Var}(x - L\breve{\eta}^*)\Sigma_e^{-1}L1_n}{(1_n'L'\Sigma_e^{-1}L1_n)^2} \tag{5.8}$$

We can now test the significance of the bias by computing $t = \hat{b}/\mathrm{SE}(\hat{b})$ and treating t as approximately $N(0, 1)$. The covariance matrix between $\hat{b}$ and $\breve{\eta}^*$ is

$$\mathrm{Cov}(\hat{b}, \breve{\eta}^*) = \frac{1_n'L'\Sigma_e^{-1}L\mathrm{MSE}(\breve{\eta}^*) - 1_n'\breve{L}'\Sigma_e^{-1}\breve{L}\mathrm{MSE}(\breve{\eta}^*)}{1_n'L'\Sigma_e^{-1}L1_n}. \tag{5.9}$$

The estimate of η is therefore

$$\hat{\eta} = \breve{\eta}^* - \hat{b}1_n, \tag{5.10}$$

with MSE matrix

$$\mathrm{MSE}(\hat{\eta}) = \mathrm{MSE}(\breve{\eta}^*) + 1_n1_n'\mathrm{Var}(\hat{b}) - \mathrm{Cov}(\hat{b}, \breve{\eta}^*)'1_n' - 1_n\mathrm{Cov}(\hat{b}, \breve{\eta}^*). \tag{5.11}$$

For the case of a binding benchmark, when $\Sigma_e = 0$, put $\Sigma_{\acute{e}} = \sigma_e^2\mathrm{I}_l$ and let $\sigma_e^2 \to 0$. The matrix $\left[\mathrm{I}_l - \frac{L1_n1_n'L'}{1_n'L'L1_n}\right]$ is then non invertible and we take $\left[\breve{L}\Omega\breve{L}'\right]^-$ as a generalized inverse of $\breve{L}\Omega\breve{L}'$. Equations (5.4) to (5.7) and (5.9) then reduce to

$$\breve{\eta}^* = \tilde{\eta} + \Omega\breve{L}'\left[\breve{L}\Omega\breve{L}'\right]^-\left(\breve{x} - \breve{L}\tilde{\eta}\right), \tag{5.12}$$

$$\mathrm{MSE}(\breve{\eta}^*) = \Omega - \Omega\breve{L}'\left[\breve{L}\Omega\breve{L}'\right]^-\breve{L}\Omega, \tag{5.13}$$

$$\hat{b} = -\frac{1_n'L'(x - L\breve{\eta}^*)}{1_n'L'L1_n}, \tag{5.14}$$

$$\mathrm{Var}(x - L\breve{\eta}^*) = \left\{\mathrm{I}_l - L\breve{B}\left[\mathrm{I}_l - \frac{L1_n1_n'L'}{1_n'L'L1_n}\right]\right\} L\Omega L'$$
$$\times \left\{\mathrm{I}_l - L\breve{B}\left[\mathrm{I}_l - \frac{L1_n1_n'L'}{1_n'L'L1_n}\right]\right\}', \tag{5.15}$$

$$\mathrm{Cov}(\hat{b}, \breve{\eta}^*) = \frac{1_n'L'L\mathrm{MSE}(\breve{\eta}^*) - 1_n'\breve{L}'\breve{L}\mathrm{MSE}(\breve{\eta}^*)}{1_n'L'L1_n}, \tag{5.16}$$

with $\breve{B}$ now defined as $\Omega\breve{L}'\left(\breve{L}\Omega\breve{L}'\right)^-$.

In principle it would have been possible at the outset to have applied the bias correction to $\tilde{\eta}$ instead of η by putting $\tilde{\eta} = \eta + b1_n$ in (5.1). This gives

$$\log p(\eta, x \mid y) = \text{constant} - \frac{1}{2}\left[(\eta - \tilde{\eta} + b1_n)'\,\Omega^{-1}(\eta - \tilde{\eta} + b1_n)\right.$$
$$\left. + (x - L\eta)'\,\Sigma_e^{-1}(x - L\eta)\right] \tag{5.17}$$

Differentiating (5.17) with respect to b and equating to zero gives as the estimate of b in terms of the unknown η,

$$\tilde{b} = -\frac{1_n'\Omega^{-1}(\eta - \tilde{\eta})}{1_n'\Omega^{-1}1_n}. \tag{5.18}$$

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

We now concentrate out b from (5.17) by replacing it by $\tilde{b}$. The term $\eta - \tilde{\eta} + b1_n$ becomes

$$\begin{aligned}
\eta - \tilde{\eta} + b1_n &= \eta - \tilde{\eta} - \frac{1_n 1_n' \Omega^{-1} (\eta - \tilde{\eta})}{1_n' \Omega^{-1} 1_n} \\
&= \left[\mathrm{I}_n - \frac{1_n 1_n' \Omega^{-1}}{1_n' \Omega^{-1} 1_n} \right] (\eta - \tilde{\eta}) \\
&= M_\Omega (\eta - \tilde{\eta}) \text{ say.}
\end{aligned}$$

The concentrated form of (5.17) is therefore

$$\begin{aligned}
\log p_c (\eta, x \mid y) &= \text{constant} - \frac{1}{2} \left[(\eta - \tilde{\eta})' M_\Omega' \Omega^{-1} M_\Omega (\eta - \tilde{\eta}) \right. \\
&\quad \left. + (x - L\eta)' \Sigma_e^{-1} (x - L\eta) \right] \qquad (5.19)
\end{aligned}$$

It follows from (5.19) that

$$\hat{\eta} = \left[M_\Omega' \Omega^{-1} M_\Omega + L' \Sigma_e^{-1} L \right]^{-1} \left[M_\Omega' \Omega^{-1} M_\Omega \tilde{\eta} + L' \Sigma_e^{-1} x \right]. \qquad (5.20)$$

We see that $\tilde{b}$ from (5.18)and $\hat{\eta}$ from (5.20) require the inversion of the $n \times n$ matrix Ω where n is at least $12l$ when the observations are monthly and the benchmarks annual, while $\hat{b}$ from (5.6) requires only the inversion of the $l \times l$ matrix Σ_e; also, (5.20) cannot be used for binding benchmarks; finally, we cannot express $\hat{\eta}$ in (5.20) as the sum of $\tilde{\eta}$ and a correction factor as is done in (5.4) because $|M_\Omega| = 0$. We conclude that the former approach is preferable.

It is straightforward to extend the treatment of the bias to the situation where the series is divided into segments with the bias assumed to be constant within the segments but different between the segments.

