WORKING PAPER NO. TRSRA-87-010E
TIME SERIES RESEARCH & ANALYSIS DIVISION
METHODOLOGY BRANCH

CORRECTING THE REFERENCE PERIODS OF ANNUAL AND QUARTERLY DATA

by

Pierre A. Cholette

This is a preliminary version. Do not quote without author's permission.
Comments are welcome.

LA RECTIFICATION DES PÉRIODES COUVERTES PAR DES DONNÉES ANNUELLES ET TRIMESTRIELLES

- RÉSUMÉ -

Le document présente une méthode pour rectifier les périodes couvertes par des données annuelles et trimestrielles. Par exemple, des données annuelles disponibles peuvent se rapporter aux douze mois commençant en avril et se terminant en mars; les valeurs corrigées couvrent les périodes de janvier à décembre. Des variantes de la méthode conviennent aux séries de flux, de stock et d'indice. Les périodes couvertes par les données disponibles peuvent aussi varier dans le temps.

CORRECTING THE REFERENCE PERIODS OF ANNUAL AND QUATERLY DATA

by

Pierre A. CHOLETTE

Statistics Canada
Methodology Branch
Time Series Research and Analysis Division
Coats Building, 13th floor "J"
OTTAWA, Canada
K1A 0T6

(613) 991-1600

August 1987

- SUMMARY -

A method is presented to correct the reference periods of yearly and quarterly data. For instance, the available yearly data may cover the 12 months from April to March; the corrected values would pertain to January to December. Variants are proposed for flow, stock and index series. The reference periods of the available may also change through time.

INTRODUCTION

In many cases, statistical agencies receive annual data which pertain to the financial year (or "fiscal year") of the respondent. The reference period of that year may range from June 1 to May 31 for instance. However the statistician actually needs figures which pertain to the conventional year (or "calendar year") ranging from January 1 to December 31. The yearly data published statistical agencies are indeed supposed to reflect the conventional year.

The problem is also aggravated by the fact that the reference period of the financial years may change from occasion to occasion for a given respondent: In 1982 (say), one company's report covers from April 1981 to November 1982; in 1983, from December 1982 to December 1983; in 1985 (no report in 1984), from January 1984 to June 1985; and so on.

The problem would be simple if the statistician had the sub-annual breakdown of the financial year values available. Unfortunately, this is generally not the case. Here is an example of how this may happen. In order to keep the response burden to questionaires to a minimum, statistical agencies ask a greater number of their respondents to fill detailled questionnaires on a yearly basis; and a lesser number of respondents to fill summary questionnaires on a monthly or quarterly basis. One variable available yearly for a respondent may therefore not be available at all sub-annually or be available with a lower degree of reliability.

This paper provides a general method to convert financial year data into conventional year values. The problem is seen as a special application of interpolation between annual benchmarks, according to a variant of the modified Denton method (Denton, 1971; Cholette, 1984). The method consists of dis-aggregating the financial data into sub-annual values; and, of re-aggregating the sub-annual values into the desired conventional year values. The dis-aggregation and re-aggregation are implicit however and do not actually have to be carried out: The desired annual values are simply a weighted average of the available financial year data.

Variants of the method are derived for flow series, for stock series and for index series. An approach to correct the reference periods of

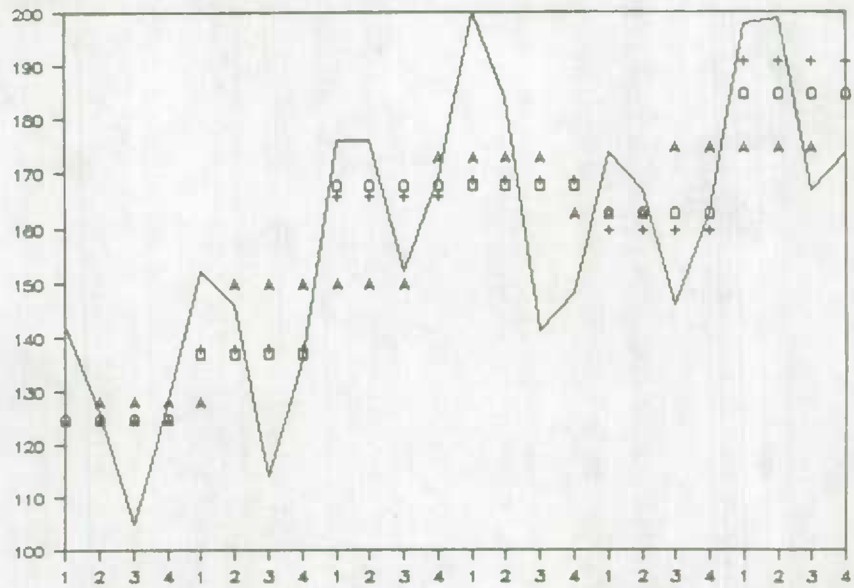
quarterly data is outlined. The conversion of (lumps of) weekly data into monthly values is covered in another paper (Cholette and Higginson, 1987), because too specific.

The paper unfolds according to Ehrenberg's (1982) recommendation about technical papers. Illustrations and results are presented first. Methodological details come after.

1. ILLUSTRATION OF METHOD FOR FLOW SERIES WITH IRREGULAR FINANCIAL YEARS

Flow series are such that they sum to the corresponding annual values, e.g. Retail Trade. The financial years of the flow series in Figure 1 have reference periods which vary from occasion to occasion. Their statistical usefulness is limited by the fact they are not comparable through time. The problem consists of converting those **five** available **irregular financial years** into **six** conventional values, each of which would cover 4 quarters.

The available financial year data are actually represented by their averages over their reference periods, i.e. by their value divided by the number of periods covered. This allows their display on the same scale as the underlying quarterly series. The latter series is not available in practise. For the sake of illustration however, it is assumed known. The underlying series contains an obvious seasonal pattern, with a seasonal peak in the first quarter of each year and a trough in the third. The trend-cycle component is not linear. It goes up from year 1 to year 3; down, for year 4 and 5; and then up again.



LEGEND: A: available financial year data; o: desired true conventional year values; +: estimated conventional year values; __: underlying quarterly series

Figure 1: Estimated conventional year values when the financial years cover sometimes more and sometimes less than four consecutive quarters (or 12 months).

The figure finally contains the true conventional year values (computed from the underlying quarterly series) and the corresponding estimated values. Like the financial years, the true and the estimated conventional years are represented by their averages over their reference period. The estimated conventional year values were obtained by means of the conversion method proposed in this paper. These estimates allow comparisons through time (contrary to the original irregular financial year data). Furthermore their percentage errors with respect to the true conventional year values are small: 0.1%, 1.0% -1.4% 0.8% -1.6% and 3.5% for years 1 to 6 respectively.

