Working Paper No TRSA-87-015E
Time Series Research and Analysis Division
Methodology Branch

A NOTE ON AGGREGATING DATA REFERRING TO DIFFERENT FINANCIAL YEARS

by

Pierre A. CHOLETTE

and

John HIGGINSON

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

Telephone: (613) 951-0059

October 1987

Résumé

Cette note analyse l'effet d'agréger des données annuelles d'exercices financiers différents (ex.: juin à mai, septembre à août, etc). Les chiffres annuels résultant de pareille aggrégation sont des estimateurs biaisés des composantes cyclique et tendancielle des séries. La nature et la taille du biais varient avec les phases de la composante cyclique (l'expansion, le retournement, la récession et la reprise). Les biais diffèrent également d'une variable socio-économique à l'autre, détruisant ainsi les relations entre variables qui prévalaient avant l'agrégation. On conclut à la nécessité de convertir les chiffres d'exercices financiers en chiffres d'années civiles avant leur agrégation.

INTRODUCTION

Annual data reported to statistical agencies often pertain to the financial years of the respondents. Those financial years (generally) consist of any twelve consecutive months, for instance from June of one year to May of the following year. In most cases however, statisticians need annual data which reflect to the conventional year, extending from January to December. One practice with respect to financial year data is described by Bustros (1986): The data of respondents with any of the twelve possible financial years ending between April (say) of year i and March of year i+1 are assigned to year i; and for the variable considered, the annual value y_i is ruled to be the sum of all the data pertaining to twelve financial years assigned to year i. This note examines the effect of such a scheme on the trend-cycle and the seasonal components of series.

Annual estimates constructed according to that rule will be regarded as equivalent to taking weighted sums of the underlying monthly data. The underlying monthly data are not known, but - under certain assumptions - the weights of the weighted sum are known. The inherent properties of those weighted sums are analyzed. The weighted sums are found to produce biases in the trend-cycle component of the resulting annual series. For stock series, they also introduce seasonality.

1. FLOW SERIES WITH EQUALLY IMPORTANT FINANCIAL YEARS

Flow series are such that the annual values are the sum of the monthly values. For instance, the annual sales are the sum of the monthly sales. The rows of Table 1 represents the twelve possible financial years. For the financial year ending in April of year i (row 1) the "annual" sales of the financial year group is the sum of the sales in the months of May of year i-1, June of year i-1, and so forth to April of year i. Similarly, for the financial year ending in May of year i, the annual sales of the group is the sum of the sales in June of year i-1, to May of year i. Each number "1" in each row can be regarded as the weight applied to the unknown underlying monthly sales of the financial year group. The set of 1's in each row, can be regarded as the weights of the weighted sum (of the unkonwn underlying monthly sales) yielding the annual sales of the group considered. These weighted sums are trivial for the time being, but this will not be the case later on.

The results in this section are based on two assumptions.

1) The twelve financial year groups are equally important in terms of volume of the socio-economic variable considered - as opposed to number of respondents. For a given month, the sales in one group are equal to those in any other group. This assumption is equivalent to assigning a common value, for instance 1, to the different rows of Table 1. This unrealistic assumption will be relaxed in Section 2.

2) The unknown monthly sales are identical for each financial year group. For any given month considered, e.g. for July of year i-1, all groups realize the same sales. That volume may change from July to August of year i-1, but still in August all groups have same volume. In other words, the unknown monthly sales of each group are given by a common monthly time series. This assumption allows one to factorize the weights from each monthly value and to aggregate them column-wise into the implicit weights of Table 1.

Deriving the annual estimates y_i, described in the Introduction, is equivalent to applying the implicit weights, 1, 2, ..., 11, 12, 11, ..., 1, to the unknown underlying monthly sales common to all financial year groups. The inherent properties of that 23-term weighted sum are now analysed.

