Survey Methodology, December 1990
Vol. 16, No. 2, pp. 271–277
Statistics Canada

Benchmarking of Economic Time Series

NORMAND LANIEL and KIMBERLEY FYFE[1]

ABSTRACT

Benchmarking is a method of improving estimates from a sub-annual survey with the help of corresponding estimates from an annual survey. For example, estimates of monthly retail sales might be improved using estimates from the annual survey. This article deals, first with the problem posed by the benchmarking of time series produced by economic surveys, and then reviews the most relevant methods for solving this problem. Next, two new statistical methods are proposed, based on a non-linear model for sub-annual data. The benchmarked estimates are then obtained by applying weighted least squares.

KEY WORDS: Survey errors; Non-linear model; Weighted least squares.

1. INTRODUCTION

Traditionally benchmarking has been defined as the method of adjusting monthly or quarterly figures derived from one source to annual values (benchmarks) obtained via another source (see Denton 1971, Cholette 1988a, and Monsour and Trager 1979). For example, the monthly shipments of Canadian Manufacturers could be adjusted so that they add up to the Annual Census of Manufacturers shipments figures. Another definition of benchmarking is the more general one of improving sub-annual estimates derived from one source with annual estimates obtained via a second source (see Hillmer and Trabelsi 1987). This definition assumes that the annual values are subject to error, which is not the case with the first definition. For example, the monthly inventories of Canadian Retailers derived from a sample survey could be improved using the end of year inventories obtained from the annual retail trade sample survey. This second definition of benchmarking corresponds to the situation encountered with many economic time series and is the one dealt with in this paper.

The purpose of this article is twofold, first it describes in detail, the benchmarking problem as it appears for many time series produced by large scale economic surveys. Then, two well known benchmarking methods dealing with a single time series are presented and discussed. Since both of these methods fail in some respects to resolve the problem, two other methods which use a non-linear weighted least squares approach are proposed. Finally, two of the above mentioned methods are illustrated with some simulated data and the results are discussed.

2. PROBLEM DESCRIPTION

The problem of improving a two-way table of sub-annual series of estimates with annual series of estimates from business surveys is described here, accompanied by the characteristics of the original data and a list of the features desired from a benchmarking procedure.

[1] Normand Laniel and Kimberley Fyfe, Business Survey Methods Division, Statistics Canada, Ottawa, Ontario, K1A 0T6.

The sub-annual estimates are often biased due to frame coverage deficiencies. Undercoverage is caused by delay in the inclusion of new businesses and no representation of non-employer businesses (usually small ones) on the frame. These sub-annual estimates are usually derived from relatively small overlapping samples, implying that sampling variances are relatively large and that sampling covariances exist between sub-annual estimates of different time periods. In addition, most economic sub-annual surveys produce series of estimates for a number of industrial activities within a number of geograghical regions. These are published sub-annually in the form of industry by geographical region tables, where the cells as well as the marginals and the grand totals need to be benchmarked.

As regards annual estimates, they can be assumed to be unbiased since in practice their frames do not suffer much from coverage deficiencies. Also, the annual estimates usually come from relatively large samples or censuses and thus have relatively small or no sampling errors associated with them, while their sampling covariances tend to be large because of substantial sample overlap between years. Another point to note about the annual estimates is that these figures come in approximately two years after the time to which they refer. For example, annual data for 1988 will not be released until some time in 1990, while sub-annual data are usually available a few months after the time period to which they refer. Therefore, when the sub-annual estimates are to be benchmarked, there will be no annual benchmarks for some of the sub-annual periods.

There are a number of features that a benchmarking procedure should have in order to be used for large scale survey estimates. First, the procedure should be simple enough that it can be used without too much data analysis. Second, it must be possible to produce preliminary benchmarking factors for periods for which benchmarks are not yet available. This feature allows benchmarking to be performed as the sub-annual data are produced. Otherwise discontinuities will be introduced in the sub-annual data. It is also desirable that the method maintain consistency between the grand-totals, marginal totals, and cell estimates for the benchmarked estimates in a table.

More discussion on the last two features can be found in Laniel and Fyfe (1989) and (1990) and Cholette (1988a) and (1988b). The rest of this paper deals with the problem of benchmarking a single time series in the context described above.

3. BENCHMARKING A SINGLE SERIES

Four approaches to benchmarking a single time series of sub-annual flow or stock estimates are described in the following sub-sections.