2. SIMPLE SOLUTION FOR FLOW SERIES WITH REGULAR FINANCIAL YEARS

Figure 2 describes a simpler and probably more common situation where all the available financial year values $y^a{}_i$ (for year i) refer to 4 consecutive quarters. All values cover from the second quarter of one year to the first of the following year. In cases of regular financial years, estimation is greatly simplified. Each estimated conventional year value is a convex combination of the two closest financial year data:

(2.1) $$y^d{}_i = (K\ y^a{}_{i-1} + (J-K)\ y^a{}_i) / J, \quad i=2,...,N-1,$$

or more specifically for the situation in Figure 2,

(2.1') $$y^d{}_i = (1\ y^a{}_{i-1} + 3)\ y^a{}_i) / 4, \quad i=2,...,5.$$

In equation (2.1'), J is the number of months per year (4 for quarterly and 12 for monthly series); and K is the number of "months" not in the conventional year reference period. A positive K means that K quarters wanted in year i are in fact in year i+1; and a negative K, in year i-1. (For Figure 2, K=1 and J=4.) Variable N stands for the total number of conventional years referred to -however partially- by the financial years. (For instance a financial year covering from avril 1981 to March 1982 refers to two conventional years, 1981 and 1982.)

```
financial years:             ya1      ya2      ya3
                           |---+---|---+---|---+---|
quarters:                1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4
years:                   |   1   |   2   |   3   |   4   |
                         |---+---|---+---|---+---|---+---|
conventional years:         yd1     yd2     yd3     yd4
```

Figure 3: The reference periods of conventional years 2 and 3 (central years) are embedded in the reference periods of financial years 1 and 2, and of 2 and 3, respectively. The reference periods of conventional years 1 and 4 are not entirely embedded. In that sense, the corresponding conventional year estimates are forecasts (as opposed to interpolations).

The first and last desired values are respectively obtained by:

(2.2) $$y^d_1 = ((J+K)\ y^a_1 + K)\ y^a_2)\ /\ J,$$

(2.3) $$y^d_N = ((J-K)\ y^a_{N-2} + (2J-K)\ y^a_{N-1})\ /\ J;$$

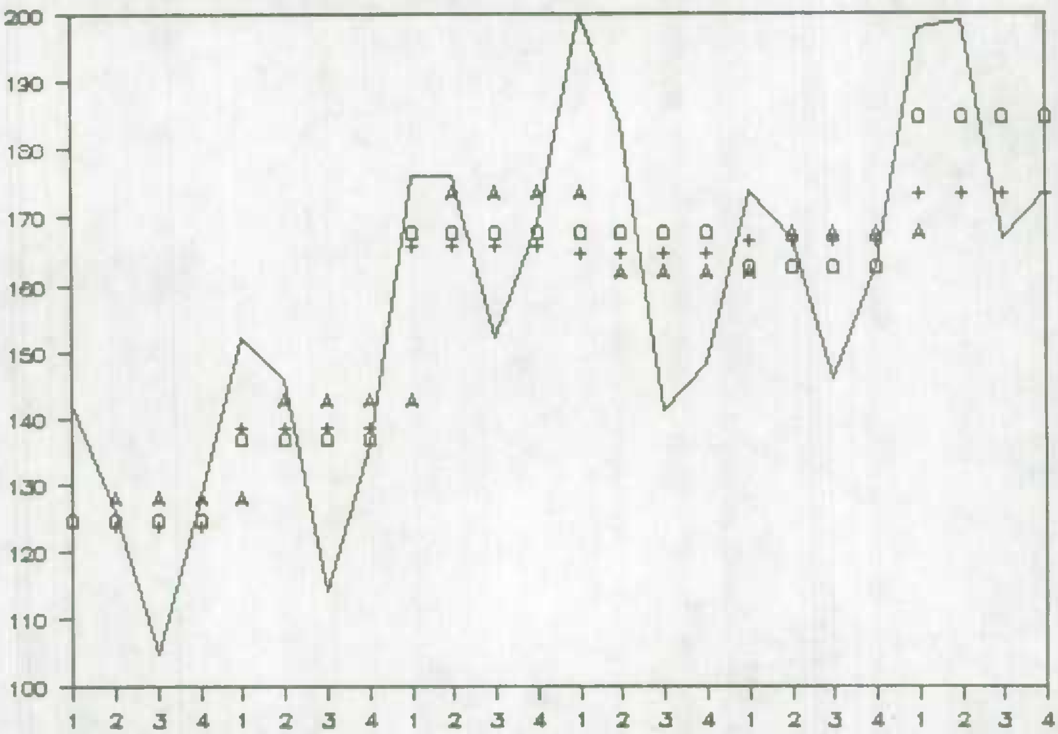
or more specifically in the situation of Figure 2,

(2.2') $$y^d_1 = (5\ y^a_1 - 1\ y^a_2)\ /\ 4,$$

(2.3') $$y^d_6 = (-3\ y^a_4 + 7\ y^a_5)\ /\ 4.$$

These **end estimates** of the conventional years are actually extrapolations or forecasts. This is illustrated in Figure 3. Their reliability is therefore lower than that of the **central estimates** of equation (2.1). Note that the end estimates do not have to be used in practise.

Figure 2 displays the conventional year values estimated by means of (2.1') to (2.3'). The percentage errors (with respect to the true desired conventional values) are -0.9%, 1.7%, -1.2%, -2.0%, 2.6% and 5.9% for years 1 to 6 respectively. The error is bigger for the last year (5.9%) because that estimate is in fact an extrapolation.



LEGEND: A: available financial year data; o: desired true conventional year values; +: estimated conventional year values; __: underlying quarterly series

Figure 2: Estimated conventional year values when each financial year value cover four consecutive quarters (or 12 months)

One attitude with respect to regular financial year values is to ignore the fact that their reference periods are wrong. In other words, the first available financial value is considered an estimate of the first conventional value; the second available value, of the second conventional value; etc. The resulting percentage errors are 2.1% for year 1, instead of -0.9% with the proposed method; 4.4% for year 2, instead of 1.7%; 3.5%, instead of -1.2%; -3.8%, instead of -2.0%; and 3.7% for year 5, instead of 2.6%. (Doing this provides no estimate for year 6.) The errors are larger than with the method proposed and biased. The bias is positive in periods of increase in the (trend-cycle of) the underlying series; and negative in periods of decrease.

Equations (2.1) to (2.3) correspond to the solution for flow series and regular financial years, when the method proposed is applied on 3-year series intervals (two years of financial data). Appendix A provides the weights corresponding to a 5-year specification. Section 4 explains how the weights are derived.

3. ASSUMPTIONS ABOUT THE UNDERLYING SUB-ANNUAL SERIES

In order to convert financial year data into conventional year values, hypotheses about the underlying sub-annual series have to be made.