5.2 *Estimation of Survey Bias Under the Multiplicative Model*

The benchmarking relations in the multiplicative case are

$$\begin{aligned}
x &= LN + e \\
&= L \exp(\eta) + e \\
&= L \exp(\check{\eta} - 1_n b) + e \\
&= L \exp(\check{\eta}) \exp(-b) + e \\
&= L\check{N} \exp(-b) + e,
\end{aligned}$$

where $\check{N} = \exp(\check{\eta})$. As in (5.1), the log joint density of $\check{\eta}$ and x given y is

$$\begin{aligned}
\log p (\check{\eta}, x \mid y) &= \text{constant} - \frac{1}{2} \left[(\check{\eta} - \tilde{\eta})' \Omega^{-1} (\check{\eta} - \tilde{\eta}) \right. \\
&\quad \left. + \left(x - L\check{N} \exp(-b) \right)' \Sigma_e^{-1} \left(x - L\check{N} \exp(-b) \right) \right] \qquad (5.21)
\end{aligned}$$

We can estimate the bias in terms of the unknown $\check{N}$ by differentiating (5.21) with respect to b and equating to zero, giving the estimate of $\exp(-b)$ as

$$\exp(-\check{b}) = \frac{\check{N}' L' \Sigma_e^{-1} x}{\check{N}' L' \Sigma_e^{-1} L \check{N}} \qquad (5.22)$$

The concentrated form of (5.21) is thus

$$\log p_c (\check{\eta}, x \mid y) = \text{constant} - \frac{1}{2} \left[(\check{\eta} - \tilde{\eta})' \Omega^{-1} (\check{\eta} - \tilde{\eta}) \right.$$

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

$$+\left(x - L\breve{N}\exp(-\breve{b})\right)' \Sigma_e^{-1}\left(x - L\breve{N}\exp(-\breve{b})\right)\Bigg] \qquad (5.23)$$

To get the PME of $\breve{\eta}_t$ it is necessary to differentiate (5.23) with respect to $\breve{\eta}_t$, equate the resulting equation to zero for $t = 1, \dots, n$, and solve for $\breve{\eta}_t$. In section 4, we linearised the derivatives around a trial value and solved the resulting equations to find an improved solution. Here, we use a slightly different but equivalent formulation. We first linearise the nonlinear factors $x - L\breve{N}\exp(-\breve{b})$ in the likelihood around a trial value of $\breve{\eta}$ and maximise the resulting linearised expression with respect to $\breve{\eta}$. The process is continued until adequate convergence is achieved.

Let $\acute{\eta} = \left[\acute{\eta}_1, \dots, \acute{\eta}_n\right]'$ be a trial value of $\breve{\eta}$ and let $\acute{N} = \exp(\acute{\eta})$. Expanding $\exp(\breve{\eta}_t)$ about $\acute{\eta}_t$ we have

$$\exp(\breve{\eta}_t) = \exp(\acute{\eta}_t)(1 - \acute{\eta}_t + \breve{\eta}_t)$$

to a first approximation. A first approximation to $x - L\breve{N}\exp(-\breve{b})$ is

$$\begin{aligned} x - L\breve{N}\exp(-\breve{b}) &= x - L\acute{K}\exp(-\acute{b})\left(1_n - \acute{\eta} + \breve{\eta}\right) \\ &= x - L\acute{K}\exp(-\acute{b})\left(1_n - \acute{\eta}\right) - L\acute{K}\exp(-\acute{b})\breve{\eta} \\ &= \acute{x} - \acute{L}\breve{\eta} \end{aligned}$$

where $\acute{K}$ is a diagonal matrix with $\exp(\acute{\eta})$ on the main diagonal and

$$\begin{aligned} \exp(-\acute{b}) &= \frac{\acute{N}'L'\Sigma_e^{-1}x}{\acute{N}'L'\Sigma_e^{-1}L\acute{N}}, \\ \acute{L} &= L\acute{K}\exp(-\acute{b}), \\ \acute{x} &= x - \acute{L}\left(1_n - \acute{\eta}\right). \end{aligned}$$

It follows that our linearised form of (5.23) is

$$\log p_c(\breve{\eta}, x \mid y) = \text{constant } -\frac{1}{2}\left[(\breve{\eta} - \tilde{\eta})'\Omega^{-1}(\breve{\eta} - \tilde{\eta}) + \left(\acute{x} - \acute{L}\breve{\eta}\right)' \Sigma_e^{-1}\left(\acute{x} - \acute{L}\breve{\eta}\right)\right], \qquad (5.24)$$

which is maximised at

$$\acute{\breve{\eta}} = \tilde{\eta} + \Omega\acute{L}'\left[\acute{L}\Omega\acute{L}' + \Sigma_e\right]^{-1}\left(\acute{x} - \acute{L}\tilde{\eta}\right). \qquad (5.25)$$

The value of $\acute{\breve{\eta}}$ is then taken as an improved approximation to $\breve{\eta}^*$ and is used as the starting value for the next iteration. The iterative process is continued until convergence. The MSE matrix of $\breve{\eta}^*$ is approximated by formula (3.3) with L replaced by $\acute{L}$ giving

$$\text{MSE}(\breve{\eta}^*) = \Omega - \Omega\acute{L}'\left[\acute{L}\Omega\acute{L}' + \Sigma_e\right]^{-1}\acute{L}\Omega. \qquad (5.26)$$

The resulting final estimate $\breve{N}^*$ of $\breve{N}$ is $\exp(\breve{\eta}^*)$ whose approximate MSE matrix has its (t,s)-th element given by (4.6) with $\hat{N}$ and $\hat{\eta}$ replaced by $\breve{N}^*$ and $\breve{\eta}^*$.

The final estimate of the bias is

$$\hat{B}(\breve{N}^*) = \frac{(\breve{N}^*)'L'\Sigma_e^{-1}L\breve{N}^*}{(\breve{N}^*)'L'\Sigma_e^{-1}x},$$

and the final estimate of N_t is

$$\hat{N} = \frac{\breve{N}^*}{\hat{B}(\breve{N}^*)}.$$

The MSE of $\hat{B}(\breve{N}^*)$ and $\hat{N}$ are approximated by linearizing the estimates. Expanding $\hat{B}(\breve{N}^*)$

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

around $\breve{N}$ we get

$$\hat{B}(\breve{N}^*) - \hat{B}(\breve{N}) = \sum_{t=1}^{n} \frac{\partial \hat{B}(\breve{N}^*)}{\partial \breve{N}_t^*} \left(\breve{N}_t^* - \breve{N}_t\right), \tag{5.27}$$

from which we deduce that, approximately,

$$\operatorname{Var}\left(\hat{B}\right) = E\left(\hat{B}(\breve{N}^*) - \hat{B}(\breve{N})\right)^2 = \left[\frac{\partial \hat{B}(\breve{N}^*)}{\partial \breve{N}^*}\right] \operatorname{MSE}(\breve{N}^*) \left[\frac{\partial \hat{B}(\breve{N}^*)}{\partial \breve{N}^*}\right]', \tag{5.28}$$

where

$$\left[\frac{\partial \hat{B}(\breve{N}^*)}{\partial \breve{N}^*}\right] = \left[\frac{\partial \hat{B}(\breve{N}^*)}{\partial \breve{N}_1^*}, \dots, \frac{\partial \hat{B}(\breve{N}^*)}{\partial \breve{N}_n^*}\right]'. \tag{5.29}$$

