LINKING AND BENCHMARKING THE REDESIGNED RETAIL AND WHOLESALE TRADE SERIES

by Pierre A. CHOLETTE

Revised May 15 1990

Time Series Research and Analysis Division
R.H Coats, 7th floor
Ottawa, Canada
K1A 0T6

tel.: (613) 951-1601

Working Paper TSRA-90-004E
Methodology Branch

LINKING AND BENCHMARKING THE REDESIGNED RETAIL AND WHOLESALE TRADE SERIES

by Pierre A. CHOLETTE

Revised May 15 1990

Time Series Research and Analysis Division
R.H Coats, 7th floor
Ottawa, Canada
K1A 0T6

tel.: (613) 951-1601

1. Introduction

The Retail and Wholesale trade time series are key measures of economic activity and provide major components of the National Accounts and of the Input-Output tables. During 1989 Statistics Canada conducted two parallel monthly Retail and Wholesale surveys: the *old survey*, based on earlier Standard Industrial Classifications (SIC), which has provided the estimates until now, and a *new survey*, based on a new sample design, on a new frame and on the 1980 SIC. Starting in 1990, the Retail and Wholesale estimates published for 1989 and for the ensuing years will originate from the new survey. For the years 1981 to 1989, the *old series* will be revised; the revised series (1) will reflect the 1980 SIC, (2) will be linked to the *new series* and (3) will be benchmarked to annual census values. The goals of the revision are to provide the best possible historical Retail and Wholesale estimates consistent with the new current series and to facilitate seasonal adjustment of the new series starting in January 1989. (Seasonal adjustment requires at least three years of continuous data.) This paper describes the techniques and methods used in the revision process. As for concepts, survey methodology and other details, one should refer to the Retail Trade and Wholesale Trade publications for January 1990 (Statistics Canada, 1990a 1990b).

For the purpose of linkage and benchmarking (as opposed to publication), the revised Retail Trade sales series form a system of 247 series for 19 *Trade Groups* and 13 *Provinces* or territories, listed in Table 1. These 247 series include the 19 *trade group totals* at the Canada level, the 13 *provincial totals* (of all trade groups). The first trade group is defined as the sum of the other *component trade groups* and the first "province" as the sum of the other *component provinces*, so that the first first trade group for the first province is in fact the Canada *grand total*. The 216 (18x12) *sub-provincial series* can be collapsed to any higher level of aggregation, by combining trade groups and/or provinces, and the collapsed system will still be consistent. It is also possible - without inconsistency - to publish some series in a given province where most of the other series are not published. At the time of writing, the decision had been made to publish 17 trade group totals (four trade groups being collapsed into two) and the 13 provincial totals; and the decision to publish at more detailed levels of aggregation would be made at a later date.

In the case of the Wholesale trade, the revised series form a system of 21 series, listed in Table 2, consisting of 10 trade group totals, 12 provincial totals, the first trade group for the first province being the grand Canada total. All 21 series will be published.

The strategy adopted to revise the old series was essentially the following:

1) Reclassify the old series according to the 1980 SIC. This involved going from a larger number of trade groups in the old series to a smaller number in the new. Details are provided in Section 2.
2) In order to minimize movement discontinuity, link each reclassified old series to the observation of January 1989 of the new series. The linked series coincides with the new series starting in January 1989 and is equal to old reclassified changed by a factor before that date. This factor change preserves the month-to-month growth rates of the old. Linkage is examined in Section 3.
3) Benchmark each linked series to *annual benchmarks* available for some years. Since the benchmarks originated from annual censuses and the sub-annual series from a survey, this step aimed at improving the reliability of the sub-annual series - in terms of level especially.

Benchmarking was performed in such a way to change the level of the linked series to that of the benchmarks, while preserving the original month-to-month growth rates as much as possible. Benchmarking is examined in Section 4.

4) Balance the system of series by means of "raking", so that the additivity of the system of series is restored. Indeed both linking and benchmarking mildly disturb additivity. Raking is explained in Section 5.

Section 6 discusses the whole revision exercise, in terms of other methods and stategies which could have been used and in terms of what oucomes to expect from those systems of series as more new data become available. The Appendix A and B provide mathematical details on benchmarking and raking; while Appendix supplies small but complete examples of benchmarking and raking programmed in SAS (SAS Institute, 1985).

2. Reclassifying the Old Series

Until 1989, the Retail and Wholesale series contained a larger number of trade *groupings*. Without entering to much detail, these groupings were called kinds of business (KOBs). The Retail Trade KOBs were a break-down of the 2-digit 1960 SIC classes (Standard Industrial Classification); whereas the Wholesale Trade KOBs were combinations of the 1970 3-digit SIC classes. The new trade groups are combinations of 4-digit 1980 SIC classes both cases.

In the case of the monthly Retail Trade series, 29 KOBs had to be mapped into 19 trade groups. Since some KOBs overlapped with more than one trade group, the reclassification entailed retabulating the sample files at a finer level of KOBs and combining the finer tabulations to obtain the 19 trade groups for each province.

In the case of the monthly Wholesale trade series, each KOB was included within one trade group. The reclassification process merely consisted of combining the 17 old KOB series into nine trade groups series.

For the annual benchmarks, the census files of both Retail and Wholesale were available on the basis of the 4-digit 1980 SIC. The new trade groups corresponded to combination of the SIC classes. This more precise KOB information - along with the fact that they originate from censuses - explains why the benchmarks are indeed considered as benchmarks.

3. The Linking Method Used

In order to minimize movement discontinuity, the old reclassified series was linked to the new series. The method chosen simply changed the level of the old series to that of the new series by one factor. This technique is identically equivalent to applying the month-to-month growth rates (the "trends") of the old to backcast the new series.

For Retail trade the ratio of the new series to the old reclassified series was calculated for each series at the *link point*, January 1989. Each linked series was then defined as the old reclassified series multiplied by that one linkage ratio, before January 1989; and, as the new series, starting in January 1989.

For Wholesale trade, ratios were calculated for each month of 1989 and the average was taken for each series. Each linked series was then defined as the old reclassified series multiplied by the *average* ratio, before January 1989; and, as the new series starting at that date. The difference of treatment between Retail and Wholesale is justified by the fact that the ratios were relatively erratic in the latter case.

For both Retail and Wholesale, the linkage procedure assumes that the new and the old series differ only in terms of level. The validity of the procedure therefore depends 1) on whether this assumption is true and 2) on whether the level difference observed - over 1 and 12 months respectively - is representative of the that really existing. As for the latter consideration, it is obvious that the linkage factors are subject to sampling error, which is larger at lower levels of aggregation.

It should also be pointed out that the linkage factors represent the *net* conceptual, methodological and other differences between the new and the old series. Consequently, a factor of 1.0 would not necessarily imply that the new and old series are equivalent in all respects.

4. The Benchmarking Method Used

The benchmarking method used is an extension of the proportional Denton (1971) method. The original proportional Denton method basically sets the benchmarked series proportional to the un-benchmarked (here the linked) series, with the restriction that the *annual* sums of the benchmarked series are equal to the annual benchmarks. The applicable extensions are the following:

1) The annual benchmarks may refer to fiscal years instead of calendar years. The restriction simply becomes that fiscal year (weighted) sums of the benchmarked series are equal to the "annual" benchmarks. Ignoring the fiscal attribute of the annual benchmarks (i.e. pretending they cover the calendar year) impairs their ability to reflect business cycle movements (Cholette and Higginson, 1987).
2) Sub-annual benchmarks referring to one month instead of one year may be specified. This feature proposed by Helfand, Monsour and Trager (1977) allows linkage to be performed within the benchmarking operation: The benchmarked series has to comply with the historical annual benchmarks, but also has to lead to a selected current value of the new series. The feature can also force the benchmarked series to start from a given value, ruled as un-alterable by the statistical agency (Ibid.).