The first assumption states that the series consists of the sum of 4 components: 1) the trend-cycle, 2) seasonality, 3) the irregular or random component and 4) the trading-day component:

(3.1) $$z_t = c_t + s_t + I_t + D_t$$ Assumption 1

By definition, the seasonal component tends to cancel out an any 4 or 12 successive periods:

(3.2) $$\sum_{j=1}^{J} s_{(i-1)J+j+K} \rightarrow 0$$ Assumption 2

where J is the number of months per year and i and j respectively stand for the year and the month considered.

The irregular or noise component has an expectation of zero. This implies it will tend to cancel out on K successive periods:

(3.3) $$E(I_t) = 0 \Rightarrow \sum_{k=1}^{K} I_{t+k} \rightarrow 0$$ Assumption 3

The trading-day component is also centered on zero. It is therefore reasonable to assume that it tend to cancel out on any K successive periods:

(3.4) $$\sum_{k=1}^{K} D_{t+k} \rightarrow 0$$ Assumption 4

This last assumption is relevant only for monthly flow series. In that case, it is always valid if K equals 3. Indeed, (3.4) then constitute a quarter. Since all quarters have (almost) the same number of days, they

cannot display any trading-day variations.

The two first assumptions are common in all time series decomposition methods. (See Dagum, 1980; Shiskin et al., 1967; Harvey and Todd, 1983) Assumption 3 and 4 are valid if K is sufficiently large or if the component considered is small.

One more assumption is needed. The trend-cycle is locally linear:

(3.5) $$c_t = 2c_{t-1} - c_{t-2} + a_t, \quad a_t \to 0$$ Assumption 5

This equation states that each value of the trend-cycle basically lies on the straight line running through its previous two values. This variant of linearity is that chosen by Boot, Feibes and Lisman (1967) for the interpolation of time series; by Leser (1961, 1963) and Schlicht (1981) for specifying the trend-cycle component in time series decomposition methods.

The available financial years $y^a{}_i$ and the desired conventional year $y^d{}_i$ may be respectively written as:

(3.6) $$y^a{}_i = z_{(i-1)J+1+Ki} + z_{(i-1)J+2+Ki} + \dots + z_{iJ+Ki}$$

(3.7) $$y^d{}_i = z_{(i-1)J+1} + z_{(i-1)J+2} + \dots + z_{iJ}$$

For $i=5$, $K_i=3$, $J=12$ (3.6) and (3.7) are respectively:

(3.6') $$y^a{}_5 = z_{52} + z_{53} + \dots + z_{60} + z_{61} + z_{62} + z_{63}$$

(3.7') $$y^d{}_5 = z_{49} + z_{50} + z_{51} + z_{52} + \dots + z_{59} + z_{60}$$

For K_i constant and positive ($K_i=K>0$), i.e. for regular financial years containing J "months", the difference between the desired and available yearly values is

(3.8) $$y^d{}_i - y^a{}_i = \Big[\sum_{k=1}^{K} (c_{(i-1)J+k} - c_{iJ-k+K+1})\Big] + \Big[\sum_{j=1}^{J} s_{(i-1)J+j} - \sum_{j=1}^{J} s_{(i-1)J+j+K}\Big]$$
$$+ \Big[\sum_{k=1}^{K} (I_{(i-1)J+k} - \sum_{k=1}^{K} I_{iJ-k+K+1})\Big] + \Big[\sum_{k=1}^{K} D_{(i-1)J+k} - \sum_{k=1}^{K} D_{iJ-k+K+1}\Big]$$

by virtue of Assumption (3.1). For $i=5$, $K=3$ and $J=12$ that difference reads:

(3.8') $$y^d{}_5 - y^d{}_5 = [c_{49} - c_{61} + c_{50} - c_{62} + c_{51} - c_{63}]$$
$$+ [(s_{49} + s_{50} + \dots + s_{60}) - (s_{52} + s_{53} + \dots + s_{63})]$$
$$+ [(I_{49} + I_{50} + I_{51}) - (I_{61} + I_{62} + I_{63})]$$
$$+ [(D_{49} + D_{50} + D_{51}) - (D_{61} + D_{62} + D_{63})]$$

Assumption 2 -if true- cancels the second term in brackets (in eq. (3.8) or (3.8')). Assumption 3 cancels the third term if I_t is small or K is large.

Assumption 4 cancels the fourth term, if D_t is small or K is large or equal to 3.

Under those assumptions, the difference between the conventional and the financial years is solely determined by the first term, that is by the year-to-year differences in the trend-cycle c_t. If c_t is linear, i.e. Assumption 5 is true, the financial years also behave linearly. Furthermore, the line in the latter -which is observable- is the same as that in the trend-cycle. The difference between the conventional and the financial can then be known exactly. More generally, the extent to which the trend-cycle (linear or not) of the underlying sub-annual series can be measured from the annual data, to the same extent the conversion will be exact. The conversion is garanteed to be exact under linearity. This is the principle of the conversion method proposed for regular financial years, and this is why it works even if the trend-cycle is not linear.

4 SPECIFICATION FOR FLOW SERIES WITH REGULAR FINANCIAL YEARS

This Section considers flow series, when all financial years refer to four consecutive quarters, that is when K is constant. This the more simple situation depicted in Figure 2. The conversion from financial to conventional years is then simplified, because neither the available financial year data nor the desired conventional year estimates contain seasonality, by virtue of Assumption 2. The problem reduces to dis-aggregating the financial year into non-seasonal values ($c_t + I_t + D_t$) and to re-aggregating the sub-annual values into conventional years. The underlying series may contain seasonality. Its estimates are not required however, because they would cancel out in the re-aggregation process.

For similar reasons Assumptions 3 and 4 further reduce the problem, namely to dis-aggregating the financial data into sub-annual trend-cycle values. The difference between the conventional and the financial year through (3.8), provide that the trend-cycle is linear. Linearity is maximizied by minimizing the following objective function:

$$f(c) = \sum_{t=3}^{IJ} (c_t - 2c_{t-1} + c_{t-2})^2 \qquad (4.1)$$

where I is the number of conventional years in the series interval considered (I<=N). The extent to which linearity is possible is governed by constraints:

$$\sum_{j=1}^{J} (c_{(i-1)J+j+K}) = y^a{}_i, \quad i=1,..,I-1. \qquad (4.2)$$

These state that the values sought have to sum to the available financial year values.