The approximate MSE of $\hat{N}$ is

$$\operatorname{MSE}(\hat{N}) = \operatorname{MSE}\left(\frac{\breve{N}^*}{\hat{B}}\right) = \frac{1}{\hat{B}^2}\left\{\hat{N}\hat{N}'\operatorname{Var}(\hat{B}) - 2\hat{N}\operatorname{Cov}(\hat{B}, \breve{N}^*) + \operatorname{MSE}(\breve{N}^*)\right\}, \tag{5.30}$$

where we have written $\hat{B}$ for $\hat{B}(\breve{N}^*)$ and where, using (5.27), we get

$$\operatorname{Cov}(\hat{B}, \breve{N}^*) = \left[\frac{\partial \hat{B}(\breve{N}^*)}{\partial \breve{N}}\right] \operatorname{MSE}(\breve{N}^*). \tag{5.31}$$

To compute (5.28), (5.30) and (5.31) we need to evaluate (5.29), the derivatives of the estimator of the bias with respect to $\breve{N}_t^*, t = 1, \dots, n$. In practice, it is easier to use numerical derivatives because $\hat{B}(\breve{N}^*)$ is a complicated function of $\breve{N}_t^*, t = 1, \dots, n$.

6 Single-Stage Benchmarking

In this section we develop the second method of benchmarking mentioned at the beginning of Section 3. We assume that we have the same set-up as in Section 3 but instead of analysing the observations y_t first to give $\tilde{\eta}$ and then combining this with the benchmarks x_i to produce $\hat{\eta}$, we arrange the data into a single series and obtain $\hat{\eta}$ by a single filtering and smoothing operation. The series is constructed by inserting each benchmark into the monthly series immediately after the last monthly value to which the benchmark refers. For simplicity of presentation our treatment will be based on the case where the benchmarks start in the first year and they consist either of calendar yearly totals or accurate December values. The series is arranged in the form

$$y_1, \ \dots, \ y_{12}, \ x_1, \ y_{13}, \ \dots, \ y_{24}, \ x_2, \ y_{25}, \ \dots .$$

If a benchmark is available for the final year the series ends with x_l; otherwise it ends with y_n.

Let us consider the model defined by (2.1) to (2.5) with u_t an AR(1) but in place of (2.6) we assume that the trading day coefficients are updated only once a year each January; this will still suffice in most cases and it provides a considerable saving in the length of the state vector. Thus we assume that

$$\delta_{j,12i+1} = \delta_{j,12i} + \zeta_{j,12i+1}, \qquad j = 1, \dots, k, \qquad i = 1, \dots, l^*, \qquad \delta_{j,t} = \delta_{j,t-1} \text{ otherwise},$$

where l^* is the last year for which a January value is available. Let us regard the time point in the series at which benchmark x_i occurs as $t = (12i)'$; thus the point $t = (12i)'$ occurs in the series between $t = 12i$ and $t = 12i + 1$. We re-define the state vector as

$$\alpha_t = \left[\mu_t, \ \dots, \ \mu_{t-11}, \ \gamma_t, \dots, \gamma_{t-11}, \ \delta_{1t}, \dots, \delta_{kt}, \ \epsilon_t, \dots, \epsilon_{t-11}, \ u_t\right]' \tag{6.1}$$

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

for $t = 1, \ldots, n$ and $\alpha_t \equiv \alpha_{12i}$ for $t = (12i)'$, $\quad i = 1, \ldots, l$. The transition matrix T_t is a block diagonal matrix with block elements

$$\begin{bmatrix} 2 & \text{-1} & 0_{1,10} \\ \multicolumn{2}{c}{I_{11}} & 0_{11,1} \end{bmatrix}, \begin{bmatrix} 1'_{11} & 0_{1,1} \\ I_{11} & 0_{11,1} \end{bmatrix}, I_k, \begin{bmatrix} \multicolumn{2}{c}{0_{1,12}} \\ I_{11} & 0_{11,1} \end{bmatrix}, \phi \tag{6.2}$$

for $t = 1, \ldots, n$ and is I_{37+k} when $t = (12i)'$, $\quad i = 1, \ldots, l$. The state error is $\nu_t = [\xi_t,\ \omega_t,\ \epsilon_t,\ \chi_t]$, $t = 12i + 2, \ldots, 12(i + 1)$, $\nu_t = [\xi_t,\ \omega_t,\ \zeta_{1t}, \ldots, \zeta_{kt},\ \epsilon_t,\ \chi_t]'$, $t = 12i + 1$, $i = 1, \ldots, l^*$ and $\nu_t = 0$ for $t = (12i)'$, $\quad i = 1, \ldots, l$. We take

$$\tilde{Z}_t = \begin{bmatrix} 1 \ 0_{1,11} \ 1 \ 0_{1,11} \ w_{1t} \ \ldots \ w_{kt} \ 1 \ 0_{1,11} \ \ k_t \end{bmatrix}, \tag{6.3}$$

$$Z_t = \begin{bmatrix} 1 \ 0_{1,11} \ 1 \ 0_{1,11} \ w_{1t} \ \ldots \ w_{kt} \ 1 \ 0_{1,11} \ 0 \end{bmatrix} \quad t = 1, \ldots, n, \tag{6.4}$$

and

$$\tilde{Z}_t = \begin{bmatrix} \tilde{l}_i \ \tilde{l}_i \ \tilde{l}_i w_{1(i)} \ \ldots \ \tilde{l}_i w_{k(i)} \ \tilde{l}_i \ 0 \end{bmatrix}, \quad t = (12i)', \quad i = 1, \ldots, l \tag{6.5}$$

where $\tilde{l}_i = \begin{bmatrix} l_{i,12i} \ldots l_{i,12i-11} \end{bmatrix}$ with $l_{i,t}$ defined as the (it)th element of L, and $w_{j(i)} = \begin{bmatrix} w_{j,12i}, \ \ldots, \ w_{j,12i-11} \end{bmatrix}'$ for $j = 1, \ldots, k$ and $i = 1, \ldots, l$. We therefore have $y_t = \tilde{Z}_t\alpha_t$, $\eta_t = Z_t\alpha_t$, $t = 1, \ldots, n$ and $x_i = \tilde{Z}_t\alpha_t + e_i$, $t = (12i)'$, $\quad i = 1, \ldots, l$. We use a modified form of the state space model introduced in section 2 in which the error variance of the observation equation is zero for $t = 1, \ldots, n$ and is the variance of the benchmark for $t = (12i)'$, $\quad i = 1, \ldots, l$. We are assuming that Σ_e is a diagonal matrix since otherwise the state vector becomes too large. The state transition equation is

$$\alpha_t = T_t\alpha_{t-1} + R_t\nu_t, \qquad t = 1, \ldots, 12, (12)', 13, \ldots$$

where R_t is a selection matrix which is made up of appropriate columns of I_{37+k}.