For both Retail and Wholesale series, the annual benchmarks actually referred to mixtures of the fiscal years of the respondents. A respondent (to the annual census) with any of the twelve possible fiscal years ending between April 1986 (say) and March 1987 was classified in 1986. The benchmark for 1986 thus potentially covers 23 months: from May 1985 to March 1987. In the simple case where all respondents have a fiscal year ranging (say) from February 86 to January 87, the extended Denton method assumes that the benchmark covers those 12 months with weights equal to 1. In the case of a mixture of fiscal years, the benchmark covers more than twelve month, with some weights lower than 1. Department Stores in Ontario provide an informative example: in 1986, 20% of the retail sales were done by firms having a fiscal year coinciding with the calendar year; and 80%, by firms having a fiscal year ranging from February 86 to January 87. The benchmark is then specified to cover January 86 with weight 0.20; the months of February 86 to December 86, with weights equal to 1.00; January 87, with weight 0.80; and the other time periods with weight 0.0. The 1986 constraint then becomes that the fiscal year *weighted* sum of the benchmarked series is equal to the benchmark value. The appropriate calendar year 1986 value is then the calendar year sum of the benchmarked series.

At higher levels of aggregation, over trade groups especially, the weights tended to be positive for all the 23 reference periods. For the Canada grand total, the weights of the 1986 benchmark attributed to the months of May 85 to March 87 were respectively 0.04, 0.06, 0.09, 0.14, 0.17, 0.21, 0.24, 0.25, 0.66, 0.89, 0.94 1.00 0.96 0.94, 0.91, 0.86, 0.83, 0.79, 0.76, 0.75, 0.34, 0.11 and 0.06. The fiscal years were much more homogeneous for a given trade group than for provincial totals. This fact will be taken into account in the revison exercise.

More information on the derivation of the fiscal year weights can be found in Cholette and Higginson (1987). Mathematical and other details about the extended Denton benchmarking method are given in Appendix A.

4.1 Benchmarking the Retail Trade Series

In the case of Retail Trade Series, the benchmarks of two trade groups, Supermarket and Grocery Stores and Other Food Stores, pertained to the aggregate of the two groups. For benchmarking purposes, those two groups were then collapsed into one group: Food Stores. This momentarily reduced the number of series to 234: 18 trade groups and 13 provinces.

For Retail Trade, only one benchmark was used, that of year 1986. However, for the Canada grand total no benchmark was used: The benchmarked series was set equal to the linked. The philosophy was that the new - and therefore the linked - series was satisfactory for the grand total, and that only a distribution problem of the grand total over trade groups and over provinces existed. The trade group and provincial distribution indicated by the new series was not satisfactory in terms of level, and the distribution of the 1986 benchmark was deemed preferable. In other words benchmarking was used as a vehicle to impose the 1986 annual distribution on the system of series. However, the linked series of Department Stores and Liquor Wine and Beer Stores in all provinces and of Other General Merchandise Stores in Prince Edward Island were also deemed preferable and excluded from this redistribution: the benchmarked series were set equal to the linked. (The benchmarks of the "provincial" totals had been modified so the amounts excluded from the redistribution would not reappear in the benchmarked provincial totals. The amounts are equal to the dollar discrepancies between the benchmark and the fiscal year sums of the excluded linked sub-provincial series.)

Table 1 provides the *annual proportional discrepancies*, between the annual benchmarks and the fiscal year (weighted) sums of the linked Retail series, observed for the trade group totals and for the provincial. For the sub-provincial series, the discrepancies (not tabulated) were distributed in a more scattered manner.

Benchmarking the Retail Trade series produced a system of 234 series, where some of the components coincided with the new series starting in 1989 and other components were higher or lower than the new series by a multiplicative factor. Since there was only one benchmark that factor was equal (under the Denton method) to the corresponding annual proportional discrepancy, displayed in Table 1 for the totals.

Table 1: Annual proportional discrepancies observed for the trade group and provincial totals, in the case of the Retail Trade series (entries in %)

Trade Group Totals:	All Trade Groups (0)	Canada Total	(0)	.
	Food Stores (15 (10 + 20))	Canada Total	(0)	97.46
	Supermarket and Grocery Stores			.
	Other Food Stores			.
	Drug and Medicine (30)	Canada Total	(0)	98.59
	Shoe Stores (40)	Canada Total	(0)	110.59
	Mens Clothing (50)	Canada Total	(0)	109.69
	Women Clothing (60)	Canada Total	(0)	93.80
	Other Clothing (70)	Canada Total	(0)	125.62
	Furniture and Appliances (80)	Canada Total	(0)	107.25
	Other Household Furnishing (90)	Canada Total	(0)	101.00
	Motor Vehicles (100)	Canada Total	(0)	101.76
	Gas. Service Stations (110)	Canada Total	(0)	97.21
	Automotive Parts (120)	Canada Total	(0)	91.48
	Department Stores (131)	Canada Total	(0)	.
	Other General Merchandise (132)	Canada Total	(0)	108.82
	Other Semi-Durable Goods (140)	Canada Total	(0)	108.05
	Other Durable Goods (150)	Canada Total	(0)	103.43
	Liquor, Wine and Beer (161)	Canada Total	(0)	.
	All Other Retail (162)	Canada Total	(0)	134.43
Provincial Totals:	All Trade Groups (0)	Newfoundland (10)		107.66
	All Trade Groups (0)	Prince Edward Island (11)		98.48
	All Trade Groups (0)	Nova Scotia (12)		113.99
	All Trade Groups (0)	New Brunswick (13)		105.50
	All Trade Groups (0)	Quebec (20)		101.01
	All Trade Groups (0)	Ontario (30)		99.75
	All Trade Groups (0)	Manitoba (40)		108.57
	All Trade Groups (0)	Saskatchewan (47)		113.50
	All Trade Groups (0)	Alberta (48)		84.30
	All Trade Groups (0)	British Columbia (50)		107.94
	All Trade Groups (0)	Yukon (61)		105.01
	All Trade Groups (0)	North West Territories (62)		84.22

4.2 Benchmarking the Wholesale series

In the case of Wholesale Trade, annual benchmarks were available for the years 1981 to 1987. The benchmarks of 1981 were dropped however, because they covered time periods not included in the linked monthly series (May 1980 to December 1980). For Wholesale, the new series was much better than the old, so that the first observation of the each new series was considered and specified as a sub-annual benchmark. The resulting benchmarked series thus satisfied six annual benchmarks, each covering 23 overlapping periods (May 1981 to March 1983, May 1982 to March 1984, ..., May 1986 to March 1988), and one sub-annual benchmark forcing the benchmarked series to lead to the link point observation of January 1989. Table 2 contains statistics on the annual discrepancies.

Contrary to Retail, each of the 21 Wholesale Trade benchmarked series in the system coincided with the new series starting in January 1989. Under the linkage method chosen, identical series would have been achieved by benchmarking the old reclassified without generating a linked series. Indeed proportional benchmarking is insensitive to a proportional factor. Linkage was carried out, because we wanted to keep open the option of using another linkage method entertained (see discussion), which did not involve a mere proportional factor; as well as the option of publishing the linked instead of the benchmarked series.