The solution to that constrained minimization problem is developped in Appendix B (for more general variants of the method). The desired trend-cycle values are weighted averages of the available financial

values:

$$c_t = \sum_{m=1}^{(I-1)} w^K_{t,m}\, y^a_i \quad t,\ldots,IJ \tag{4.3}$$

The desired conventional year values are then simply the annual sums of the sub-annual trend-cycles values:

$$y^d_i = \sum_{m=1}^{(I-1)} c_{(i-1)J+j}, \quad i=1,\ldots,I. \tag{4.4}$$

Since the trend-cycle estimates are of no interest in themselves, the desired conventional values can be expressed directly in terms of the available ones (by subtituting (4.3) in (4.4)):

$$y^d = \sum_{m=1}^{(I-1)} p^K_{i,m}\, y^a_m, \quad i=1,\ldots,I \tag{4.5}$$

The weights $p^K_{i,m}$ depend only on I, the number of years in the series interval; on J, the number of months per year; and on K, the number of months not in the conventional year. They do not depend on the observations y^a_i. They can therefore be applied to any series interval -or any series- with the same values of J and K. In other words, they can be calculated once and for all.

This is why Appendix A provides tables of weights for I=5, J=4 and 12 and for relevant values of K. We would recommend their implementation in a 5-year moiving average manner:

	$y^d_1 = \Sigma\, w^K_{1,m}\, y^a_m$	end year	(extrapolation)
	$y^d_2 = \Sigma\, w^K_{2,m}\, y^a_m$	non-central year	(interpolation)
(4.6)	$y^d_n = \Sigma\, w^K_{3,m}\, y^a_{n-3+m}$, n=3,...,N-2,	central years	(interpolation)
	$y^d_{N-1} = \Sigma\, w^K_{4,m}\, y^a_{N-5+m}$	non-central year	(interpolation)
	$y^d_N = \Sigma\, w^K_{5,m}\, y^a_{N-5+m}$	non-central year	(extrapolation)

where all summations take place over m going from 1 to 4. For I=3, the weights and their implementation are given by equations (2.1), (2.2) and (2.3).

5. SPECIFICATION FOR FLOW SERIES WITH IRREGULAR FINANCIAL YEARS

This Section considers flow series when the reference periods of the financial year data are irregular. This is the more complex case depicted in Figure 1. Financial year values covering sometimes more and sometimes less than 12 months (4 quarters) contain the seasonal values of the months in excess or missing in the reference period. Consider for instance a year refering to periods from March to May (of the following year). The 12

months which appear once cancel out in the financial year, by virtue of Assumption 2. However, the seasonal values of the extra March, April and May are part of the financial year considered. (Seasonality cancels out over 12 months but not necessarily over three.) In that sense, such financial year data do contain seasonality. And this fact must be reckoned with in dis-aggregating into sub-annual values. In other words, the dis-aggregated sub-annual values must reflect seasonality.

As for the irregular and the trading-day components, Assumptions 3 and 4 are still valid. (That statement will deserve qualification later on.) In other words, these components still tend to cancel out. The sub-annual values needed must therefore only comprise the trend-cyle (like before) and the seasonal components $\zeta_t = c_t + s_t$. An appropriate objective function is:

(5.1) $$f(\zeta) = \sum_{t=3}^{IJ} (\zeta_t/x_t - 2\zeta_{t-1}/x_{t-1} + \zeta_{t-2}/x_{t-2})^2$$

where x_t is an approximative seasonal pattern supplied by the subject matter expert of the series. For simplicity the values of x_t may be the same for each same month over the whole series interval considered. A value equal to 1.5 means that the month is 50% higher than an average month; equal to 0.6, 40% lower. A value of 1.0 specifies the month to be average; and a value of 0.0001, practically nill. Only the relative values of x_t matter.

The meaning of objective function (5.1) is then the following. The desired sub-annual trend-cycle-seasonal values ζ_t should be proportional to the seasonal pattern x_t; and the proportion ζ_t/x_t should change as linearly as possible. (Strictly speaking, Assumption 1 is being approximated by $z_t = (\zeta_t = c_t s_t) + I_t + D_t$.) The new objective function is minimized subject to the following constraints:

(5.2) $$\sum_{t=\tau_m}^{\rho_m} \zeta_t = y^a_m, \quad \rho_m >= \tau_m, \quad m=1,\dots,M.$$

These constraints state that the desired sub-annual values ζ_t must sum to the available financial data, whatever the reference periods $[\rho_m,\dots,\tau_m]$.

The solution is developped in Appendix B. As in the case of regular financial years, the desired sub-annual trend-cycle-seasonal values ζ_t are weighted averages of the available financial year values:

(5.3) $$\zeta_t = \sum_{m=1}^{M} w^x_{t,m}\, y^a_i \quad t,\dots,IJ$$

The desired conventional year values are then simply the annual sums of the sub-annual trend-cycles values:

(5.4) $$y^d_i = \sum_{m=1}^{(I-1)} \zeta_{(i-1)J+j}, \quad i=1,\dots,I.$$

Since the [illegible] in themselves, the desired con[illegible] directly in terms of the available [illegible]:

(5.5) [illegible]

The [illegible]. They do depend [illegible] on the seasonal pattern [illegible] to calculated for [illegible]

In many cases [illegible] realistic. Indeed a[illegible] trading-day component will not cancel especially if [illegible] is small and different from [illegible]. Objective function [illegible] defined as the product of [illegible] pattern [illegible]. The [illegible] all the daily rates of activity [illegible] seasonal [illegible]. Thus rates are expressed [illegible].

The [illegible] with the variant [illegible] pattern [illegible].

6. GENERALIZATION [illegible]

The two variants [illegible] until now only addressed flow series. This section [illegible] changes required to accommodate stock and index series. [illegible] is special case of (5.1), where x_t equals 1 and [illegible] developments will therefore start from the latter. The objective function [illegible]:

$$(6.1) \qquad [illegible] \sum_{t=1} [illegible] (\ldots_{t-2}/x_{t-2})^2$$

The unknown ζ_t represents [illegible] in case of regular financial years, i.e. for K constant, [illegible] the aggregate of the trend-cycle and the seasonal components for K not constant. Variable x_t is equal to 1.0 for K constant; or, to an expert supplied seasonal pattern for K not constant.

6.1 Stock series

The annual values of stock series (considered here) correspond to one of the sub-annual values. Unless the underlying sub-annual series is non-seasonal, it can no longer be assumed that the financial year and the conventional years do not contain seasonality. Indeed their values will always depend on the quarter or month chosen. A seasonal pattern x_t different from 1.0 is then supplied to objective function (6.1).

The appropriate [illegible] the particular [illegible] (6.1).

Since the sub-annual estimates are of no interest in themselves, the desired conventional values may be expressed directly in terms of the available ones (by subtituting (5.3) in (5.4)):

$$(5.5) \qquad y^d_i = \sum_{m=1}^{M} p^{K,x}_{i,m} \, y^a_m, \quad i=1,\ldots,I$$

The weights do not depend on the financial year values. They do depend however on the distribution of their references periods, on the seasonal pattern chosen and on J and I. This implies they have to calculated for each new series interval.