When the survey error u_t is ARMA(p,q) and not AR(1) we proceed by replacing k_t in (6.3), 0 in (6.4) and (6.5), u_t in (6.1) and ϕ in (6.2) by the appropriate vectors and matrix as indicated in Section 2 below equation (2.10).

The model is now in standard state space form so the required benchmarked values $\hat{\eta}_t$ can be obtained by standard Kalman filtering and smoothing for $t = 1, \ldots, 12, (12)', 13, \ldots, 24, (24)', 25, \ldots$. A referee asked if this formulation was anything more than a routine application of standard material in Harvey's (1989) book. We believe it is. Our starting point was that it is not practical just to define a new state vector equal to a stacked form of the twelve state vectors for the twelve months in a particular year and then try to apply the KFS in a routine fashion. We found that this was unmanageable. We therefore set out to employ as much ingenuity as we could muster to reduce the size of the composite model as much as possible. This led to unorthodox devices such as putting the observation error into the state vector.

The most important case in practice is that in which the benchmark is an annual total, $x_i = \sum_{t=12i-11}^{12i} \eta_t$. Here, the solution is obtained by substituting $\tilde{l}_i = 1'_{12}$ in the general formulae. A second important case is that in which the benchmark is a particular monthly value, say December. For this case some simplification is possible by taking

$$\begin{aligned}
\alpha_t &= \begin{bmatrix} \mu_t,\ \mu_{t-1},\ \gamma_t,\ \ldots,\ \gamma_{t-10},\ \delta_{1t},\ \ldots,\ \delta_{kt},\ \varepsilon_t,\ \varepsilon_{t-1},\ u_t \end{bmatrix}' \\
\tilde{Z}_t &= \begin{bmatrix} 1 \ 0 \ 1 \ 0_{1,10} \ w_{1t} \ \ldots \ w_{kt} \ 1 \ 0 \ k_t \end{bmatrix} \\
Z_t &= \begin{bmatrix} 1 \ 0 \ 1 \ 0_{1,10} \ w_{1t} \ \ldots \ w_{kt} \ 1 \ 0 \ 0 \end{bmatrix}, \qquad t = 1, \ldots, n \\
\tilde{Z}_t &= \big[l_{i,12i} \ 0 \ l_{i,12i} \ 0_{1,10} \ l_{i,12i} w_{1,12i} \ \ldots \\
&\qquad l_{i,12i} w_{k,12i} \ l_{i,12i} \ 0 \ 0 \big], \qquad t = (12i)', \ i = 1, \ldots, l \\
T_t &= \text{a diagonal matrix with blocks}
\end{aligned}$$

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

$$\begin{aligned}
& \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} \multicolumn{2}{c}{-1'_{11}} \\ \mathrm{I}_{10} & 0_{10,1} \end{bmatrix}, \mathrm{I}_k, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \phi, \qquad t = 1, \dots, n \\
T_t &= \mathrm{I}_{16+k}, \quad t = (12i)', \quad i = 1, \dots, l \\
x_i &= \eta_{12i} + e_i, \quad i = 1, \dots, l.
\end{aligned}$$

Estimation of the bias is easier in single-stage benchmarking than in two-stage benchmarking. All that needs to be done is add a term b_t to the observation equation for $t = 1, \dots, n$, exclude it for $t = (12i)'$, include the term b_t in the state vector and add the equation $b_t = b_{t-1}$ to the state transition equation.

7 Single-Stage Benchmarking for the Multiplicative Model

We now consider single-stage benchmarking for the case where the monthly observations behave multiplicatively but the benchmarking relations are additive. As in section 4, the observation model is

$$Y_t = N_t U_t, \qquad t = 1, \dots, n$$

where $N_t = \exp(\eta_t)$ and U_t is the multiplicative sampling error. On taking logs we have

$$y_t = \eta_t + k_t u_t$$

where k_t and u_t are as before. Similarly, the benchmark constraints are

$$x = LN + e. \tag{7.1}$$

Our approach will be to obtain by iteration the PME's of $\alpha_1, \dots, \alpha_n$ such that the constraints (7.1) are satisfied.

In order to develop the method we first consider the PME approach to the additive model of section 2 in which (7.1) is replaced by

$$x = L\eta + e = [l_1 \ \dots \ l_n\,][\eta_1, \dots, \eta_n]' + e \tag{7.2}$$

where $\eta_t = Z_t \alpha_t$. The log joint density of $\alpha_1, \dots, \alpha_n, y_1, \dots, y_n$ and $x_1, \dots, x_l$ given α_0 is, on noting that $R_t Q_t R_t'$ and $R_t Q_t^{-1} R_t'$ are Moore–Penrose generalised inverses of each other, as is easily confirmed,

$$\begin{aligned}
\log p &= \text{constant} - \frac{1}{2} \sum_{t=1}^{n} \left[(\alpha_t - T_t \alpha_{t-1})' R_t Q_t^{-1} R_t' (\alpha_t - T_t \alpha_{t-1}) \right. \\
&+ \left. \left(y_t - \tilde{Z}_t \alpha_t\right)' H_t^{-1} \left(y_t - \tilde{Z}_t \alpha_t\right) \right] \\
&- \frac{1}{2} \left(x - \sum_{t=1}^{n} l_t Z_t \alpha_t \right)' \Sigma_e^{-1} \left(x - \sum_{t=1}^{n} l_t Z_t \alpha_t \right)
\end{aligned}$$

Here we have put $\mathrm{Var}\left(y - \tilde{Z}_t \alpha_t\right)$ equal to the nonsingular matrix H_t which later we can allow to converge to a null matrix, and similarly for Σ_e. Differentiating with respect to α_t and equating to zero gives the equations

$$\begin{aligned}
-R_t Q_t^{-1} R_t' (\alpha_t - T_t \alpha_{t-1}) &+ T_{t+1}' R_{t+1} Q_{t+1} R_{t+1}' (\alpha_{t+1} - T_{t+1} \alpha_t) + \tilde{Z}_t' H_t^{-1} \left(y_t - \tilde{Z}_t \alpha_t\right) \\
&+ Z_t' l_t' \Sigma_e^{-1} \left(x - \sum_{t=1}^{n} l_t Z_t \alpha_t \right) = 0
\end{aligned} \tag{7.3}$$

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

for $t = 1, \ldots, n-1$ together with the equation

$$-R_nQ_n^{-1}R_n'\left(\alpha_n - T_n\alpha_{n-1}\right) + \tilde{Z}_n'H_n^{-1}\left(y_n - \tilde{Z}_n\alpha_n\right) + Z_n'l_n'\Sigma_e^{-1}\left(x - \sum_{t=1}^{n} l_tZ_t\alpha_t\right) = 0,$$

which we shall ignore from now on for simplicity of discussion since it is easily dealt with. The solution of equations (7.3) is the PME of $\alpha = \left[\alpha_1', \ldots, \alpha_n'\right]'$ since the derivative of the log conditional density of α given y and x is the same as the derivative of the log joint density. But since all variables are normally distributed and equations (7.3) are linear the solution is also normal. But modes of normal distributions are equal to means. It follows that the solution of (7.3) is the same as the value $\hat{\alpha} = E\left(\alpha|y, x\right)$ obtained by KFS in section 6.