First, the implicit weights eliminate stable (and most moving) seasonality. Indeed, each same-month receives same sum of weights. As testified by the last row of Table 1, the two Mays in the 23-term interval receive a total sum of weights equal to 12; and that holds for all the other months. The same conclusion is reached on noticing that each row sum is free from seasonality (which cancels on 12 successive months); and that any linear combination of non-seasonal values is non-seasonal.

Second, as shown in Appendix, the implicit weights reproduce a linear trend-cycle of the underlying series over the 23-term interval. In other words, should the underlying series behave in a linear manner, the estimates would also behave linearly. However, the estimates y_i pertain to April instead of the middle of each target year i. Indeed April of year i is the month located in the middle of the 23-term interval. This can be visualized in Figure 1: The weights corresponding to the uniform distribution (equally important groups) display a pyramidal shape centered in April. (The weights appear in relative terms, i.e. divided by their sum.) The ideal weights, also displayed, would provide a central estimate. This defect of the implicit weights translates into phase-shifts or biases in the estimate y_i of the linear trend-cycle. By lagging behind, the weighted sum under-estimates the desired value, in periods of rise; and in periods of decline, over-estimates.

Third, the implicit weighted sum is much too rigid for socio-economic time series. Assume for a moment that the four phases of a business cycle encompass exactly 23 months - which is in reality possible. (Think of a sine wave of periodicity 23.) Any straight line specification - the implicit weights in particular - would completely eliminate such a cycle. Similarly it would eliminate at least part of longer cycles: it would cut through the peaks and troughs. The implicit weights therefore very seriously dampen the amplitude of business cycles. For those knowledgeable in moving averages, the implicit weights are shown in Appendix to be equivalent to a 12 by 12 moving average. It would be unthinkable to use such an average to estimate the trend-cycle of monthly series.

The two types of bias described , phase-shift and cycle reduction, clearly depend on the various phases of the business cycle of the target socio-economic variable. This is most serious since time series are supposed to support decision makers in their assessement and dating of the same cycles.

2. FLOW SERIES WITH UNEQUALLY IMPORTANT FINANCIAL YEARS

One could argue the conclusion reached in Section 1 is due to the two unrealistic assumptions made at the outset. Table 2 relaxes one of the assumptions. The distribution of the respondents over the ending-month groups therein is no longer uniform. The distribution selected is that of the Canadian retail trade sales in 1983. In that year, 31% of the sales were done by companies with financial year ending in December; 20%, in January; and so forth. Assumming that distribution is equivalent to filling each row of Table 2 with the relative weight of the group. Assumption 2 is maintained. As a result of the two assumptions, each financial group behaves identically except for a factor, and that factor is the relative importance of the financial year group. Put differently, all groups have a common **serial pattern**. The second assumption allows one to factorize and aggregate the weights to form the implicit weights.

The implicit weights of Table 2 are also represented in Figure 1, along with the uniform weights (of Table 1) and with the ideal weights. Thinking along the same lines as for the uniform distribution, one reaches the following conclusions. First, the implicit weights of Table 2 eliminate seasonality. Second (as explained in appendix), they are equivalent to a 12-term weighted sum of 12-term simple sums, that is to a 12-term weighted sums of central straight line estimates. They are therefore very rigid and cause reduction in the amplitude of the business cycle.

Note that a valid way to reason with the scheme is the following: If all the respondents were concentrated in the financial group ending in December, the implicit weights of Table 2 would coincide with the ideal weights, which are those of a simple 12-term average applied from January to December of year i. If the respondents were equally distributed (in terms of volume) between the December and January financial groups, the scheme would be equivalent to a 13-term 2 by 12 moving average centered on July; if they were equally distributed between the November, December and January groups, to a 14-term 3 by 12 moving average centered in the middle of year i; and so forth. The wider the distribution, the more rigid the implicit weights are.