In many cases Assumption 4 may not be realistic. Indeed an important trading-day component will not cancel especially if K is small (and different from 3). Objective function (5.1) is still relevant if x_t is defined as the product of the seasonal and trading-day pattern ($x_t=s_tD_t$). The trading-day value D_t for a month is the sum of all the daily rates of activity d_k in the month. Similarly to the seasonal pattern, those rates are expressed with respect to the average of the week.

The estimated conventional year values of Figure 1 were estimated with the variant presented in this Section, with I=5, J=4 and a very well seasonal pattern x_t.

6. GENERALIZATION TO STOCK AND TO INDEX SERIES

The two variants of the method presented until now only addressed flow series. This section indicates the changes required to accomodate stock and incex series. Objective function (4.1) is special case of (5.1), where x_t equals 1 and ζ_t equals c_t. Further developments will therefore start from the latter. The objective function is then:

$$(6.1) \qquad f(\zeta) = \sum_{t=3}^{IJ} (\zeta_t/x_t - 2\zeta_{t-1}/x_{t-1} + \zeta_{t-2}/x_{t-2})^2$$

The unknown ζ_t represents the trend-cycle in case of regular financial years, i.e. for K constant; or, the aggregate of the trend-cycle and the seasonal components for K not constant. Variable x_t is equal to 1.0 for K constant; or, to an expert-supplied seasonal pattern for K not constant.

6.1 Stock series

The annual values of stock series (considered here) correspond to one of the sub-annual values. Unless the underlying sub-annual series is non-seasonal, it can no longer be assummed that the financial year and the conventional years do not contain seasonality. Indeed their values will always depend on the quarter or month chosen. A seasonal pattern x_t (different from 1.0) is then required in objective function (6.1).

The appropriate constraints for stocks series are particular cases of (5.2):

$$(6.2) \qquad \zeta_{tm} = y^a_m, \quad m=1,\ldots,M$$

According to Appendix B, the conventional year estimates are weighted averages of the financial values like in equation (5.5). These have to be recalculated for each series (interval).

6.2 Index series

The annual values of **index series** correspond to the average of the sub-annual values. (Flow and stock series expressed as indexes , e.g. the Index of Industrial production in the first case, the Consumer Price Index and Unemployment in the second case, are considered as index series for the purpose of the method.) The appropriate constraints are:

$$(6.3) \qquad \sum_{t=\tau_m}^{\rho_m} \zeta_t \,/\, (\rho_m - \tau_m + 1) = y^a_m, \quad \rho_m > \tau_m, \quad m=1,\ldots,M,$$

or more conveniently,

$$(6.3') \qquad \sum_{t=\tau_m}^{\rho_m} x_t = (\rho_m - \tau_m + 1)\, y^a_m, \quad \rho_m > \tau_m, \quad m=1,\ldots,M.$$

The latter equation is particular case of (5.2) if y^a_m is appropriately redefined.

When the reference period of the financial year are irregular, x_t is seasonal index and solution (5.5) applies. For index series, however, financial years should be regular. In that case, assumptions 2 is valid, so that x_t may be set to 1.0 and ζ_t to c_t (unless Assumption 4 cannot be justified). The solution is then the same as (4.5) except y^a_t is redefined. The weights have the same convenient properties described thereafter.

7. CORRECTING THE REFERENCE PERIOD OF QUARTERLY DATA

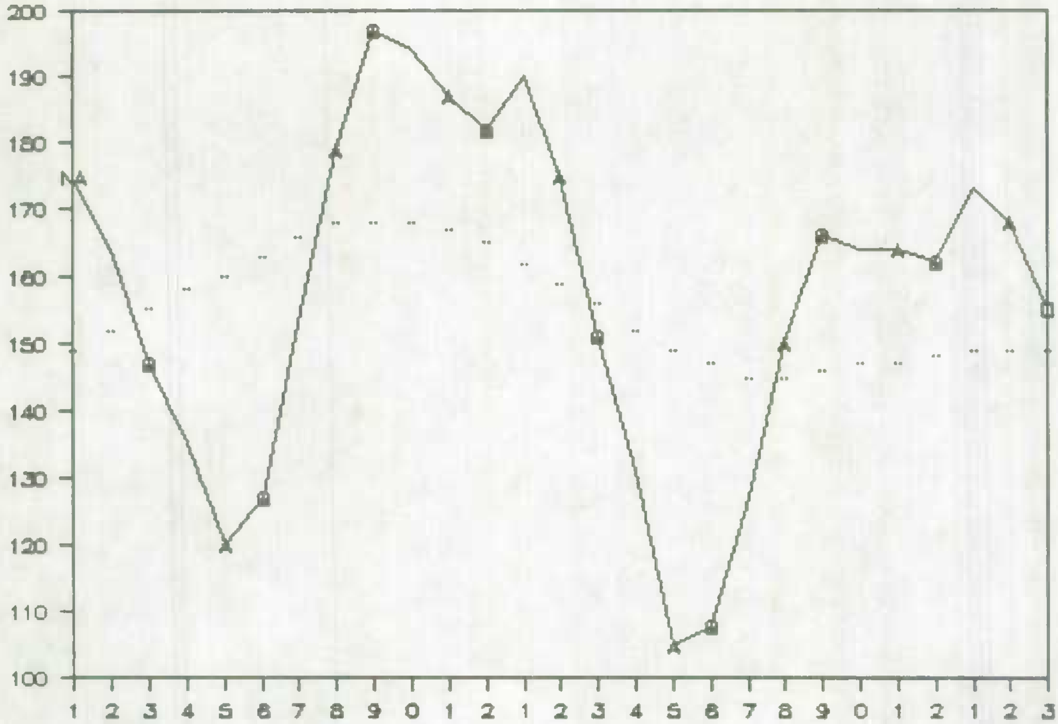
The problem discussed for yearly data is also encountered for quarterly data. For instance, the reference period of available quarterly data may cover December to February, March to May, etc., instead of January to March, April to June, etc. The reference periods may also be irregular.

A variant of the conversion method proposed can correct the reference periods of available quarterly data. These are dis-aggregated into monthly values which are then re-aggregated into the desired conventional quarters. An appropriate objective function is still a particular case of (5.1)

$$(7.1) \qquad f(\zeta) = [\sum_{t=3}^{3I} (\zeta_t/x_t - 2\zeta_{t-1}/x_{t-1} + \zeta_{t-2}/x_{t-2})^2]$$

where I now stand for the number of quarters (I>5) in the series interval considered. The objective function specifies that the desired monthly values should remain proportional to some expert-supplied monthly seasonal-trading-day pattern x_t. Objective function (7.1) is minimized subject to constraints (5.2), where y^a_i now represent the available

quarterly data. The solution is derived in Appendix B. Like in the other variants of conversion, the corrected quarterly estimates are a linear combination of the available quarters. Those weights depend on xt, and I.