We now apply the PME approach to the multiplicative case (7.1). The log joint density is

$$\begin{aligned}\log p = \text{constant} \quad & - \quad \frac{1}{2}\sum_{t=1}^{n}\left[\left(\alpha_t - T_t\alpha_{t-1}\right) R_tQ_t^{-1}R_t'\left(\alpha_t - T_t\alpha_{t-1}\right)\right. \\ & + \quad \left.\left(y_t - \tilde{Z}_t\alpha_t\right)' H_t^{-1}\left(y_t - \tilde{Z}_t\alpha_t\right)\right] \\ & - \quad \frac{1}{2}\left(x - \sum_{t=1}^{n} l_t\exp(\eta_t)\right)' \Sigma_e^{-1}\left(x - \sum_{t=1}^{n} l_t\exp(\eta_t)\right)\end{aligned}$$

Differentiating with respect to α_t and equating to zero gives the equations

$$\begin{aligned}-R_tQ_t^{-1}R_t'\left(\alpha_t - T_t\alpha_{t-1}\right) \quad & + \quad T_{t+1}'R_{t+1}Q_{t+1}R_{t+1}'\left(\alpha_{t+1} - T_{t+1}\alpha_t\right) + \tilde{Z}_t'H_t^{-1}\left(y_t - \tilde{Z}_t\alpha_t\right) \\ & + \quad \exp\left(\eta_t\right) Z_t'l_t'\Sigma_e^{-1}\left(x - \sum_{t=1}^{n} l_t\exp(\eta_t)\right) = 0 \qquad (7.4)\end{aligned}$$

for $t = 1, \ldots, n-1$ together with an analogous equation for $t = n$ which, as in the additive case, we ignore. These equations are nonlinear. As in section 4 we solve them by linearising and putting the linearised equations into the same form as the additive equations (7.3) so that we can solve the linearised equations by KFS.

Let $\tilde{\alpha}_t$ be a trial value of α_t for $t = 1, \ldots, n$ and let $\tilde{\eta}_t = Z_t\tilde{\alpha}_t$. Expanding $\exp\left(\eta_t\right)$ about $\tilde{\eta}_t$ gives $\exp\left(\eta_t\right) = \exp\left(\tilde{\eta}\right) + \left(\eta_t - \tilde{\eta}_t\right)\exp\left(\tilde{\eta}_t\right) = \exp\left(\tilde{\eta}_t\right)\left(1 - \tilde{\eta}_t + \eta_t\right)$ to a first approximation. Using this approximation the final term of (7.4) becomes

$$\exp\left(\tilde{\eta}_t\right) Z_t'l_t'\Sigma_e^{-1}\left(x - \sum_{t=1}^{n} l_t\exp\left(\tilde{\eta}_t\right)\left(1 - \tilde{\eta}_t + \eta_t\right)\right) = Z_t'\bar{l}_t'\Sigma_e^{-1}\left(\bar{x} - \sum_{t=1}^{n}\bar{l}_t\eta_t\right) \qquad (7.5)$$

where $\bar{x} = x - \sum_{t=1}^{n} l_t\exp\left(\tilde{\eta}_t\right)\left(1 - \tilde{\eta}_t\right)$ and $\bar{l}_t = l_t\exp\left(\tilde{\eta}_t\right)$. We observe that $\bar{x}$ and $\bar{l}_t$ have the same form as in section 4. Substituting for the last term of (7.4) we see that the linearised equations have the same form as in (7.3) and so can be solved by the KFS. An improved value of α_t and hence η_t can therefore be obtained by means of the theory of section 6 provided that x and l_t are replaced by $\bar{x}$ and $\bar{l}_t$. This is taken as a new trial value and the process is repeated until satisfactory convergence has been achieved. The iterative procedure is initialised with $\tilde{\alpha} = E\left(\alpha_t|y\right)$ and is stopped when the maximum absolute difference between successive iterations is smaller than a preassigned value. The process will be slower than for the two-stage method since the series has to be put through the KFS at each iteration. Denoting the final value of α_t by $\hat{\alpha}_t$, we take as our final monthly values $\hat{N}_t = \exp\left(\hat{\eta}_t\right)$, where $\hat{\eta}_t = Z_t\hat{\alpha}_t$ for $t = 1, \ldots, n$. It would be possible to develop an approximate modified estimate of the form $\exp\left(\hat{\eta}_t + \frac{1}{2}\hat{v}_t\right)$ analogous to that given in section 4 but we shall not pursue this here.

Bias is estimated exactly as in the additive single-stage case by including the term b_t in the

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

observation equation and state vector and adding the equation $b_t = b_{t-1}$ in the state transition equation for $t = 1, \ldots, n$.

8 Application to Canadian Retail Sales Data

The methods developed earlier will now be illustrated by applying them to monthly values of the Canadian Retail Trade Sales series from January 1980 to December 1989. These values are set out in Table A1 of the Appendix. The first four benchmarks are annual totals from February to the following January 1985–88. Retail trade data are totalled in this way because of the special nature of the transition from December to January in the retail trade. The last three benchmarks are values for October, November and December 1989 observed in a separate and redesigned monthly survey intended to replace the previous one and they form part of a linking exercise. The benchmark values are set out in Table A2 of the Appendix. Sources and other features of the data are discussed in Mian & Laniel (1993). The monthly data and the benchmarks are displayed in Figure 1.