Spectral analysis (Chatfield, 1975; Koopmans, 1974) could reveal the exact properties of the implicits weights of Table 2. That is not necessary however to draw conclusions: The implicit weights (and the resulting spectral properties) change from variable to variable - even for a given population of respondents. Indeed the relevant entries in the table are the relative importance of the twelve financial year groups <u>for the variable considered</u>. One could find the optimal (i.e. the least damageable) definition of year i for one variable, e.g. from June of year i-1 to May of year i, and that definition would be inappropriate for other variables pertaining to the same population. (The various respondent firms, being involved in different activities, do not display the same relations between variables.) Since the year definition cannot be changed from variable to variable, the scheme produces different biases which depend on the variable considered. One faces a set of variables which are consistent at the respondent level (i.e which satifactorily describes the respondent) but become inconsistent at the aggregate level. This is doubly unacceptable: First, the decision making process is based on the simultaneous consideration of several socio-economic variables. Second, each variable in isolation displays biases which depend on the phases of the trend-cycle.

Now the second assumption is relaxed: the financial year groups may follow different serial patterns. The situation becomes more speculative. A few assertions can nevertheless be made. First, adding different financial year data is still equivalent to some 23-term weighted average. The weights of the average can no longer be calculated, because they cannot be factorized from the underlying monthly values. Second, the weights change from year to year and from variable to variable. One would have to resort to simulations to do a more detailed analysis. It is very doubfull that any general and durable conclusions could be drawn.

3. STOCK SERIES

Table 3 and Figure 2 pertain to the situation which prevails for stock series under the assumption of common serial pattern. The distribution is still that of the retail trade sales (although it should that of the stock variable considered). The situation is worst than for flow series, because seasonality is neither eliminated nor preserved. Stock series are essentially seasonal, because they pertain to one month (usually the last) of the year. The ideal weight pattern should then preserve seasonality and would be a single spike equal to 100% in December (not represented in the figure).

CONCLUSION

Aggregating financial year data with different reference periods, produces

1) biased annual values for each variable considered in isolation - and the bias changes with the phase of the business cycle;

2) and mutually inconsistent annual values for variables considered together.

For stock series, the problem is compounded by the fact that seasonality is neither preserved nor eliminated.

It is therefore our opinion that financial year data must be converted to conventional year ("calendarized") before aggregation.

APPENDIX

Taking a simple average of twelve chronological observations is equivalent to fitting a constant to those observations by Ordinary Least Squares. It can be shown (Brown, 1963) that for the central point of the 12-term interval, a simple average is also equivalent to fitting a straight line (a first degree time polynomial) by O.L.S. In other words, the fitted values of the constant and of the straight line coincide for the central middle point. (The central point is located between the 6th and the 7th observation.)

In particular, taking the sum of the twelve monthly values defining a financial year is equivalent to fitting a straight line to those months, except for a factor equal to 12. Taking the sum of 12 chronologically ordered financial year totals is therefore equivalent to fitting a straight line on 12 central straight line estimates, except for a factor equal to 12. In the case of equal importance of the financial year groups of Table 1, adding 12 financial year values is then equivalent to a 12 by 12 average of the underlying monthly values (except for a factor equal to 144.) If the underlying series behaves linearly, so will the estimates. In the case of unequal importance of the financial year groups of Table 2, the scheme is equivalent to a 12-term weighted average of simple 12-term average. The weights of the weighted average are the relative importances of the groups.

REFERENCES

- Brown, R.G. (1963), **Smoothing, Forecasting and Prediction of Discrete Time Series**, Prentice-Hall.

- Bustros, J. (1986), "Reporting Arrangements Reference Period", **Statistics Canada**, Sub-Project Team 4.

- Chatfield, C. (1975), **The Analysis of Time Series: Theory and Practice**, Wiley.

- Koopmans, L.H. (1974), **The Spectral Analysis of Time Series**, Academic Press.