LEGEND: A: available quarterly data; o: desired true quarterly values; +: estimated corrected quarterly values; __: underlying monthly series ; ..: underlying trend-cycle

Figure 4: Quarters corrrected by the method proposed for a stock series

Figure 4 illustrates an example of quarterly stock data corrected for reference periods. The available quarterly values refer to the February, May, August, November, February, etc., instead of March, June, September, December, March, etc. The estimation error of the corrected quarters (with respect to the true quarters) read -0.2%, -0.1%, 0.0%, 0.0%, 0.0%, 0.2% -0.1%, -0.1% and 0.4% for the nine quarters. When considering the available quarters as estimates of the desired ones, the errors are much bigger 12.0%, -6.1%, -9.1%, 2.9%, 16.0%, -2.6%, -9.4%, 0.8% and 8.3%.

8. DISCUSSION

The approach proposed for correcting the reference periods of yearly (and quarterly) data is based on assumptions about the underlying monthly series and in many cases on user-supplied seasonal or seasonal-trading-day patterns. Sometimes, an underlying montly series does exists, but is inconsistent with the yearly data. The coverage of respondents may be

different; the reliability, lower; etc. When such a monthly indicator is available, the conversion problem should be solved in the context of benchmarking, in our opinion. This recommendation is especially relevant when the corrected data are generated to be used as benchmarks for the corresponding sub-annual series.

Carrying out the conversion in the context of benchmarking would work as follows.

1) For one socio-economic variable, respondents with same financial years are aggregated. Obviously, the level of aggregation must be lower than that of publication, and the aggregation must involve respondents with presumably common seasonal and trading-day patterns (for the variable considered).

2) The same aggregation is carried out "annually" and sub-annually. Annual discrepancies are then observed between the annual and the sub-annual data.

3) The sub-annual data is benchmarked to the "annual" data. In the benchmarking process, the annual data are specified to pertain to the reference periods which they actually cover (and not to the conventional years). The result at this point is a sub-annual benchmarked series consistent the available yearly data.

4) The conventional year values are then calculated as the annual sums (or relevant operation) of the benchmarked series.

Denton (1971) provides an appropriate objective function for proportional benchmarking (subject to (5.2)):

$$\sum_{t=2}^{T} (z_t/x_t - z_{t-1}/x_{t-1})^2 \qquad (8.1)$$

Variable x_t is the available sub-annual information. The unkown z_t is the desired sub-annual benchmarked series consistent with the financial year data. This function specifies that the ratio of the benchmarked series to the original is as constant as possible. This specification imputes the sub-annual missing respondents (in $z_t - x_t$) proportionnaly the the ones available (in x_t). In other words, the missing respondents behave like the available ones. That behaviour is also governed by the observed annual discrepancies.

An objective function for additive abenchmarking (also from Denton, 1971) is:

$$\sum_{t=2}^{T} ((z_t - x_t) - (z_{t-1} - x_{t-1}))^2 \qquad (8.2)$$

This function assumes that the behaviour of the missing respondents (i.e. of $z_t - x_t$) is governed exclusively by the movement of the annual discrepancies. The behaviour would be constant if the annual discrepancies were constant; linear; sine-like; etc.

The advantages of correcting the reference periods in the context of benchmarking are the following.

1) The conversion is no longer based on assumptions but on facts about the sub-annual series. Those facts are explicitly and operationally incorporated in the conversion process. (Note that benchmarking does not assume that the movement of the original sub-annual series is reliable, but that it is better than nothing; see Cholette, 1987.)

2) When the variables involved in the conversion have to be benchmarked, both the conversion and benchmarking are achieved in a single operation.

When a respondent has a unique reporting pattern and/or provides no sub-annual values: The sub-annual values of other respondents may be used as sub-annual indicator.

The correction of the reference periods of quarterly data in the context of benchmarking is handled in a way similar to that of annual data.

The benchmarking approach to conversion proposed in this section is in a sense already used in some surveys conducted by Statistics Canada. The S.I.O. surveys for instance collect sub-annual and annual stocks series. The reference periods of the financial year values are corrected on the basis of the movement in the available sub-annual data. If for instance the variable considered rose by 5% sub-annually from December to March, then the financial year value pertaining to March is reduced by 5% to make it refer to December. That procedure implicitely assummes that objective function (8.1) has reached its minimum of zero; in other words, that a sub-annual series consistent with the financial year values would have the same growth rates as the sub-annual series actually available. (Two series exactly proportional to each other have identical growth rates.) The benchmarking strategy proposed here is a refinement of that procedure: It uses the exact growth rates required for consistency.

More information about benchmarking may be found in Denton (1971), Cholette, (1984, 1987) and Lanierl (1986).

CONCLUSION

It is well possible to convert financial year data, with reference period ranging from April to March for instance, into conventional year data, with reference period ranging from January to December. Under the method proposed the desired conventional year estimates are weighted moving averages of the available financial year data.

In all variants considered, assumptions about the underlying sub-annual series have to be made (see Section 3). Namely the trend-cycle component is linear and the seasonality is constant.

For flow and index series, these assumptions are sufficient to derive acceptable conventional year estimates, provided the reference periods of the financial years are regular and cover no more than 12 months (e.g. always June to May). The computations involved are also tremendously simplified: the weights of the moving averages are known in advance (see Appendix A) and do not require calculation for each new application.

For stock series or when the reference periods of the financial years are irregular (e.g. one year January to March, the next year April to March,

April to December, etc.), the assumptions about the underlying series are not sufficient to derive acceptable estimates. The user has to supply a seasonal pattern and possibly a trading-day pattern for the underlying sub-annual series. The computations involved are also much more time consuming: The weights of the moving averages are not known in advance and have to be recalculated for each application.

The same approach can also be used to correct the reference period of quarterly data. Both assumptions and user-supplied monthly seasonal and/or trading-day pattern are required.

In cases where some sub-annual data is available, we recommend solving the conversion problem within the framework of benchmarking. The conversion does not any more rely on assumptions about the underlying series and on user-supplied seasonal-trading patterns, but on facts. As explained in Section 8, the desired conventional year values are then the annual sums (or relevant operation) of the sub-annual benchmarked series.

ACKNOWLEDGEMENT

I am most thankful to my colleague Jocelyn Plourde for his valuable assistance.

REFERENCES

Boot,J.C.G., Feibes,W., Lisman, J.H.C. (1967), "Further Methods of Derivation of Quarterly Figures formj Annual Data", Applied Statistics, Vol. 16, No. 1, pp. 65-75.

Cholette, P.A. (1978), "A Comparison and Assessment of Various Adjustment Methods of Sub-annual Series to Yearly Benchmarks", Statistics Canada, Time Series Research and Analysis Division, Research Paper 78-03-001B.

Cholette, P.A. (1984), "Adjusting Sub-Annual Series to Yearly Benchmarks", Survey Methodology, Vol. 10, No. 1, pp., 35-49.