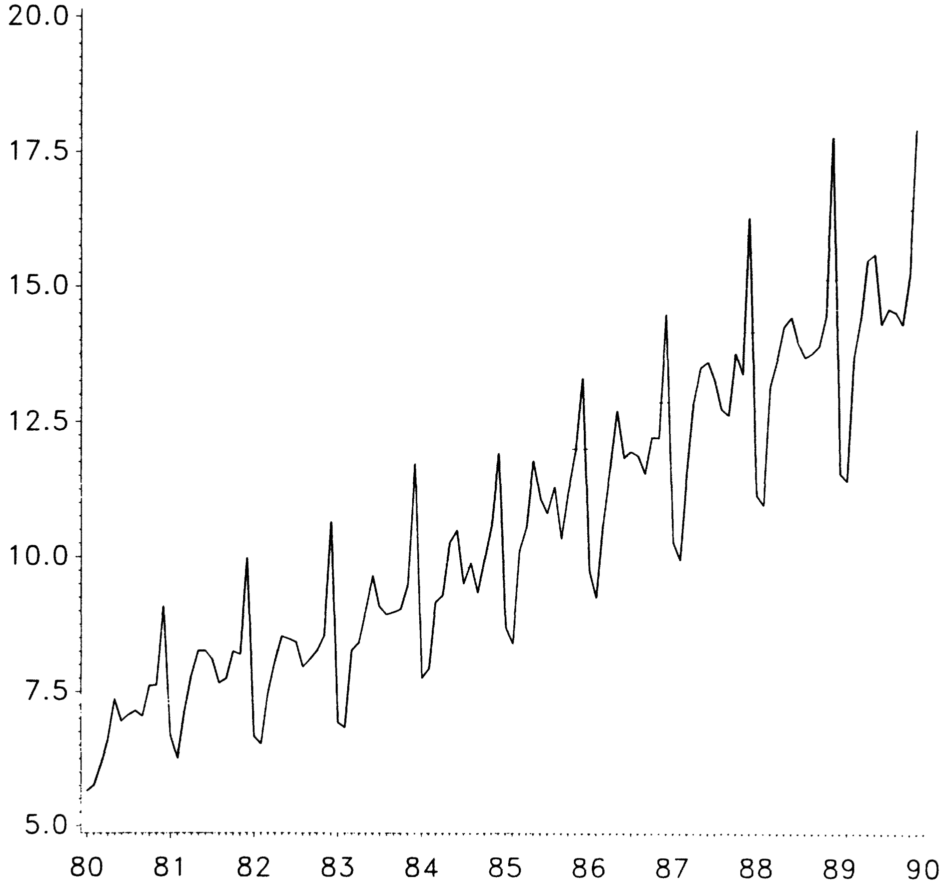

Figure 1. *Canadian retail sales series from January 1980 to December 1989 and average values of the benchmarks, graphed as horizontal lines. Values are expressed in millions.*

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

Table 1

Number of estimators available for the combination of options in benchmarking. The abbreviations 1-st. stands for single-stage and 2-st. for two-stage.

Benchmarking Methods											
Additive						Multiplicative					
Binding			Non-Binding			Binding			Non-Binding		
No Bias		Bias	No Bias		Bias	No Bias		Bias	No Bias		Bias
1-st.	2-st.	2-st.	1-st.	2-st.	2-st.	1-st.	2-st.	2-st.	1-st.	2-st.	2-st.
1	1	1	1	1	1	1	3	1	1	3	1

The graph of the monthly data shows an upward trend with increasing amplitude, which is typical behaviour of series that behave multiplicatively. Also, a comparison of the data with the benchmarks demonstrates clearly that the observations are negatively biased. These characteristics suggest strongly that the most suitable model for benchmarking these data is the multiplicative model with bias. However, since our objective is illustrative we have in fact analysed the data with the methods we developed. The various combinations of options we considered are set out in Table 1. For example, to estimate the underlying true value N_t under the multiplicative + binding + no bias + two-stage combination there are three methods: (i) compute the posterior mode $\hat{\eta}_t$ of $\eta_t = \log(N_t)$ and estimate N_t by $\exp(\hat{\eta}_t)$, (ii) compute $\hat{\eta}_t$ and its variance $\hat{v}_t$ and estimate N_t by $\exp(\hat{\eta}_t + \frac{1}{2}\hat{v}_t)$, and (iii) estimate N_t by its own posterior mode.

We modelled the survey error term u_t in (2.7) by the seasonal autoregressive model

$$(1 - 0.9387B)\left(1 - 0.8927B^{12}\right)u_t = \chi_t, \tag{8.1}$$

with $\text{var}(\chi_t)$ chosen to make $\text{var}(u_t) = 1$. We calculated the coefficients of the model from the autocorrelations of the survey errors $k_t u_t$ provided by Mian and Laniel (1993), Table 2. Since k_t is regarded as non-stochastic the autocorrelations of u_t are the same as those of $k_t u_t$.

Table 2

Estimates of the hyperparameters.

	Additive	Multiplicative
σ_ξ^2	2.5267×10^8	3.293×10^{-4}
σ_ω^2	1.8382×10^{10}	1.10×10^{-8}
σ_ϵ^2	5.0083×10^9	1.2195×10^{-4}

Table 3

Estimates of the bias parameters and their coefficients of variation expressed as percentages.

Additive	Multiplicative
$\hat{b} = -1215099$	$\hat{B} = 0.9140659$
$cv\left(\hat{b}\right) = 4.0833$	$cv\left(\hat{B}\right) = 0.02386$

For the trading day and leap year coefficients we used model (2.6); we took $\sigma_\zeta^2 = 0$ since work by Dagum, Quenneville & Sutradhar (1992) indicated that a fixed trading day pattern is suitable for this type of data. The method of maximum likelihood of Engle & Watson (1981) was used for the estimation of the hyperparameters $\sigma_\epsilon^2, \sigma_\xi^2$ and σ_ω^2. The estimates obtained are given in Table

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

2. Estimates of the bias parameters under both the additive and multiplicative models are given in Table 3. When tests of significance of the bias were carried out we found that for the additive case $|t| = 24.5$ while for the multiplicative case $|t| = 4191$. The fact that the value of $|t|$ in the multiplicative case is much larger than in the additive case indicates the superiority of the multiplicative model for these data.

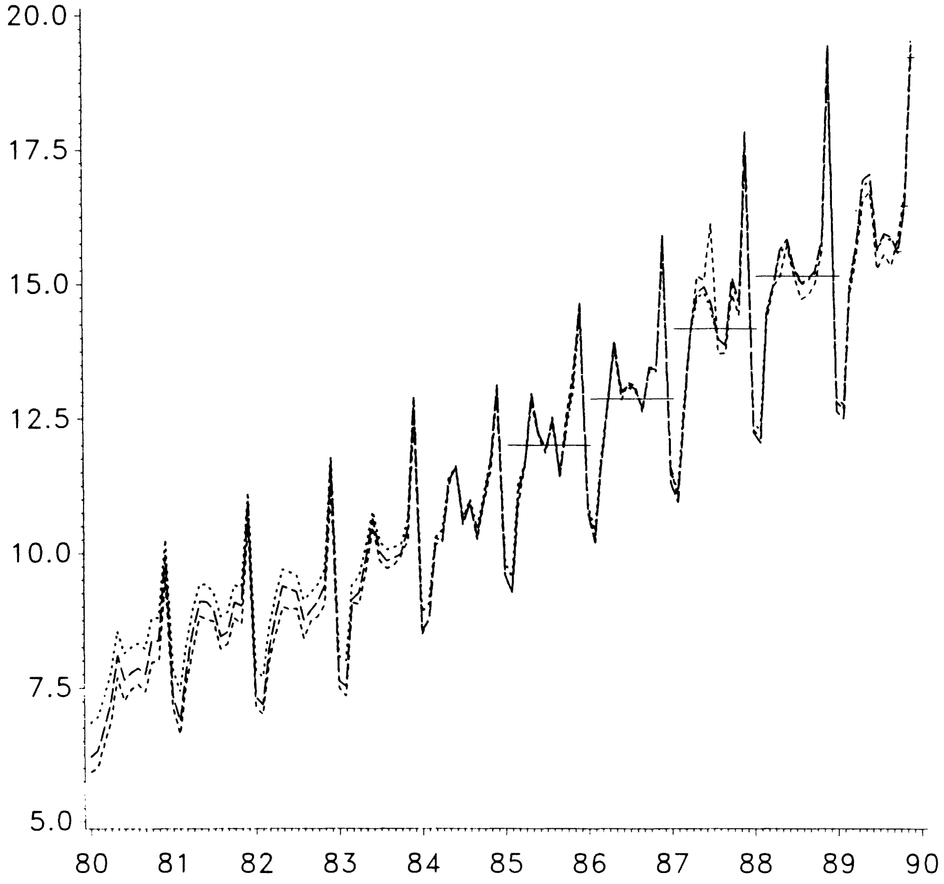

Figure 2. *Benchmarked values for the Canadian retail sales series and average values of the benchmarks: Additive Without Bias (AWB) is the dashed line; Multiplicative with Bias (MB) is the broken line; Additive with Bias (AB) is the dotted line. Values are expressed in millions.*