Table 2: Statistics on the annual proportional discrepancies observed for the Wholesale Trade series

(entries in %)

			mean	std. dev.	min.	max.
Trade Group Totals:	All trade groups (0)	Can. Tot (0)	101.89	2.52	97.86	105.32
	(10) Food, Beverage, Drug, etc.	Can. Tot (0)	105.97	3.00	101.06	110.22
	(20) Apparel and Dry Goods	Can. Tot (0)	74.23	11.90	57.34	89.62
	(30) Household Goods	Can. Tot (0)	95.05	9.83	84.35	112.60
	(40) Motor Vehicles, Parts, etc.	Can. Tot (0)	85.93	2.10	82.10	89.17
	(50) Metals, Hardware, etc.	Can. Tot (0)	99.69	4.89	91.48	106.18
	(60) Lumber & Build. Mat. etc.	Can. Tot (0)	96.52	5.45	87.07	104.29
	(70) Farm Machinery, Equ., etc.	Can. Tot (0)	87.29	7.91	75.71	97.95
	(80) Other Machinery, etc.	Can. Tot (0)	101.95	4.53	95.04	108.77
	(90) Other Products	Can. Tot (0)	116.19	2.53	111.86	119.69
Provincial Totals:	All trade groups (0)	N.F.L (10)	99.65	5.31	88.66	104.91
	All trade groups (0)	P.E.I. (11)	93.44	10.15	82.05	112.71
	All trade groups (0)	N.S. (12)	84.29	2.43	80.40	87.91
	All trade groups (0)	N.B. (13)	88.43	5.80	82.06	98.66
	All trade groups (0)	Que. (20)	100.98	3.82	94.91	107.88
	All trade groups (0)	Ont. (30)	105.15	4.42	97.26	109.37
	All trade groups (0)	Man. (40)	103.70	7.21	89.06	111.61
	All trade groups (0)	Sask. (47)	104.22	8.85	89.83	114.30
	All trade groups (0)	Alb. (48)	103.21	2.99	99.12	106.44
	All trade groups (0)	B.C. (50)	96.80	4.37	90.29	103.75
	All trade groups (0)	Y. & N.W.T (60)	67.54	9.29	53.67	79.32

5. The Raking Method Used

When applied to systems of series, both proportional linking and benchmarking produce series which are mildly inconsistent: *Aggregation discrepancies* are observed between the (sum of the component) trade group totals and the grand total; between the provincial totals and the grand total; and in the case of the Retail Trade system, between the sub-provincial series and the provincial totals and between sub-provincial series and the trade group totals. Consistency was restored by means of raking. The raking variant is based on Generalized Least Squares adjustment, that is on an explicit and well defined objective function (see Appendix B). Raking was used in steps of one-dimensional raking followed, in the case of Retail Trade, by steps of two-dimensional raking.

One-dimensional raking is applicable when considering the Canada grand total against the trade group totals *or* against the provincial totals. *Two-dimensional raking* is applicable when considering the trade group totals and the provincial totals on the one hand, against the sub-provincial series on the other hand.

5.1 Raking the Retail Trade series

In the case of the Retail Trade series, benchmarking produced a system of 234 inconsistent series, classified in 18 trade groups and 13 provinces (including the totals). The raking steps were as follows.

Firstly, one-dimensional raking imposed the grand national total (namely the values of the linked series) onto the 17 component trade group totals. The objective function was such that the grand total, the Department Stores total, and the Wine and Beer Stores total were not *alterable* and so that the alterable values would change as little as possible in percentage. If for a month considered the aggregation discrepancy between the grand total and the trade group totals was +0.40% for instance, then each alterable trade group total was raised by a same percentage (e.g. 0.70%) and the unalterable totals were left untouched. The raking corrections made to the alterable benchmarked series had an *algebraic* mean of -0.52%, a standard deviation of 0.60%, a minimum of -2.11% and a maximum of 1.26%. For the corresponding linked series, the same statistics were respectively 0.82%, 0.64%, 0.00% and 3.18%. These statistics were calculated over time, that is on 108 observations.

Table 3: Statistics on the 108 aggregation discrepancies of the benchmarked series between the total indicated and the corresponding set of sub-provincial series (entries in %)

total		mean	std. dev.	min.	max.
Food Stores (15 (10 + 20))	Canada Total (0)	-0.46	0.58	-1.91	1.31
Drug and Medicine (30)	Canada Total (0)	-0.44	0.45	-1.46	1.00
Shoe Stores (40)	Canada Total (0)	-0.90	1.10	-4.23	0.86
Men Clothing (50)	Canada Total (0)	-0.55	0.75	-3.87	1.27
Women Clothing (60)	Canada Total (0)	-0.62	0.79	-2.62	1.45
Other Clothing (70)	Canada Total (0)	-0.66	0.84	-3.39	1.22
Furniture and Appliances (80)	Canada Total (0)	-0.89	0.80	-3.73	0.39
Other Household Furnishing (90)	Canada Total (0)	-1.39	1.29	-4.99	0.92
Motor Vehicles (100)	Canada Total (0)	-0.60	0.74	-2.87	1.39
Gas. Service Stations (110)	Canada Total (0)	-0.98	1.02	-3.72	1.21
Automotive Parts (120)	Canada Total (0)	-0.81	0.71	-2.65	1.18
Department-Stores (131)	Canada Total (0)	-0.04	0.05	-0.17	0.06
Other General Merchandise (132)	Canada Total (0)	-0.19	0.69	-1.68	1.63
Other Semi-Durable Goods (140)	Canada Total (0)	-0.67	0.84	-3.50	1.02
Other Durable Goods (150)	Canada Total (0)	-0.75	0.85	-3.33	0.86
Liquor, Wine and Beer (161)	Canada Total (0)	0.11	0.56	-0.72	1.99
All Other Retail (162)	Canada Total (0)	-0.65	0.59	-2.46	0.84
All Trade Groups (0)	Newfoundland (10)	-0.89	0.89	-2.36	1.96
All Trade Groups (0)	Prince Edward Island (11)	-2.37	1.99	-7.67	2.21
All Trade Groups (0)	Nova Scotia (12)	-0.38	0.50	-2.58	0.48
All Trade Groups (0)	New Brunswick (13)	-0.83	0.41	-2.11	-0.19
All Trade Groups (0)	Quebec (20)	-0.35	0.79	-2.70	1.92
All Trade Groups (0)	Ontario (30)	-0.35	0.58	-1.83	0.96
All Trade Groups (0)	Manitoba (40)	-0.55	0.42	-1.85	0.31
All Trade Groups (0)	Saskatchewan (47)	-1.71	1.54	-5.95	1.29
All Trade Groups (0)	Alberta (48)	-0.56	0.95	-8.03	1.06
All Trade Groups (0)	British Columbia (50)	-1.06	0.75	-3.05	0.79
All Trade Groups (0)	Yukon (61)	-3.04	8.26	-18.45	11.79
All Trade Groups (0)	North West Territories (62)	-1.58	1.53	-6.08	1.17

Secondly, one-dimensional raking imposed the grand Canada total onto the 12 component provincial totals in an identical manner, except only the grand total was unalterable. The objective function was such that raking amounted to the following very simple arithmetic procedure. If for a month considered the aggregation discrepancy between the grand total and the provincial totals was 0.50% for instance, then each provincial total was raised by 0.50% and the grand total was left untouched. The raking corrections made to the altered benchmarked series had an algebraic mean of -0.53%, a standard deviation of 0.04%, a minimum of -0.68% and a maximum of -0.40%. For the linked series, the same statistics were -0.07%, 0.10%, -0.31% and 0.06%.

In order to determine whether raking introduced spurious seasonality in the trade group and provincial totals, a two-factor analysis of variance (ANOVA) of the raking corrections was performed. The two factors were the month and the year. In the case of the trade group totals, both factors were significant with F-values of 10.5 and 78.0 for the month and the year respectively; in the case of the provincial totals, only the year effect was significant, with F-values of 0.8 and 14.4. Raking thus spuriously modified the seasonal pattern of the component trade group totals. The seasonal peaks and trough of that spurious seasonality occurred in November with a value of +0.5% (i.e. each November got raised by 0.5% on average) and in May with a value of -0.8%. This is further discussed in Section 6.4.