Cholette, P.A. (1987a), "Concepts, Definitions and Principles of Benchmarking and Interpolation of Time Series", Statistics Canada, Time Series Reasearch and Analysis, Research Paper forthcoming.

Cholette, P.A. (1987b) "Converting Weekly Data into Seasonal Monthly Values", Statistics Canada, Time Series Research and Analysis, Research Paper Forthcoming.

Dagum, E.B. (1980), "The X-11-ARIMA Seasonal Adjustment Method", Statistics Canada, Cat. no 12-564EF.

Denton, F.T. (1971), "Adjustment of Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization", J.A.S.A., Vol 66, No 333, pp. 99-102.

Ehrenberg, A.S.C. (1982), "Writing Technical Papers and Reports", The American Statistician, Vol. 36, No. 4, pp. 326-329.

Harvey, A.C., Todd, P.H.J. (1983), "Forecasting Economic Time Series with Structural and Box-Jenkins Models: A Case Study", J.B.E.S, pp. 299-365.

Laniel, N. (1986), "Adjustment of Economic Production Sub-Annual Series to Corresponding Annual Series", Statistics Canada, Business Survey Methods Division, Working Paper No. BSMD-86-006E

Leser, C.E.V. (1961), "A Simple Method of Trend Construction", J.R.S.S. B, Vol. 23, pp. 91-107.

Leser, C.E.V. (1963), "Estimation of Quasi-Linear Trend and Seasonal Variation", J.A.S.A., Vol. 58, pp. 1033-1043.

Schlicht, E. (1981), "A Seasonal Adjustment Principle and a Seasonal Adjustment Method Derived from this Principle", J.A.S.A., Vol. 76, No 374, pp. 374-378

Shiskin, J., Young, A.H., Musgrave, J.C. (1967), "The X-11 Variant of the Census Method II Seasonal Adjustment Program", U.S. Department of Commerce, Technical Paper No 15.

APPENDIX A: Five-year weights for converting financial year flow data when the financial years all cover 4 quarters or 12 monthls.

The row numbers in the tables indicate which conventional year in the 5-year interval is estimated by the weights in the row. The column number indicates to which available financial year the weight applies. For instance for J=4 and K=1

$$y^d{}_2 = 0.15152y^a{}_1 + 0.97760\ y^a{}_2 - 0.15976\ y^a{}_3 + 0.03064\ y^a{}_4$$

Parameter K stands for the number of months or quarters which should be in conventional year m, but are actually in year m+1 for positive K or in year m-1 for negative K. The weights are the same for K=-1 and K=J-1; for K=-2 and K=J-2; and so on. The weights are therefore given only for positive values of K.

Furthermore, the weights are the same for K=1 and K=J-1, for K=2 and K=J-2, except the rows and the columns are in the reverse order. This can be verified for the quarterly weights for K=0 and 4 and K=1 and 3. In order to save space, the weights will not be given in the monthly case for K greater than 6.

<u>Quarterly weights</u>:

for K=0:

	1	2	3	4
1	1.00000	0.00000	-0.00000	0.00000
2	0.00000	1.00000	0.00000	-0.00000
3	-0.00000	0.00000	1.00000	0.00000
4	0.00000	-0.00000	0.00000	1.00000
5	-0.14901	0.72594	-2.00483	2.42791

for K=1:

	1	2	3	4
1	1.34997	-0.48459	0.16926	-0.03464
2	0.15152	0.97760	-0.15976	0.03064
3	-0.04342	0.26595	0.84837	-0.07090
4	0.02719	-0.13433	0.43710	0.67005
5	-0.11064	0.53921	-1.49650	2.06793

for K=2:

	1	2	3	4
1	1.70795	-0.98817	0.35249	-0.07227
2	0.38013	0.77921	-0.19881	0.03947
3	-0.07663	0.57663	0.57663	-0.07663
4	0.03947	-0.19881	0.77921	0.38013
5	-0.07227	0.35249	-0.98817	1.70795

for K=3:

	1	2	3	4
1	2.06793	-1.49650	0.53921	-0.11064
2	0.67005	0.43710	-0.13433	0.02719
3	-0.07090	0.84837	0.26595	-0.04342
4	0.03064	-0.15976	0.97760	0.15152
5	-0.03464	0.16926	-0.48459	1.34997

for K=4:

	1	2	3	4
1	2.42791	-2.00483	0.72594	-0.14901
2	1.00000	0.00000	-0.00000	0.00000
3	0.00000	1.00000	0.00000	-0.00000
4	-0.00000	0.00000	1.00000	0.00000
5	0.00000	-0.00000	0.00000	1.00000

Monthly Weights:

for K=0:

	1	2	3	4
1	1.00000	0.00000	-0.00000	0.00000
2	0.00000	1.00000	0.00000	-0.00000
3	-0.00000	0.00000	1.00000	0.00000
4	0.00000	-0.00000	0.00000	1.00000
5	-0.16394	0.78371	-2.07561	2.45584

for K=1:

	1	2	3	4
1	1.11660	-0.16169	0.05690	-0.01181
2	0.04094	1.01407	-0.06762	0.01261
3	-0.01405	0.07947	0.96653	-0.03196
4	0.01085	-0.05251	0.15580	0.88586
5	-0.14971	0.71584	-1.89921	2.33308

for K=2:

	1	2	3	4
1	1.23510	-0.32791	0.11717	-0.02437
2	0.09098	1.00797	-0.12221	0.02326
3	-0.02881	0.16900	0.91511	-0.05530
4	0.02048	-0.09944	0.30412	0.77485
5	-0.13549	0.64797	-1.72281	2.21033

for K=3:

	1	2	3	4
1	1.35502	-0.49752	0.17998	-0.03748
2	0.15002	0.98155	-0.16315	0.03159
3	-0.04344	0.26629	0.84775	-0.07060
4	0.02859	-0.13947	0.44317	0.66771
5	-0.12127	0.58012	-1.54643	2.08758

for K=4:

	1	2	3	4
1	1.47597	-0.66960	0.24463	-0.05100
2	0.21777	0.93516	-0.19030	0.03737
3	-0.05697	0.36865	0.76693	-0.07861
4	0.03490	-0.17123	0.57109	0.56524
5	-0.10706	0.51233	-1.37013	1.96487

for K=5:

	1	2	3	4
1	1.59763	-0.84340	0.31056	-0.06480
2	0.29380	0.86960	-0.20393	0.04053
3	-0.06833	0.47310	0.67547	-0.08023
4	0.03915	-0.19338	0.68599	0.46825
5	-0.09289	0.44469	-1.19404	1.84224

for K=6:

	1	2	3	4
1	1.71979	-1.01836	0.37736	-0.07878
2	0.37752	0.78605	-0.20466	0.04109
3	-0.07645	0.57645	0.57645	-0.07645
4	0.04109	-0.20466	0.78605	0.37752
5	-0.07878	0.37736	-1.01836	1.71979

APPENDIX B: Derivation of the solution of the general variant of the conversion problem

The problem is specified on a series interval containing I years and J "months" per year. The solution is first developped for correcting the reference periods of financial years; and then, of quarterly data.