Figure 2 shows the results of benchmarking using the additive model without bias (AWB). All other methods without bias are indistinguishable on the graph from this case. The results of benchmarking for the additive model with bias (AB) and the multiplicative model with bias (MB) are also shown. These exhibit differences between each other and from the AWB case, particularly at the beginning of the series. Differences here are to be expected since there are no benchmarks at the beginning of the series. To explain the differences we note that as we move away from the benchmarks, the benchmarking correction converges to the bias parameter and this is zero for AWB, a constant for AB and a multiple of the level for MB.

There are further differences in years 87, 88 and 89 between the methods with and without bias.

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

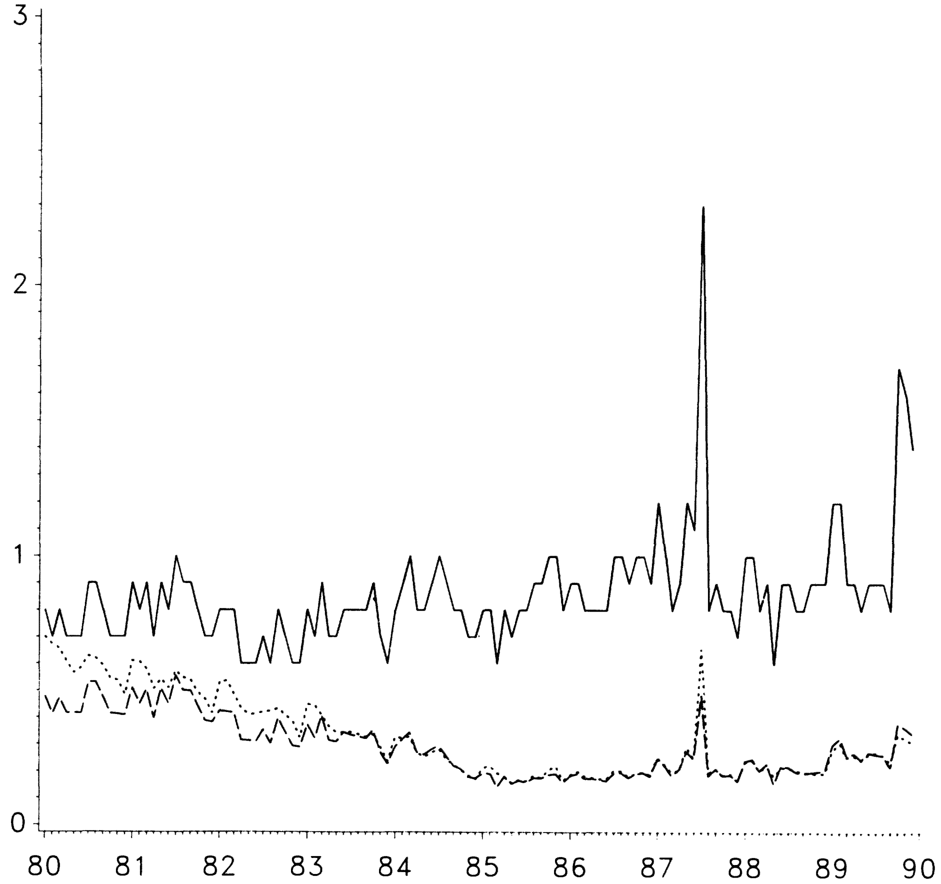

Figure 3. *Coeffiecients of variation (cv) expressed as percentages: original data is the solid line; AB is the dotted line; MB is the broken line*

These are explained by the exceptionally large cv for the monthly observation in July 87, as shown in Figure 3. This is a correct value and is due to an intentional adjustment to the sample that was made by survey statisticians for technical reasons. The effect has been to distort the benchmark adjustment in 87 and, because of autocorrelation in the survey errors, in subsequent years also. In a real-life analysis the value concerned would have been treated as an outlier in the cv series and would have been adjusted accordingly. We left the value unchanged in order to illustrate what happens in such cases.

As convergence criterion, we stopped iterating when the maximum relative absolute change was less than 10^{-6}. It took only five iterations to achieve convergence for both the two-stage and single-stage multiplicative methods.

The graphs show that even with a biased series, benchmarking with a model without bias can work well in years containing benchmarks. However, the models with bias handle the outlier in the cv series better in the sense that the monthly pattern in the benchmarked series when bias is included is closer to the pattern in the original series than with AWB.

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

Figure 3 shows the cv's of the original series (also provided in Table A3) and of AB and MB; all other methods give graphs which are indistinguishable from that of MB. It is clear that benchmarking reduces the cv by an amount which has considerable intrinsic worth quite apart from the general desirability of adjusting the monthly observations to satisfy the benchmark constraints. The fact that in the first part of the series the cv of AB is higher than that of MB is presumably due to the higher cv of the estimate of the bias parameter for the additive model compared with that for the multiplicative model.

For this particular series our conclusion is that the most appropriate model is the multiplicative model with bias correction.

9 Conclusions

The use of state space models for time series analysis is increasing for a variety of reasons. The main explanation is that the great flexibility of these models enables the analyst to model components or attributes of the series in ways that are specifically appropriate to the particular data set. For example, change in behaviour over time is relatively easily accommodated. Also, explanatory variables can readily be incorporated to allow for the influence of other factors on the series. Moreover the models are very general, so many other types of models, for example ARIMA models, emerge as special cases. Benchmarking is an important problem in applied time series analysis, particularly for government statistical agencies. We therefore claim that to tackle the benchmarking problem by state space methods, even for a single univariate series, is an important step forward.

In this paper we have provided two solutions to the problem. The first of these was a two-stage method using the approach that Hillmer & Trabelsi employed for ARIMA models. In the first stage we fitted a state space model to the monthly series and in the second stage we combined the results of this with the benchmarks. In the second method we put the monthly observations and the benchmarks together to form a single series and then analysed this by an appropriate state space model. Using both formulations we were able to deal with the situation, common with economic series, where the monthly data behave multiplicatively but the benchmarks are additive in nature. We also showed how to estimate a constant bias in the monthly series. Taken together, we regard these results as providing an effectively complete solution to the benchmarking problem for a single series.