3.1 Denton's Method

In his 1971 paper, Denton proposed procedures for benchmarking based on a Quadratic Minimization approach, each of which corresponds to a specific penalty function. One of these penalty functions is the proportionate first difference between the original and benchmarked series and is often used for the problem of benchmarking time series that was described in section 2. Denton's procedure can be presented in statistical terms by first stating that the sub-annual estimates follow the model:

$$\frac{\theta_t}{y_t} = \frac{\theta_{t-1}}{y_{t-1}} + \epsilon_t, \qquad t=1, 2, \ldots, n \tag{3.1}$$

subject to the restriction to the annual data:

$$z_T = \sum_{t \in T} \theta_t, \qquad T = 1, 2, \ldots, m, \tag{3.2}$$

where:

- t refers to a sub-annual period,
- T refers to an annual period,
- $\{y_t\}$ is a sequence of biased estimates of the sub-annual parameters (levels),
- $\{\theta_t\}$ is a sequence of fixed sub-annual parameters (true values of the levels),
- $\{\epsilon_t\}$ is a sequence of uncorrelated and identically distributed errors with mean vector and covariance matrix $(\mathbf{0}, \sigma^2 \boldsymbol{I})$ and,
- $\{z_T\}$ is a sequence of annual benchmarks.

To find the benchmarked estimates, least squares are applied to the above restricted model.

It is important to note that Denton's approach assumes that the bias follows a random walk and that both the sub-annual and annual data are observed without sampling errors. Unfortunately, these assumptions are unlikely to be satisfied by economic time series (see section 2).

3.2 Hillmer and Trabelsi's Method

In 1987, Hillmer and Trabelsi proposed an alternative approach to benchmarking based on the Box-Jenkins (1976) ARIMA models. They assumed that the sub-annual estimates follow the model:

$$y_t = \theta_t + u_t \qquad t = 1, 2, \ldots, n \tag{3.3}$$

and the annual estimates follow the model:

$$z_T = \sum_{t \in T} \theta_t + a_T \qquad T = 1, 2, \ldots, m, \tag{3.4}$$

where:

- $\{\theta_t\}$ is a sequence of stochastic sub-annual parameters (true values of levels) following an ARIMA model,
- $\{y_t\}$ is a sequence of unbiased estimates of the sub-annual parameters,
- $\{u_t\}$ is a sequence of sub-annual dependent sampling errors with mean vector and covariance matrix $(\mathbf{0}, \boldsymbol{\Sigma}_u)$,
- $\{z_T\}$ is a sequence of annual unbiased estimates, and
- $\{a_T\}$ is a sequence of annual dependent sampling errors with mean vector and covariance matrix $(\mathbf{0}, \boldsymbol{\Sigma}_a)$.

Using the above models, they obtain the benchmarked sub-annual estimates by applying stochastic least squares. That is, they minimize $E(\hat{\theta}_t - \theta_t)^2$, the mean squared error. This technique is also referred to in time series terminology as signal extraction, and the derivation of the solution can be found in the paper written by Hillmer and Trabelsi.

As it is stated with the models, this method takes into account the sampling variances and covariances of the sub-annual and annual estimates. Unfortunately, the approach does not accommodate biases in the sub-annual data. Also, since ARIMA modelling is being used in this method, it would be costly to implement for large scale surveys dealing with hundreds of series. Therefore it would be best to use this type of approach for only a small number of very important economic indicators. There would also be risks of oversmoothing the data if the ARIMA models are not properly specified.

Cholette and Dagum(1989) modified the Hillmer and Trabelsi approach by introducing an "intervention" model instead of an ARIMA model. This allows the modelling of systematic effects in the time series, but according to the authors, this approach still possesses the same weaknesses as the original Hillmer and Trabelsi method.

3.3 Model on Trends

The following method was developed in an attempt to meet the benchmarking requirements of the economic surveys. It is based on the assumption that the sub-annual estimates follow the model:

$$\frac{y_t}{y_{t-1}} = \frac{\theta_t}{\theta_{t-1}} + v_t \qquad t = 1, 2, \ldots, n \tag{3.5}$$

and the annual estimates follow model (3.4), where:

$\{\theta_t\}$ is a sequence of sub-annual parameters (true values), as in Denton's method,

$\{v_t\}$ is a sequence of dependent sub-annual sampling errors of the trends with mean vector and covariance matrix $(\mathbf{0}, \Sigma_v)$.

Least squares theory is applied to the above models to produce benchmarked estimates. The description of the Gauss-Newton algorithm necessary to solve this problem and the calculation of the covariance matrix of the benchmarked estimates are given in Laniel and Fyfe (1989) or (1990).

This method can be used when the benchmarks come from either a census or annual overlapping samples and when the sub-annual level estimates are biased, if the relative bias is a constant. The assumption of a constant relative bias will be verified in practice if the rate of the frame maintenance activities is relatively stable, that is, when the proportion of frame coverage deficiencies is fairly constant over time. This assumption also implies that the undercovered businesses behave like the ones covered by the frame.