TABLE 1: Monthly weight distribution implicit in assigning to year i all respondents with financial year ending between April of year i and March of year i+1, assuming common serial pattern and equal importance of the financial yaear groups

Year:	i-1								i												i+1			Weight of
Month:	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	Group
Financial year groups																								
April	1	1	1	1	1	1	1	1	1	1	1	1												1
Mai		1	1	1	1	1	1	1	1	1	1	1	1											1
June			1	1	1	1	1	1	1	1	1	1	1	1										1
July				1	1	1	1	1	1	1	1	1	1	1	1									1
August					1	1	1	1	1	1	1	1	1	1	1	1								1
September						1	1	1	1	1	1	1	1	1	1	1	1							1
October							1	1	1	1	1	1	1	1	1	1	1	1						1
November								1	1	1	1	1	1	1	1	1	1	1	1					1
December									1	1	1	1	1	1	1	1	1	1	1	1				1
January										1	1	1	1	1	1	1	1	1	1	1	1			1
February											1	1	1	1	1	1	1	1	1	1	1	1		1
March												1	1	1	1	1	1	1	1	1	1	1	1	1
	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	Total
Implicit weights	1	2	3	4	5	6	7	8	9	10	11	12	11	10	9	8	7	6	5	4	3	2	1	144
Implicit same-month weights	12	12	12	12	12	12	12	12	12	12	12	12												

TABLE 2: Monthly weight distribution implicit in assigning to year i all respondents with financial year ending between April of year i and March of year i+1, assuming common serial pattern and unequal importance of the financial year groups (distribution of retail trade sales)

Year:	i-1								i												i+1			Weight of Group
Month:	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	
Financial year groups																								
April	3	3	3	3	3	3	3	3	3	3	3	3												3
Mai		4	4	4	4	4	4	4	4	4	4	4	4											4
June			3	3	3	3	3	3	3	3	3	3	3	3										3
July				4	4	4	4	4	4	4	4	4	4	4	4									4
August					5	5	5	5	5	5	5	5	5	5	5	5								5
September						4	4	4	4	4	4	4	4	4	4	4	4							4
October							5	5	5	5	5	5	5	5	5	5	5	5						5
November								2	2	2	2	2	2	2	2	2	2	2	2					2
December									31	31	31	31	31	31	31	31	31	31	31	31				31
January									...	20	20	20	20	20	20	20	20	20	20	20	20			20
February									...	...	7	7	7	7	7	7	7	7	7	7	7	7		7
March									...	...	...	12	12	12	12	12	12	12	12	12	12	12	12	12
	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	Total
Implicit weights	3	7	10	14	19	23	28	30	61	81	88	100	97	93	90	86	81	77	72	70	39	19	12	1200
Implicit same-month weights	100	100	100	100	100	100	100	100	100	100	100	100												

TABLE 3: Monthly weight distribution implicit in assigning to year i all respondents with financial year ending between April of year i and March of year i+1, assuming common serial pattern and unequal importance of the financial year groups, for a stock series

Year:	i-1								i												i+1			Weight of Group
Month:	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	
Financial year groups																								
April	0	0	0	0	0	0	0	0	0	0	0	3												3
Mai		0	0	0	0	0	0	0	0	0	0	0	4											4
June			0	0	0	0	0	0	0	0	0	0	0	3										3
July				0	0	0	0	0	0	0	0	0	0	0	4									4
August					0	0	0	0	0	0	0	0	0	0	0	5								5
September						0	0	0	0	0	0	0	0	0	0	0	4							4
October							0	0	0	0	0	0	0	0	0	0	0	5						5
November								0	0	0	0	0	0	0	0	0	0	0	2					2
December									0	0	0	0	0	0	0	0	0	0	0	31				31
January										0	0	0	0	0	0	0	0	0	0	0	20			20
February											0	0	0	0	0	0	0	0	0	0	0	7		7
March												0	0	0	0	0	0	0	0	0	0	0	12	12
	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	Total
Implicit weights	0	0	0	0	0	0	0	0	0	0	0	3	4	3	4	5	4	5	2	31	20	7	12	10[illegible]
Implicit same-month weights	4	3	4	5	4	5	2	31	20	7	12	3												