To illustrate the methods we applied them to Canadian retail sales data; this is a series that is a little unusual in having two types of benchmarks, both annual totals and accurate monthly values. We found that the monthly data behaved multiplicatively and were biased. However, we had no difficulty in fitting any of our models to the data. The cv series for these data contained an outlier which we left untouched so that we could see what effect it had on the analysis. We found that when the bias parameter was excluded from the model there was some distortion of the seasonal pattern in the neighbourhood of the outlier but when the bias parameter was included the distortion disappeared from the benchmarked series. Inclusion of bias estimation is additionally useful since we can use the estimate of bias for provisional benchmarking of current observations pending arrival of the next benchmark. Our overall conclusion is that state space models provide a practical and useful basis for benchmarking time series.

Potential areas for further work include the following:

(i) Extension to multivariate series. This would, for example, enable regional series to be adjusted so that regional and national series satisfy appropriate regional and national benchmarks. Similarly, if an overall series is subdivided among categories it would enable overall and subseries to be benchmarked across categories.

(ii) Development of models for slow change of bias over time. If bias is in fact changing over time, this would enable it to be estimated more accurately; it would also assist the current adjustment of monthly values, particularly when benchmarks arrive substantially later than the

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

monthly values to which they relate.

(iii) Development of better time series models for survey errors with benchmarking in mind, possibly including allowance for nonsampling errors.

References

1. Cholette, P.A. & Dagum, E.B. (1994). Benchmarking Time Series with Autocorrelated Survey Errors. *International Statistical Review*, **62**, 365–377.
2. Dagum, E.B., Quenneville, B. & Sutradhar. B. (1992). Trading-day Variations Multiple Regression Models with Random Parameters. *International Statistical Review*, **60**, 57–73.
3. Durbin, J. & Cordero, M. (1993). Handling structural shifts, outliers and heavy-tailed distributions in state space time series models. (Unpublished).
4. Durbin, J. & Koopman, S.J. (1992). Filtering, smoothing and estimation for time series when the observations come from exponential family distributions. (Unpublished).
5. de Jong, P. (1989). Smoothing and Interpolation With the State-Space Model. *Journal of the American statistical Association*, **84**, 1085–1088.
6. de Jong, P. (1991). The diffuse Kalman filter. *The Annals of Statistics*, **19**, 1073–1083.
7. Engle, R. & Watson, M. (1981). A One-Factor Multivariate Time Series Model of Metropolitan Wage Rates. *Journal of the American Statistical Association*, **76**, 774–781.
8. Farmeir, L. (1992). Posterior mode estimation by extended Kalman filtering for multivariate dynamic generalised linear models. *Journal of the American Statistical Association*, **87**, 501–509.
9. Harvey, A.C. (1989). *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press, Cambridge.
10. Hillmer, C.H. & Trabelsi, A. (1987). Benchmarking of Economic Time Series. *Journal of the American Statistical Association*, **82**, 1064–1071.
11. Mian, I.U.H. & Laniel, N. (1993). Maximum Likelihood Estimation of Constant Multiplicative Bias Benchmarking Model with Application. *Survey Methodology*, **19**, 165–172.

Résumé

Nous avons une série d'observations mensuelles provenant d'une enquête par échantillonnage et nous avons aussi l'information sur les propriétés de l'erreur d'échantillonnage. De plus nous avons quelques valeurs annuelles qui sont très précises, par exemple, le vrai total annuel des valeurs mensuelles. Cet article discute du problème de l'adjustement des valeurs mensuelles afin de les rendre compatible avec les données annuelles.

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

APPENDIX

Table A1

Canadian retail sales series from January 1980 to December 1989.

5651446	5761119	6127963	6584991	7362407	6951997
7069743	7148155	7047731	7614265	7625369	9081377
6684048	6258144	7098029	7795870	8256052	8257363
8109850	7663343	7748144	8249476	8199209	9973251
6665545	6526251	7449366	8064762	8528248	8479855
8427737	7955934	8107457	8255811	8537885	10639657
6921890	6831866	8268510	8405857	8977537	9643260
9077139	8928076	8973979	9030038	9476804	11707996
7751579	7929391	9165024	9288823	10268429	10485529
9508814	9885265	9336847	9960715	10598300	11901198
8689668	8390380	10107485	10541145	11763659	11067487
10810755	11289656	10336540	11213751	11935495	13300288
9753373	9249279	10609952	11637936	12695108	11826254
11940908	11866547	11540397	12208845	12201498	14479170
10271723	9951105	11492162	12867443	13508434	13608274
13278474	12728196	12616239	13760829	13380142	16269757
11134013	10959374	13177788	13666311	14267530	14432944
13960825	13691315	13773109	13900743	14453461	17772990
11537416	11402084	13652705	14392411	15487026	15595082
14305348	14584459	14521628	14297343	15182860	17909814

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms

Table A2

Benchmarks, reference periods and coefficients of variation (expressed as percentages) for the Canadian retail sales series.

Reference period	Benchmark	cv
Feb 1985 to Jan 1986	143965400	0.03276
Feb 1986 to Jan 1987	154377100	0.03055
Feb 1987 to Jan 1988	169944600	0.19263
Feb 1988 to Jan 1989	181594000	0.13728
Oct 1989	15584920	0.5
Nov 1989	16430621	0.6
Dec 1989	19182630	0.5

Table A3

Coefficients of variation, expressed as percentages, for the Canadian retail sales series.

0.8	0.7	0.8	0.7	0.7	0.7	0.9	0.9	0.8	0.7	0.7	0.7
0.9	0.8	0.9	0.7	0.9	0.8	1.0	0.9	0.9	0.8	0.7	0.7
0.8	0.8	0.8	0.6	0.6	0.6	0.7	0.6	0.8	0.7	0.6	0.6
0.8	0.7	0.9	0.7	0.7	0.8	0.8	0.8	0.8	0.9	0.7	0.6
0.8	0.9	1.0	0.8	0.8	0.9	1.0	0.9	0.8	0.8	0.7	0.7
0.8	0.8	0.6	0.8	0.7	0.8	0.8	0.9	0.9	1.0	1.0	0.8
0.9	0.9	0.8	0.8	0.8	0.8	1.0	1.0	0.9	1.0	1.0	0.9
1.2	1.0	0.8	0.9	1.2	1.1	2.3	0.8	0.9	0.8	0.8	0.7
1.0	1.0	0.8	0.9	0.6	0.9	0.9	0.8	0.8	0.9	0.9	0.9
1.2	1.2	0.9	0.9	0.8	0.9	0.9	0.9	0.8	1.7	1.6	1.4

[*Received August 1995, accepted September 1996*]

This content downloaded from 137.122.8.73 on Mon, 27 May 2019 01:45:31 UTC
All use subject to https://about.jstor.org/terms