Table 4: Statistics on the aggregation discrepancies of the linked series between the total indicated and the corresponding set of sub-provincial series (entries in %)

		mean	std. dev.	min.	max.
Food Stores (15 (10 + 20))	Canada Total (0)	0.45	0.56	-0.17	2.72
Drug and Medicine (30)	Canada Total (0)	0.67	0.95	-0.75	5.04
Shoe Stores (40)	Canada Total (0)	0.92	0.69	-0.25	3.05
Men Clothing (50)	Canada Total (0)	-1.77	1.41	-5.70	1.35
Women Clothing (60)	Canada Total (0)	0.60	0.72	-0.71	2.76
Other Clothing (70)	Canada Total (0)	0.49	0.64	-0.68	2.53
Furniture and Appliances (80)	Canada Total (0)	1.51	0.87	-0.10	3.93
Other Household Furnishing (90)	Canada Total (0)	-0.44	1.07	-2.86	2.31
Motor Vehicles (100)	Canada Total (0)	0.68	0.83	-1.33	3.38
Gas. Service Stations (110)	Canada Total (0)	-0.15	0.77	-1.85	2.31
Automotive Parts (120)	Canada Total (0)	0.59	0.66	-0.38	3.00
Department Stores (131)	Canada Total (0)	-0.04	0.05	-0.17	0.06
Other General Merchandise (132)	Canada Total (0)	0.49	0.55	-0.32	2.48
Other Semi-Durable Goods (140)	Canada Total (0)	0.57	0.63	-0.34	2.90
Other Durable Goods (150)	Canada Total (0)	0.58	0.77	-0.47	3.69
Liquor, Wine and Beer (161)	Canada Total (0)	0.11	0.56	-0.72	1.99
All Other Retail (162)	Canada Total (0)	0.49	0.76	-1.03	2.65
All Trade Groups (0)	Newfoundland (10)	-1.26	1.40	-4.19	2.53
All Trade Groups (0)	Prince Edward Island (11)	-0.47	1.32	-4.94	2.00
All Trade Groups (0)	Nova Scotia (12)	-1.11	1.01	-3.43	0.72
All Trade Groups (0)	New Brunswick (13)	-0.05	0.58	-1.62	1.29
All Trade Groups (0)	Quebec (20)	0.47	0.62	-0.81	2.14
All Trade Groups (0)	Ontario (30)	0.38	0.37	-0.13	1.91
All Trade Groups (0)	Manitoba (40)	0.24	0.23	-0.12	0.89
All Trade Groups (0)	Saskatchewan (47)	0.15	0.88	-1.83	2.81
All Trade Groups (0)	Alberta (48)	1.83	2.38	-1.32	10.08
All Trade Groups (0)	British Columbia (50)	0.06	0.41	-1.00	0.91
All Trade Groups (0)	Yukon (61)	7.18	8.74	-8.98	27.93
All Trade Groups (0)	North West Territories (62)	-4.31	2.53	-9.87	0.79

Two-dimensional raking was then used to impose both the 17 component trade group totals and the 12 component provincial totals, obtained by one-dimensional raking, onto the corresponding 204 sub-provincial series. The objective function was such that the sub-provincial series were allowed to change equally but as little as possible in percentage, except for Department Stores and Liquor, Wine and Beer Stores. Those two sub-provincial components were allowed to change only to add to the corresponding national trade group totals. As a result the corrections made to those series were the same regardless of the province (for a given month). For instance if the aggregation discrepancy between the Department Stores total and the sum of the sub-provincial Department Stores values was 0.20%, the latter values were raised by 0.20%, even in provinces where all the other components had to be substantially reduced or increased to sum to the provincial total. It would be cumbersome to provide statistics on the corrections made to each of the 204 sub-provincial series. However, by providing statistics on the aggregation discrepancies between the totals and the sub-provincial series, Table 3 gives an idea of the corrections needed to balance the system. Not surprisingly, the corrections required are more substantial in certain provinces. Table 4 displays the corresponding statistics for the linked series.

Finally, two-dimensional raking was used again in splitting the benchmarked Food Store aggregate into its two components, Supermarket and Grocery Stores and Other Food Stores. These two food components which had been collapsed for benchmarking purposes, were first separated according to the proportion of each individual component in the old reclassified aggregate. For instance, if for a month considered Supermarket and Grocery Stores accounted for 90% of the old reclassified aggregate Food Store, the "benchmarked" Supermarket and Grocery stores was set equal to 90% of the benchmarked aggregate value for that month; and Other Food Stores, to 10% of the same aggregate value. This particular separation process (convex combination) ensured additivity of the two food components to aggregate Food Store within each province. However it did not ensure additivity of sub-provincial Supermarket and Grocery Stores and Other Food Stores series to the corresponding two national trade group totals. Two-dimensional raking imposed (or preserved) the appropriate totals in the manner explained for the other sub-provincial series, except all components were allowed to change equally. The aggregation discrepancies give an idea of the corrections required. For the Supermarket and Grocery Stores they had an algebraic mean of 0.06%, a standard deviation of 0.02%, a minimum of -0.04% and a maximum of 0.11%. For the Other Food Stores the same statistics were -0.72%, 0.24%, -1.35% and 0.50%. (No linked series had been generated for the two trade groups.)

5.2 Raking the Wholesale Trade Series

In the case of the Wholesale Trade series, benchmarking produced a systems of 21 inconsistent series, nine component trade group totals and 12 provincial totals, the first one being the Canada grand total. The raking steps were as follows.

Firstly, one-dimensional raking reconciled the grand Canada total with the 9 trade group totals. The objective function was such that, if for instance for a month considered the aggregation discrepancy was +0.50% then the grand total was reduced by 0.25% and each trade group total was raised by 0.25%. The alterability of the grand total is justified by the fact that the benchmarks have much more homogeneous fiscal years at the trade group level than at the national level. The raking corrections made to the benchmarked trade group totals had an *algebraic* mean of -0.02%, a standard deviation of 0.21%, a minimum of -0.62% and a maximum of 0.61%. For the corresponding linked series, the same statistics were -0.67%, 0.46%, -2.01% and 0.46%.

Secondly, one-dimensional raking imposed the grand Canada total just obtained to the 11 provincial totals. The objective function was such that raking amounted to the following very simple arithmetic procedure. If for a month considered the aggregation discrepancy between the grand total and the provincial totals was 0.50% for instance, then each provincial total was raised by

0.50% and the grand total was left untouched. The raking corrections made to the altered benchrnarked series had an algebraic mean of 0.17%, a standard deviation of 0.18%, a minimum of 0.00% and a maximum of 0.79%. For the linked series, the same statistics were 0.56%, 0.38%, 0.00% and 1.58%.

6. Discussion

This section discusses some aspects of the methods and strategies adopted to link, benchmark and balance the Retail and Wholesale Trade series and considers some alternative methods.

6.1 linkage

The series in the systems of series were first linked. At least three linkage methods were considered: proportional level linkage, proportional level and seasonal linkage and additive linkage.

Proportional Level Linkage - Section 3 ended by clarifying the implications and the assumptions behind the linkage method chosen: namely that in 1989 the new and the old series differ only in terms of level. If this assumption is true - or verified as more observations of the new series become available -, proportional level linkage enables successful seasonal adjustment (by means of the X-11-ARIMA method, Dagum, 1988) of the benchmarked linked series. (Benchmarking has very little impact on seasonal adjustment.) If the new and the old also differ in terms of seasoanl patter or of outliers, i.e. strong irregular movements, which are present in either series but not in the other (for the same period of time), a variety of problems may occur: distortions of the year-to-year movement between 1988 and 1989, noisy seasonally adjusted values for 1989, heavy revisons for 1988, 1989 and 1990, as more of the new observations become available. The revisions should stabilize after 1991.