One technical problem with this method is that the sampling variance-covariance matrix of the sub-annual trends cannot be calculated directly and an approximation has to be used. The first-order Taylor approximation has been tried but in some cases the resulting sampling variances and covariances were zero or negative when they should be positive. For this reason, an alternative model to (3.5) is presented in the next section.

3.4 Model on Levels

The following method is an alternative to the previous one and is suggested so that the sampling variance-covariance matrix of the sub-annual estimates would be easier to obtain. It assumes that the sub-annual estimates follows the model:

$$y_t = \alpha\theta_t + u_t \qquad t = 1, 2, \ldots, n, \tag{3.6}$$

where α is a fixed parameter taking into account the constant relative bias and u_t is the same as for equation (3.3). The annual estimates follow model (3.4).

Benchmarked estimates are obtained by applying least squares theory to the above models. The algorithm required to solve this problem is the same as for method 3.3.

3.5 Discussion

Among the methods reviewed here, the most appropriate one for benchmarking a single time series in the context of the large scale surveys is the new approach based on the model on levels. It has a statistical basis which allows us to calculate confidence regions and test the goodness of fit of the benchmarked model. To test for lack of fit one has to be careful in choosing a test since the benchmarked estimates, $\hat{\theta}_t$, have quite a small number of degrees of freedom, $m - 1$ (the number of annual observations minus one), in comparison to the number of observations, $n + m$. This small number of degrees of freedom also suggests that with the model on levels, we can expect to get benchmarked estimates with a chronological pattern similar to the one observed in the sub-annual data.

A current practical issue with benchmarking methods which take into account sampling errors such as in 3.4, is the derivation of sampling covariances between two level estimates corresponding to two different time periods. Should they be calculated directly using the sample design for all pairs of time periods or should they be modelled? From a theoretical point of view, it is better to calculate these directly, since the sequence of sampling errors is intrinsically a non-stationary stochastic process due to the population variance-covariance varying with time. However, calculating all sampling covariances can be cumbersome, thus leaving the issue of how to obtain sampling covariances still an open question.

3.6 An Example

As a comparison between Denton's method described in section 3.1 and the model on the levels approach suggested in section 3.4, these two methods were applied to a special and interesting benchmarking case. It is a situation where the annual estimates have sampling variances six times the size of the sampling variances of the corresponding monthly estimates. In such a case, the advantage of using the model on levels approach instead of Denton's method will be easily observed.

The special case, though possible in practice, was made up of simulated data. Firstly, twenty-four monthly estimates were taken from an existing economic survey. A sampling covariance matrix was arbitrarily given to these monthly estimates. The variances and covariances were calculated in by using an equal coefficient of variation through time and the following correlation pattern:

$$\rho_{ij} = 1 - \frac{|j - i|}{24} \quad \text{for} \quad i = 1, 2, \ldots, 24 \quad \text{and} \quad j = 1, 2, \ldots, 24$$

where i and j are the indexes of a pair of monthly estimates. Then, two corresponding annual estimates were constructed as follows. The first annual figure was 25% larger than the sum of the first monthly figures. Whereas the second annual figure was only 5% larger than the total of the last twelve monthly observations. The two annual estimates were given sampling variances equal to six times the variances of the corresponding monthly totals and their correlation was fixed at 0.5.

The monthly estimates are represented by the full continuous line and the annual estimates by the horizontal lines on figure 3.1. The two horizontal lines are equal to the values of the annual figures divided by twelve. On the same figure, the line with long dots represents the monthly series benchmarked with the approach based on the model on levels. The line with short dots is the benchmarked monthly series with Denton's method.

From figure 3.1, it can be observed that the series benchmarked with the model on levels approach has the same year-to-year movement as the original monthly series. Whereas the series benchmarked with Denton's method has the same year-to-year movement as the annual estimates. It can also be seen that both benchmarked series are over the original monthly series.

The difference in the year-to-year movement between the two benchmarked series can be explained as follows. The approach based on the model on levels gives the benchmarked series a year-to-year movement essentially obtained by weighting the annual and sub-annual data with the inverse of their sampling variances. Since, in this example, the sub-annual estimates are much more reliable than the annual estimates, the benchmarked series got the year-to-year movement of the monthly figures. Whereas with Denton's method, the year-to-year movement of the benchmarked series is constrained to one of the annual series regardless of its reliability. In this sense the approach based on the model on levels is better than Denton's method.

As a last comment on this example, the fact that both benchmarked series are above the original monthly series simply illustrates that both methods are providing a correction for the bias of the monthly estimates.