Conversion of Yearly Series

Objective function (6.1) may be rwitten in matrix algebra as

(B.1) $$f(Z) = Z' X^{-1} A \, X^{-1} Z$$

where

(B.2) $$Z' = [\zeta_1 \ \zeta_2 \ \cdots \ \zeta_{IJ}]$$

(B.3) $$\underset{IJ\ IJ}{X} = \begin{bmatrix} 1/x_1 & 0 & 0 \ldots & 0 \\ 0 & 1/x_2 & 0 \ldots & 0 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 0 & 0 & . & 1/x_{IJ} \end{bmatrix}$$

(B.4) $$\underset{IJ\ IJ}{A} = D_2' D_2, \quad \underset{(IJ-2)\ IJ}{D_2} = \begin{bmatrix} 1 & -2 & 1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \ldots & 0 & 0 & 0 \\ . & . & . & . & & . & . & . \\ . & . & . & . & & . & . & . \\ . & . & . & . & & . & . & . \\ 0 & 0 & 0 & 0 & \ldots & 1 & -2 & 1 \end{bmatrix}$$

The unknowns ζ_t represent the sub-annual trend-cycle values for flow and index series, provided all financial years are regular, i.e. all refer to J consecutive "months". The unknowns are the product of the trend-cycle and of the seasonal components, and of the trading-day component if applicable, in ALL other cases: stock series, and flow or index series with irregular financial years.

Symbols x_t represent known values. They are equal to 1 in case of flow and index series, provided all financial years are regular. They are equal to an expert-supplied seasonal pattern, multiplied by a trading-day pattern if applicable, in ALL other cases. Matrix D_2 is the second difference operator; and matrix A, the corresponding quadratic second difference operator.

Constraints (5.2) are written as:

(B.5) $$H^{a\prime} Z = Y^{a},$$

where

(B.6)
$$\begin{array}{ll} & \text{columns:} \quad \tau_1 \qquad\quad \rho_1 \qquad\quad \tau_2 \qquad\quad \rho_2 \\ H^{a\prime} = & \left[\begin{array}{l} 0\;0\;\ldots\;0\;1\;1\;\ldots\;1\;0\;\ldots\;0\;0\;\ldots\;0\;0\;0\;\ldots \\ 0\;0\;\ldots\;0\;0\;0\;\ldots\;0\;0\;\ldots\;1\;1\;\ldots\;1\;0\;0\;\ldots \\ .\;.\qquad .\;.\;.\qquad .\;.\;.\;.\;.\;.\;.\qquad .\;.\;. \\ .\;.\qquad .\;.\;.\qquad .\;.\;.\;.\;.\;.\;.\qquad .\;.\;. \\ .\;.\qquad .\;.\;.\qquad .\;.\;.\;.\;.\;.\;.\qquad .\;.\;. \\ \end{array}\right] \\ M\;IJ & \end{array}$$

(B.7) $$Y^{a\prime} = [\; y^{a}_{1}\; y^{a}_{2}\; \ldots\; y^{a}_{M} \;]$$

Matrix $H^{a\prime}$ is a sum operator. With τ_m equal to ρ_m and appropriate values, the matrix can also act a month selector matrix for stock series.

Vector Y^a contains the available financial year values. These values have been multiplied by J in case of index series.

For simplicity, the constraints are introduced as a quadratic term in an augmented objective function:

(B.8) $$F(Z) = Z' X^{-1} A\, X^{-1} Z \;+\; g\, (H^{a\prime} Z - Y^{a})' (H^{a\prime} Z - Y^{a})$$

where g is sufficiently large (e.g. 1000) to cause the term to reach the minimum (zero), that is to make the term equivalent to constraints in practise.

Performing the matrix operations in (B.8) yields:

(B.9) $$F(Z) = Z' X^{-1} A\, X^{-1} Z \;+\; g\, Z' H^{a} H^{a\prime} Z - 2g\, Z' H^{a} Y^{a} + g\, Y^{a\prime} Y^{a}$$

In order for (B.9) to reach its minimum with respect to ζ_t, the derivatives of the objective function with respect to ζ_t must by definition equal zero:

(B.10) $$dF/d\zeta = 2\,A\,Z \;+\; 2g\, H^{a} H^{a\prime} Z \;+\; 2g\, H^{a}\, Y = 0$$

$$\Rightarrow \; [A + g H^{a} H^{a\prime}]\, Z = g\, H^{a} Y^{a}$$

This implies the estimated sub-annual values are linear combinations, i.e. weighted averages of the available financial values:

(B.11) $$\Rightarrow \; Z = [A + g H^{a} H^{a\prime}]^{-1}\; g\, H^{a} Y^{a} = \underset{IJ\;M}{W^{\zeta}}\, Y^{a}$$

The estimated conventional year values are then the annual sum (or the appropriate operation) of the estimates sub-annual values:

(B.12) $$Y^{d} = H^{d\prime}\, Z = H^{d\prime}\, W^{\zeta}\, Y^{a} = W^{y}\, Y^{a}$$

I M

where $H^{d\prime}$ is the I IJ annual sum operator (similar H^a) for flow and index series; and the appropriate observation selection matrix in case of stock series.

Correction of Quaterly Data

For correcting quarterly data, Y^a now contain the available quarterly values, J is equal to 3, and I represents the number of quarters in the series interval. The augmented objective function is still (B.8)

Doing the same developments as for the yearly conversion, yields a similar result. The estimated sub-annual values are linear combinations, i.e. weighted averages of the available financial values:

(B.13) $$Z = [A + gH^aH^{a\prime}]^{-1} \; g\, H^aY^a = \underset{I3\ M}{W^\zeta}\, Y^a$$

The estimated quarterly values are then the quarterly sums (or the appropriate operation) of the estimates sub-annual values:

(B.14) $$Y^d = H^{d\prime} Z = H^{d\prime} W^\zeta Y^a = \underset{I\ M}{W^y}\, Y^a$$

Matrix $H^{d\prime}$ is the I 3I quarterly sum operator for flow and index series

(B.15) $$\underset{I\ I3}{H^d} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & \dots & 0 & 0 & 0 \\ . & . & . & . & . & & & . & . & . \\ . & . & . & . & . & & & . & . & . \\ . & . & . & . & . & & & . & . & . \\ 0 & & 0 & & 0 & 0 & \dots & 1 & 1 & 1 \end{bmatrix};$$

and the appropriate monthly selector for stock series.