Proportional Level and Seasonal Linkage - Another linkage method was considered which assumes that the new and the old differ both in terms of level and seasonal pattern. The method is designed to preserve year-to-year growth rates between same-months (Cholette, 1984) and amounts to this simple arithmetic operation. Over the 12-month overlap period of 1989, the 12 ratios between the new and the old are calculated for each series; all the old monthly values are then multiplied by the corresponding same-month ratio. This procedures modifies both the level and the seasonal pattern of the old series to be consistent with those of the new series. A variant of this method was used in the 1970s to link some Canadian Labour Force series: only selected months were modified (another variant would be to modify those months by a fraction of the amount). If the assumption of this alternative more general linkage method are true, namely that the discrepancy observed is one of level and seasonality, then the method allows a successful seasonal adjustment of the (benchmarked) linked series despite the change of seasonal pattern. If the new and the old also differ in terms of outliers, which are present in either one but not in the other, a variety of problems may result: distortions of the year-to-year movement between 1988 and 1989, imposition of the wrong seasonal pattern to the old series, heavy revisons for 1988, 1989 and 1990, as more of the new observations become available. The revisions should stabilize after 1992.

The preference of one linkage method over the other is determined by the probability of the assumptions underlying each method being true and by the desirability of the possible outcomes associated to each method. The simple proportional level linkage method was preferred. This simplified scenario analysis, which does not take into account trading-day variation, barely gives a flavour of the possible outcomes under the two linkage method.

It is interesting to note that with level and seasonal linkage, many of the potential problems described would tend to be experienced one year later. Indeed under such linkage, the distortions would be forced in the old and therefore interpreted by the seasonal adjustement program as seasonality, and on seasonally adjusting 1989 would seem successful. The program would reinterpret those distortions only *after* the 1990 data are incorporated, that is in 1991. With level linkage, on the other hand, any outlier or seasonal discrepancy would cause problems on seasonally adjusting 1989.

Additive Linkage - Additive level and additive level and seasonal linkage methods were also considered. These are identical to the two proportional methods examined, except the linkage factors are the differences (instead of the ratios) between the new and the old. Additive linkage has the advantage of preserving the additivity of a system of series; but entails the following unacceptable disadvantages: a) negative linked values may be (and were) obtained, b) the relative proportions between linked components of the system may be (and were) totally alien to the proportions observed in the new and in the old series. Proportional linking on the other hand imposes the relative proportions of the new on the old series for the periods of overlapping; and for the preceding periods, the proportions lie in between those of the new and those of the old.

6.2 Benchmarking

Proportional Benchmarking - The linked series were then benchmarked. For the same reasons as for linkage, proportional benchmarking (with the extended Denton method) was preferred to additive benchmarking. As explained earlier, proportional benchmarking sets the benchmarked series as proportional to the un-benchmarked (here the linked) as permitted by the benchmarks. In the case of Retail Trade, the benchmarked series was perfectly proportional to the linked, because there was only one benchmark. In such a case, the benchmarked series also has the same month-to-month growth rates as the linked, which in turn had the same growth rates as the old. In the case of Wholesale Trade, where several benchmarks were available, proportional benchmarking is an approximation to growth rate benchmarking (Monsour and Trager, 1979), although proportional benchmarking is justifiable per se.

Fiscal Years - The more innovative aspect of the benchmarking variant used in this paper, consists of specifying fiscal year benchmarks as covering the month they actually refer to. The alternative is to specify them as covering the calendar year, which is an obvious specification error, especially when all the respondents in the census (from which the benchmark originates) have a common fiscal year, e.g. April to March.

Preliminary Benchmarking - The extended Denton method also allows preliminary benchmarking. In fact all the Retail estimates obtained subsequent to 1986 are preliminarily benchmarked. Preliminary benchmarking amounts to the multiplying the un-benchmarked values by the ratio of the benchmarked to the un-benchmarked in the last reference period of the last annual benchmark, namely March 1987 in the case considered. This practice of preliminary benchmarking is common among statistical agencies (e.g. Bureau of the Census, 1986, p.5).

Future Revisions - As explained in Section 4, the variant of benchmarking used also allows sub-annual benchmarks. This feature may be used when the system of series is revised, on the availability of annual benchmarks for 1987 and 1988. The existing December 1985 (say) published value will be specified as a sub-annual benchmark, thus forcing the future revised series to start from that "historical" value. It is also possible that the 1987 benchmark will be pooled with the 1986 benchmark. The resulting 1986-87 benchmark would then refer to 35 months (May 85 to March 88) instead of 23, and the month-to-month movement of the un-benchmarked series would not be altered, because there would be only one benchmark for each series.

Simultaneous Benchmarking - It is possible to extend proportional benchmarking of the Denton type to process systems of series (Cholette, 1988). All the series in the system are benchmarked simultaneously as they would be individually, except that the benchmarked series also have to satisfy constraints of additivity. The method did not prove successful: the benchmarked series tended to drift from the un-benchmarked. In other words, the corrections made to the un-benchmarked displayed trends which were not justified by the behaviour of benchmarks.

6.3 Raking

Linkage and benchmarking disturb the additivity of a system of series; additivity was restored by means of raking. As explained in Section 5 and in Appendix B, the raking variants used are based on a specific objective function, as originally introduced by Deming and Stephan (1940). Given the computing capabilities of the time, their approach was then approximated iteratively to become what has been known as Iterative Proportional Fitting or raking. We reverted to the objective function approach and also incorporated therein the alterability coefficients of the Federal Reserve Board (1962, pp. 1400-1404). These coefficients gives control on which series will be allowed to change to satisfy additivity and by how much in relative terms. Examples are Department Stores and Liquor, Wine and Beer Stores, which the subject matter experts had good reason to keep relatively intact. In the Federal Reserve example, the alterability coefficients were given by the relative volatility of the series in the system (Taylor, 1963): the more volatile series could absorb a larger share of the adjustment necessary to satisfy additivity, compared to the smoother series. Coefficients provide a good vehicle for subject matter expertise.

It is worthwhile noting that raking was applied to the benchmarked series and to the linked series after the benchmarking operation. In other words, the linked series were not raked before they were benchmarked. This sequence of operations was designed to avoid changing the movement of the series twice: after linking and after benchmarking. Indeed as observed in Section 5, raking does affect the month-to-month movement of the series.

6.4 Systems of Time Series

In order to preserve additivity, systems of time series are usually processed at the lower levels of aggregation, and the totals are obtained by summation of the components, thus imposing the components on the total. The Canadian labour force series, for instance, are seasonally adjusted at the age-sex level; the official seasonally adjusted the national total is then defined as the sum of the seasonally adjusted components. In many situations, this *indirect approach* will lead to a total poorer than that which could have been obtained by directly seasonally adjusting the total. This strategy may also break down. In the case considered, the official national total is not equal to the sum of the seasonally adjusted provincial totals (which are adjusted directly). As already stated these aggregation issues also arise with linkage and benchmarking.

The approach to additivity adopted in this paper is somewhat opposite to the indirect approach: the totals were also processed directly but then imposed on the components. In the case of Retail Trade, the system was linked and benchmarked, at all levels of aggregation, and the grand total direct estimates were then imposed on both the trade group totals and on the provincial totals, and these marginal totals were then imposed onto the sub-provincial estimates by means of raking. The advantage of this strategy are the following:

a) the system displays no aggregation discrepancies and can be collapsed to any higher level of aggregation, without inconsistencies,
b) the totals are probably better than those obtained by summing the less reliable components.

The disadvantage of the direct approach is that, in order to restore additivity, the components have to be disturbed. As explained in Section 5.1, raking spuriously modified the seasonal pattern of the component trade group totals. One question is whether the indirect approach would have produced better totals and better sub-provincial components. Our initial intention was to adopt a mixed strategy: a) obtain the grand total indirectly by summation of the trade group totals, on the ground that the benchmarks had much more homogenous fiscal years within component trade groups and that the monthly trade group total data are good enough to sustain benchmarking; b) impose the grand indirect total just obtained onto the provincial totals; c) then impose all the totals onto the sub-provincial series. That strategy would have eliminated the need to alter any of the trade group totals (in a possibly spurious manner). But the linked grand linked total was deemed preferable as such and imposed on the trade group totals.