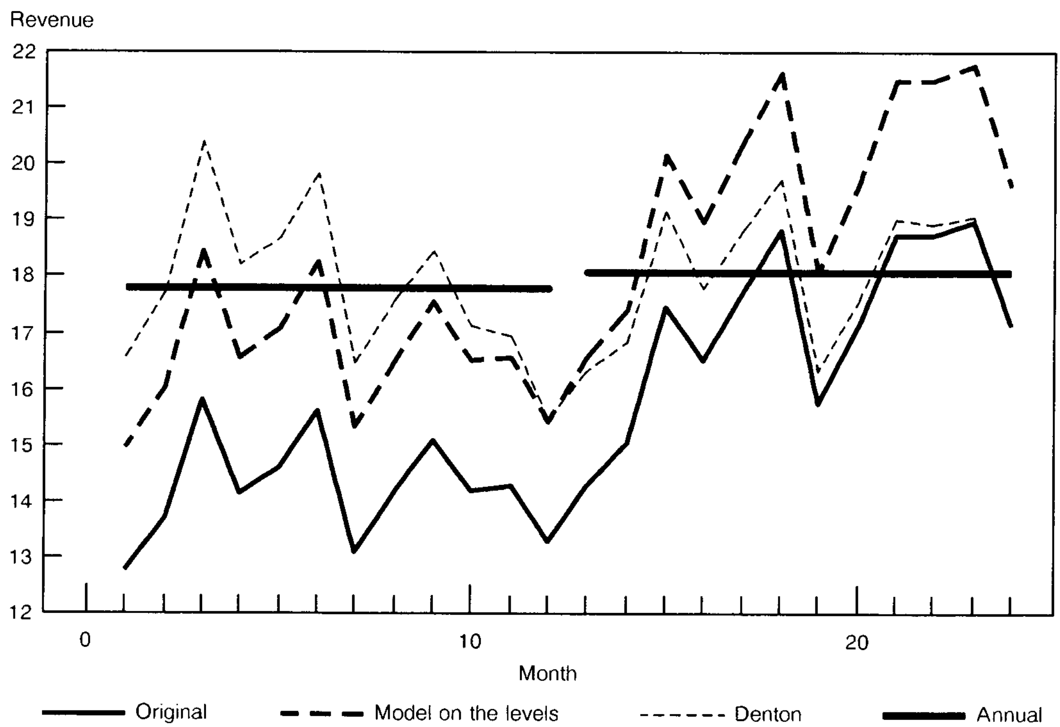

Figure 3.1 Plot of the original and two benchmarked monthly series and of the annual series

4. CONCLUSION

The problem of improving sub-annual survey estimates with the use of annual survey estimates has been examined. A new and simple procedure to benchmark a single time series has been presented. This procedure could be implemented in a computer system to allow its use in an automated mode. The advantage of the procedure over more traditional methods (*e.g.*, Denton's) is that it takes account of sampling errors. Some issues in using the proposed procedure for benchmarking a single time series have been discussed. Two important practical questions have been pointed out: benchmarking a table of series and preliminary benchmarking. Approaches to address these two topics have to be explored.

REFERENCES

BOX, G.E.P., and JENKINS, G.M. (1976). *Time Series Analysis, Forecasting and Control*. Holden-Day.

CHOLETTE, P.A. (1988a). Benchmarking and Interpolation of Time Series. Statistics Canada, Working Paper No. TSRA-87-014E.

CHOLETTE, P.A. (1988b). Benchmarking Systems of Socio-Economic Time Series. Statistics Canada, Working Paper No. TSRA-88-017E.

CHOLETTE, P.A., and DAGUM, E.B. (1989). Benchmarking Socio-Economic Time Series Data: A Unified Approach. Working Paper No. TSRA-89-006E, Statistics Canada.

DENTON, F.T. (1971). Adjustment on Monthly or Quarterly Series to Annual Totals: An approach Based on Quadratic Minimization. *Journal of the American Statistical Association*, 66, 99-102.

HILLMER, S.C., and TRABELSI, A. (1987). Benchmarking of Economic Time Series. *Journal of the American Statistical Association*, 82, 1604-1071.

LANIEL, N., and FYFE, K. (1989). Benchmarking of Economic Time Series. *Proceedings of the Business and Economic Statistics Section, American Statistical Association*.

LANIEL, N., and FYFE, K. (1990). Benchmarking of Economic Time Series. Business Survey Methods, Statistics Canada, Business Survey Redesign Project Working Paper.

MONSOUR, N.J., and TRAGER, M.L. (1979). Revision and Benchmarking of Business Time Series. *Proceedings of the Business and Economic Statistics Section, American Statistical Association*.