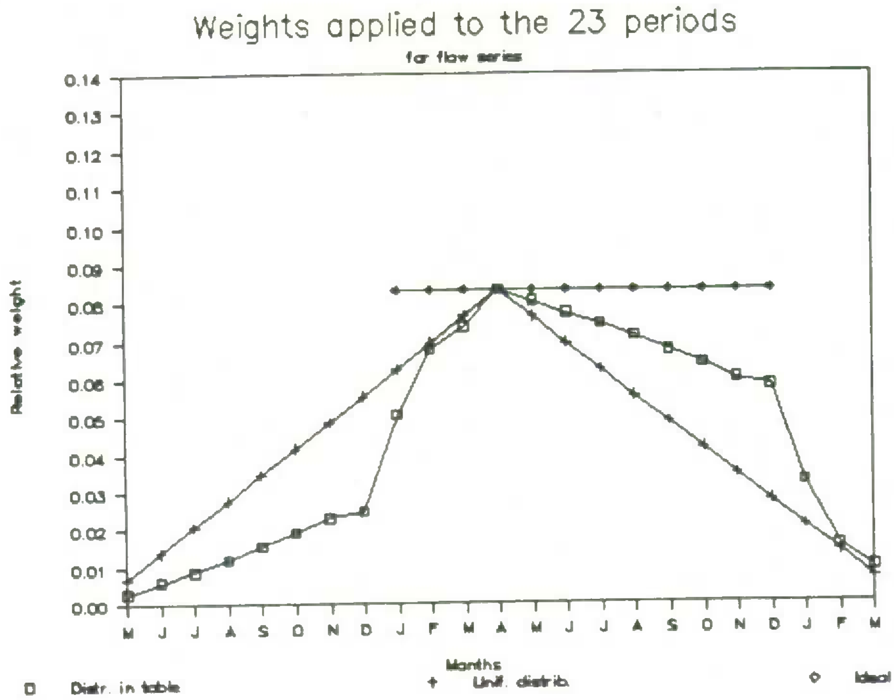

Figure 1: Relative implicit weight for flow series

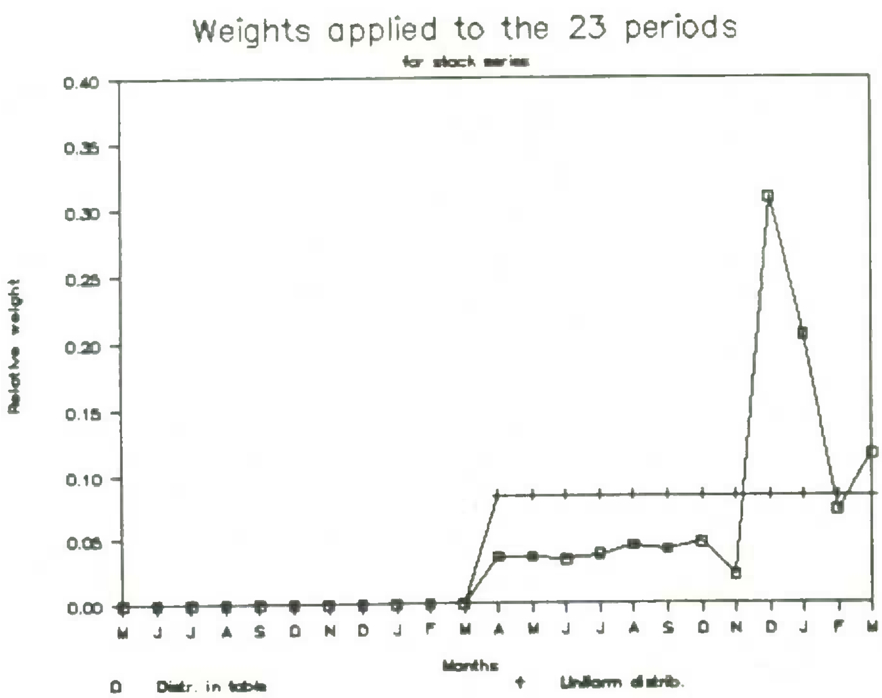

Figure 2: Relative implicit weights for stock series