Another issue is that, after raking, the individual series no longer satisfy their annual benchmarks. However, the residual annual discrepancies are very small. Past experience with artificial series indicates that, under normal conditions, iterating between benchmarking and raking converges to a solution satisfying both additivity and the benchmarks. In the case of Retail and Wholesale Trade, no such iteration was attempted. The reason is that the annual benchmarks reflect a mixture of fiscal years; consequently the calendar year sums of the raked-benchmarked series probably provides better annual estimates. It could be also argued that raking operates as a cross-validation mechanism: The resulting system of series "*incorporates[,] into each cell[,] data from the entire system and has a probability of having improved, on balance, the accuracy...*" (Federal Reserve Board, 1962, p. 1401). Although valid for the annual estimates, the argument would in our opinion not be valind to the above-mentioned seasonal distortions, for the series considered at least.

APPENDIX A: Mathematical Details on Benchmarking

For a given series, the extended Denton proportional method minimizes

$$\text{(A.1)} \quad \underline{x} \sum_{t=2}^{T} (b_t / x_t - b_{t-1} / x_{t-1})^2 + \sum_{m=1}^{M} k_m \left[\left(\sum_{t=\tau_m}^{\kappa_m} w_{t,m}\, b_t\right) - y_m\right]^2 / y_m,$$

where b_t and x_t ($t=1,\ldots,T$) respectively denote the benchmarked and the un-benchmarked (i.e. the linked here) series. The first term of the objective function thus specifies the benchmarked series proportional to the un-benchmarked. This is also equivalent to preserving the month-to-month "growth rates" in the un-benchmarked series (to the extent that the annual discrepancies are constant).

In the second term of equation (A.1), the y_m's ($m=1,\ldots,M$) denote the fiscal year benchmarks, which in this paper potentially covers 23 time periods, τ_m to κ_m, e.g. from May of the year preceding the *target year* to March of the year following the target year. The parameters $w_{t,m}$ are weights between 0 and 1 explained in Section 4. Parameters k_m ($m=1,\ldots,4$) are pre-specified constants reflecting the degree of reliability of the benchmark. A large value of k_m can make a benchmark y_m virtually binding, which is adequate for reliable benchmarks; and a low weight, non binding, which is adequate for non-reliable benchmarks. In this paper, the benchmarks are binding.

In matrix algebra, objective function (A.1) is written:

$$\text{(A.2)} \qquad F(b) = b'Q\, b + (y - W\, b)'\, R\, (y - W\, b),$$

where $Q = X^{-1}D'DX^{-1}$,

where X^{-1} is a T by T diagonal matrix with diagonal elements $\sqrt{\underline{x}} / x_t$, $\underline{x}=\Sigma x_t/T$,

where D is the quasi first difference matrix operator

$$\underset{T \text{ by } T}{D} = \begin{bmatrix} (1-\rho^2)^{1/2} & 0 & 0 & 0 & \cdots \\ -\rho & 1 & 0 & 0 & \cdots \\ 0 & -\rho & 1 & 0 & \cdots \\ 0 & 0 & -\rho & 1 & \cdots \\ . & . & . & . & \cdots \\ . & . & . & . & \cdots \\ . & . & . & . & \cdots \end{bmatrix},$$

where ρ is smaller but very close to 1.0 (e.g. $\rho=0.99999$),

where $R=K\ Y^{-1}$, where Y^{-1} and K are both M by M diagonal matrix with diagonal elements $1/y_m$ k_m respectively,

and where matrix W has dimensions M by T and contains the weights $w_{t,m}$. In the simple case of a 16-term quarterly series (starting in a 1st quarter), for which 20% of the sales are done by firms with a fiscal year coinciding with the calendar year and 80% by firms with a fiscal year ending in the first quarter of each, matrix W would be

$$W = \begin{bmatrix} 0.2 & 1.0 & 1.0 & 1.0 & 0.8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.2 & 1.0 & 1.0 & 1.0 & 0.8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.2 & 1.0 & 1.0 & 1.0 & 0.8 & 0 & 0 & 0 \end{bmatrix}.$$

Minimizing (A.2) with respect to b yields the normal equations:

$$dF/db = 2\,Q + 2\,W\,R\,y + 2\,W'R\,W\,b = 0$$

which imply solution

(A.3) $$b = (Q + W'R\ W)^{-1} W'R\ y.$$

Solution (A.3) can be written

(A.4) $$b = Q^{-1} W' (W Q^{-1} W' + R^{-1})^{-1} y,$$

where the elements of $e_{r,c}$ of matrix Q^{-1} are then known algebraically: $(x_r x_c \rho^{|r-c|})/(\underline{x}\ (1-\rho^2))$. The elements $e_{r,c}$ of $Q^{-1}W'$ and of $WQ^{-1}W'$ are then also known algebraically. They are respectively:

$$e_{r,c} = \Sigma_{n=1}^{T} w_{c,n} (x_r x_n \rho^{|r-n|})/(\underline{x}\ (1-\rho^2)),\ r=1,...,T;\quad c=1,...M,$$

$$e_{r,c} = \Sigma_{k=1}^{T} w_{r,k} (\Sigma_{n=1}^{T} w_{c,n} (x_k x_n \rho^{|k-n|})/(\underline{x}\ (1-\rho^2))),\ r=1,...,M;\quad c=1,...M.$$

The the elements of R^{-1} are also known directly. Note that solution (A.4) requires a much smaller matrix inversion than (A.3) and also allows R^{-1}=0, in which case the benchmarks are exactly binding as in this paper.

Note that (A.3) and (A.4) can respectively be written

(A.3') $$b = x + (Q + W'R\ W)^{-1} W'R\ (y - Wx),$$

(A.4') $$b = x + Q^{-1} W' (W Q^{-1} W' + R^{-1})^{-1} (y - Wx),$$

which are extensions of the more familiar forms of benchmarking solutions encountered in the literature (e.g. Denton, 1971). Solutions (A.3') and (A.4') are the ones applicable to additive benchmarking, in which case $X^{-1} = I_T$ and $Y^{-1} = I_M$.

Table C.1 displays a small benchmarking example programmed in SAS/IML (SAS Institute, 1985). The program generates a quarterly artificial series x, benchmarks y and a fiscal sum operator matrix W, and then benchmarks the quarterly series, using solution (A.4). Each benchmark refers to the four quarters of the year with weights 0.2, 1.0, 1.0 and 1.0; and, to the first quarter of the following year with weight 0.8.

APPENDIX B: Mathematical Details on Raking

The consistency of a system of series is often restored by means of raking. Raking may be specified through Generalized Least Squares (e.g. Copeland, Peitzmeier and Hoy, 1987; Luery, 1986): an objective function is minimized subject to aggregation constraint, and a linear solution is reached. Typically, the sum of squared percentage correction to the series in the system is minimized. Furthermore in this paper, an alterability coefficient (Federal Reserve Board, 1962) is incorporated in the objective function for each series. This coefficient determines by how much each series may be corrected: A coefficient of zero ensures that the series considered, whether total or component - is not altered. Positive coefficients equal for all series cause all the series to be equally modifiable in percentage. The alterability coefficients play the same role as the power allocation parameters of Bankier (1988).

In this paper, one-dimensional and two-dimensional situations are distinguished. One-dimensional situations occur when the series in the system are classified using one classifying variables, namely trade groups *or* provinces. Two-dimensional situations occur when the series in the system are classified using two classifying variables, namely trade groups *and* provinces.

B.1 One-Dimensional Raking

In a situation where the series are classified in R classes, the system has to satisfy only one constraint: $z_1 = \Sigma_{r=2}^{R} z_r$. The one-dimensional raking variant used in this paper then minimizes the following objective

$$F(z,\lambda) = \sum_{r=1}^{R} x_r/a_r \, [(z_r - x_r) / x_r]^2 + 2\lambda \, [(\sum_{r=2}^{R} z_r) - z_1],$$

or equivalently

$$\text{(B.1)} \quad F(z,\lambda) = \sum_{r=1}^{R} x_r/a_r \, [(z_r / x_r) - 1]^2 + 2\lambda \, [(\sum_{r=2}^{R} z_r) - z_1],$$

where z_1 and x_1 respectively denote the unknown (consistent) and the actual un-raked grand total values, where z_r and x_r, r=2,...,R, denote the unknown and the un-raked component values, where a_r denote the alterability coefficients and where λ stand for the Lagrangian multiplier. This function is minimized for each period of time separately (this is why the time index is dropped):

Objective function (B.1) is now written in matrix algebra:

$$\text{(B.2)} \qquad F(z,\lambda) = (X^{-1} z - \iota)' \, X A^{-1} (X^{-1} z - \iota) + 2\lambda B' z,$$

where X and X^{-1} are R by R diagonal matrix with diagonal elements x_r and $1/x_r$ respectively,
where A^{-1} is a R by R diagonal matrix with diagonal elements $1/a_r$,
where ι is a R by 1 vector with elements equal to 1,
where B is a R by 1 vector with first element equal to -1 and the other elements equal to 1.

Performing the multiplication in (B.2) and cancelling X with X^{-1} where applicable yields

$$\text{(B.2')} \qquad F(z,\lambda) = z X^{-1} A^{-1} z + \iota' X A^{-1} \iota - 2 z' A^{-1} \iota + 2\lambda B' z.$$

Minimizing with respect to z and λ yields the normal equations:

$$dF/dz = 2 X^{-1} A^{-1} z - 2 A^{-1} \iota + 2 B \lambda = 0$$

$$dF/d\lambda = 2 B' z = 0$$

which imply solution

$$\text{(B.3)} \qquad \begin{bmatrix} z \\ \lambda \end{bmatrix} = \begin{bmatrix} (XA)^{-1} & B \\ B' & 0 \end{bmatrix}^{-1} \begin{bmatrix} A^{-1}\iota \\ 0 \end{bmatrix}.$$

Using identities on the inverse of symmetric matrices (e.g. Searle and Hausman, 1970, p. 115), one can find the following solution for z

(B.4) $z = [\, X\,A + X\,A\,B\,(-B'X\,A\,B)^{-1}\,B'X\,A]\;A^{-1}\iota \;=\; x - X\,A\,B\,(B'X\,A\,B)^{-1}\,B'x.$

Since for one-dimensional raking $B'X\,A\,B$ is a scalar, one can write

(B.5) $z = x - X\,A\,B\,B'\,x \,/\, B'X\,A\,B.$

Given the contents the various vectors and matrices, and noting that $B'x = x_2 + x_3 + \ldots + x_R - x_1$ and that $B'X\,A\,B = x_1 a_1 + x_2 a_2 + \ldots + x_R a_R$, one finds an explicit algebraic formula for the desired raked values:

(B.6) $$z_r = x_r + \delta\, x_r\, a_r \left(\sum_{k=2}^{R} x_k - x_1 \right) / \left(\sum_{k=1}^{R} x_k\, a_k \right),$$

where $\delta=1$ for $r=1$ (the total) and $\delta=-1$ for $r>1$.

Note that solutions (B.6) and (B.5) do allow values of a_r and x_r equal to zero.

One-dimensional raking by means of (B.5) was used in this paper, for both Retail and Wholesale, to reconcile the trade group totals with the provincial totals.

The Wholesale Trade Series - In reconciling the grand Wholesale total with the nine (component) trade group totals, all unraked values had alterability coefficients a_r equal to 1; the grand total was then *endogenous*, i.e. modified in the raking process. If for a given month the aggregation discrepancy was 0.50%, the grand total got reduced by 0.25%; and the other trade group totals, raised by 0.25%. In reconciling the grand Wholesale total with the provincial totals, the grand total just obtained had $a_1=0$ and the provincial totals had $a_r=1$ ($r>1$); the grand total was then *exogenous* to the raking process. If for a given month the aggregation discrepancy was 0.50%, the grand total remained unchanged and the other provincial totals got raised by 0.50%.

The Retail Trade Series - For Retail Trade, the grand total (the linked values) was always exogenous, and so were the Department Stores total and the Liquor, Wine and Beer Stores total. In reconciling the grand total with the 17 trade group totals, these three exogenous totals had $a_r=0$ ($r=1$, 13, 17); and the other totals had $a_r=1$. Consequently, the three exogenous totals remained intact, and the other totals were changed equally in percentage. In reconciling the grand total with the 12 provincial totals, the former had $a_1=0$ and the latter had $a_r=1$, like for Wholesale.

The one-dimensional raking solution (B.6) (and (B.4)) is then valid for both exogenous and endogenous totals and components.

Table C.2 displays a small example of one-dimensional raking programmed in SAS/IML. As shown by the sample printout, the total is imposed onto the components by means of a zero alterability coefficient; the components are raised by 4.17%.

B.2 Two-Dimensional Raking

In the two-dimensional variant raking is used in this paper, the raked series has to satisfy the following aggregation constraints for each period of time

(B.7) $$z_{g,1} = \sum_{p=2}^{P} z_{g,p}, \quad g=1...,G,$$

(B.8) $$z_{1,p} = \sum_{g=2}^{G} z_{g,p}, \quad p=1,...,P,$$

where $z_{g,1}$ $g=1,...,G$ are the trade group totals, $z_{1,p}$, $p=1,...,P$ are the provincial totals and where $z_{g,p}$, $g=2,...,G$; $p=2,...,P$ are the sub-provincial values.

The totals $z_{g,1}$ and $z_{1,p}$ are first specified as endogenous, because this makes solution (B.4) obtained for one-dimensional raking directly applicable and for other reasons given later. Endogenous two-dimensional raking minimizes an objective function very similar to (B.1):

(B.9) $$F(z,\lambda) = \sum_{g=1}^{G} \sum_{p=1}^{P} (x_{g,p} / a_{g,p}) [(z_{g,p} / x_{g,p}) - 1]^2 + \sum_{g=1}^{G} \lambda_g \left(\sum_{p=2}^{P} z_{g,p} - z_{g,1} \right) + \sum_{p=1}^{P-1} \lambda_p \left(\sum_{g=2}^{G} z_{g,p} - z_{1,p} \right),$$

where the last constraint from (B.8) is omitted because redundant. This objective function is now written in matrix algebra:

(B.10) $$F(z,\lambda) = (X^{-1} z - \iota)' X A^{-1} (X^{-1} z - \iota) + 2 \lambda B' z,$$

where $z'=[z_{11}\ z_{21} \dots z_{G1}\ z_{21}\ z_{22} \dots z_{G2}\ z_{13}\ z_{23} \dots z_{G3} \dots]$
where X is a GP by GP diagonal matrix with diagonal elements $x_{11}\ x_{21} \dots x_{G1}\ x_{12}\ x_{22} \dots x_{G2} \dots$
where A^{-1} is a GP by GP diagonal matrix with diagonal elements $1/a_{11}\ 1/a_{21} \dots 1/a_{G1}\ 1/a_{12}\ 1/a_{22} \dots 1/a_{G2} \dots$
where ι is a GP by 1 vector with elements equal to 1,
where B is a GP by (G+P-1) trade group and provincial sum matrix. For G=3 and P=5, B would be

$$\underset{(G+P-1)\ \text{by}\ GP}{B'} = \begin{bmatrix} -1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ -1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

Since (B.10) has the same form as (B.2), solution (B.4) still applies:

(B.11) $$z = x - \underset{GP\ \text{by}\ (G+P-1)}{X A B} \ \underset{(G+P-1)\ \text{by}\ (G+P-1)}{(B' X A B)^{-1}} \ \underset{(G+P-1)\ \text{by}\ 1}{B' x} = x - XAB + Q^{-1} R.$$

Matrix $Q=B'XAB$ is symmetric, is known algebraically and will be inverted by parts. The elements $e_{m,m}$ of the G by G partition Q_{11} and of the P-1 by P-1 partition Q_{22} of Q are respectively:

$$e_{m,n} = \begin{cases} \Sigma_{p=1}^{P} x_{m,p}\, a_{m,p}, & m=n=1,\dots,G \\ 0, & m \neq n, \end{cases}$$

(B.12)

$$e_{m,n} = \begin{cases} \Sigma_{g=1}^{G} x_{g,n}\, a_{g,n}, & m=n=1,\dots,P\text{-}1 \\ 0, & m \neq n. \end{cases}$$

Note that Q_{11} and Q_{22} are diagonal so that Q_{11}^{-1} and Q_{22}^{-1} are known directly. The elements $e_{g,p}$ of the G by P-1 partition Q_{12} of Q are

(B.13) $$e_{g,p} = \delta_{g,p}\, x_{g,p}\, a_{g,p}, \quad \delta_{g,p} = \begin{cases} 1 \text{ if } g=p=1, & \text{or if } g>1 \text{ and } p>1 \\ -1 \text{ if } g=1 \text{ and } p>1, & \text{or if } g>1 \text{ and } p=1. \end{cases}$$

The corresponding partitions of $M=Q^{-1}$ are then

(B.14) $$M_{22} = (Q_{22} - Q_{12}'Q_{11}^{-1} Q_{12})^{-1}, \quad M_{12} = -Q_{11}^{-1} Q_{12} M_{22}, \quad M_{11} = Q_{11}^{-1} - M_{12} Q_{12}' Q_{11}^{-1}$$

Matrix R=B'x is also known algebraically. The elements e_m of the G by 1 partition R_1 and of the P-1 by 1 partition R_2 of R are respectively:

(B.15) $$e_g = -x_{g,1} + \Sigma_{p=2}^{P} x_{g,p},\ g=1,\dots,G; \qquad e_p = -x_{1,p} + \Sigma_{g=2}^{G} x_{g,p},\ p=1,\dots,P\text{-}1.$$

Defining $H_1 = M_{11} R_1 + M_{12} R_2$ and $H_2 = M_{12}' R_1 + M_{22} R_2$ (H_1 and H_2 being the partitions of $H=Q^{-1}R$), solution (B.11) is written

(B.11') $$z_{g,p} = x_{g,p} - x_{g,p}\, a_{g,p}\, (\gamma_g h_{1,g} + \delta_p h_{2,p}),\ g=1,\dots,G,\quad p=1,\dots,P$$

where $\gamma_g = \begin{cases} -1 \text{ if } g=1, \\ 1 \text{ if } g>1 \end{cases}$ $\qquad \delta_p = \begin{cases} -1 \text{ if } p=1, \\ 1 \text{ if } p>1 \end{cases}$

where $h_{1,g}$ and $h_{2,p}$ are the elements of vectors H_1 and H_2 and $h_{2,P}=0$. Under solution (B.11') the only actual matrix inversion required is that of the P-1 by P-1 matrix $(Q_{22} - Q_{12}'Q_{11}^{-1} Q_{12})$ in (B.14). If the system contains 18 trade groups and 13 provinces (G=18 and P=13), then matrix to be inverted has dimensions a 12 by 12.

The matrix to be inverted will be singular, if all the alterability coefficients equal zero for a given value of g or of p. However, one can impose the grand total, on the other totals; and practically impose all totals, on the sub-provincial components, by selecting $a_{11}=0$, $a_{g1}=0.001$ (g>1), $a_{1p}=0.001$ (p>1) and $a_{gp}=1$ (g,p>1). Table C.3 provides such an example. The grand and the component trade group totals obtained are virtually the same as in the example of Table C.2, which imposed the grand total on the trade group totals, by means of one-dimensional raking.

When the totals are strictly exogenous (e.g. pre-established by one-dimensional raking), the number of trade groups and of provinces, G and P, are reduced by 1. One can show that, solution (B.11) becomes:

(B.16) $$z = x - X A B\ (B' X A B)^{-1}\ [\,-z. + B' x\,] \quad = x - XAB + Q^{-1} R,$$

where z, x, X and A are as in (B.11), except that g=1 and p=1 now denote the first *component* trade group and the first *component* province (instead of the totals),
where z. contains G+P-1 of the already established exogenous totals $z.=[z_{1,.}\ z_{2,.}\ \cdots\ z_{G,.}\ z_{.,1}\ z_{.,2}\ \cdots\ z_{.,P-1}]$,
where B is a GP by (G+P-1) trade group and provincial sum matrix. For G=2 and P=4

$$\underset{(G+P-1)\ \text{by}\ GP}{B'} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}.$$

The partitions Q_{11}, Q_{22}, Q_{12} and the corresponding partitions M_{11}, M_{22}, M_{12} of Q^{-1} are still given by (B.12), (B.13) and (B.14), except $\delta_{g,p}$ in (B.13) is replaced by 1. The partitions R_1 and R_2 of R=-z.+B'x are now given by

(B.17) $e_g = -z_{g,.} + \Sigma_{p=1}^{P} x_{g,p}$, g=1,...,G; $e_p = -z_{.,p} + \Sigma_{g=1}^{G} x_{g,p}$, p=1,...,P-1.

Solution (B.16) may writen

(B.16') $z_{g,p} = x_{g,p} - x_{g,p}\, a_{g,p}\, (h_{1,g} + h_{2,p})$, g=1,...,G, p=1,...,P, where $h_{2,P}=0$.

The Retail Trade Series - Solution (B.16') was the one used, as discussed in 5.1, to impose the G=17 Retail Trade component trade group totals and the P=12 component provincial totals onto the 204 sub-provincial benchmarked and linked values. The coefficients of alterability for Department Stores and Liquor, Wine and Beer Stores were set equal to 0.001 in all provinces. As a result, the corrections made to those components were almost the same (within 0.1%) in every province, even in those where the other components were changed substantially. Coefficient of alterability thus provided simple and effective control over each series in the system.

A virtually identical solution could have been reached, by means of endogenous raking, i.e. with solution (B.11') with appropriate alterability coefficients as described above. This is illustrated by the example of Table C.4, which gives sub-provincial values virtually identical to those of Table C.3. In retrospect, endogenous raking would have been logistically easier and more economical overall. Indeed raking with exogenous totals entailed first separating and processing the totals, by means of one-dimensional raking; and then, separating the exogenous components from the endogenous and processing the endogenous sub-provincial components, by means of two-dimensional raking with exogenous totals.

Solution (B.16') was also used in the case of the food sector, as discussed in Section 5.1. The Supermarket and Grocery Stores and the Other Food Store totals (G=2) and the P=12 provincial Food Stores totals were imposed onto the 24 sub-provincial Supermarket and Grocery Stores and Other Food Stores values. All alterability coefficients were set equal to 1.

APPENDIX C: SAS/IML Statements to Perform Benchmarking and Raking

This appendix provides SAS/IML statements (version 5.0) which effect benchmarking, as described in Appendix A, and raking, as described in Appendix B. The statements are also valid under Version 6.0 of SAS/IML, if "(|" and "|)" are replaced by left and right square brackets respectively. The variable names used in the SAS/IML statements parallel the notation used in Appendix A and B. One should refer to these appendices for the content of the variables and for the explanation of the examples.

The examples provided are only intended as hints to statisticians already familiar with SAS. Namely, in a real situation, the data would of course not be embedded in the SAS coding like in the examples, but would be read from files; potential users are expected to build their own interface between the data files and the algorithms illustrated. In order to do so, the following features of SAS/IML are recommendable:

a) Each variable from a SAS dataset becomes a vector when read in IML (unless otherwise specified).
b) The "edit/read point pointer_variable" feature allows one to read only the needed observations of the SAS dataset, e.g. those pertaining to all years and months of one series, or those pertaining to one year and one month of all series.
c) Similarly the "edit/replace point" feature allows one to replace the relevant observations of the SAS dataset by values calculated in IML. This requires that the variables have been initialized in the SAS dataset.

Under that scheme, the cycle a) to c) is repeated with appropriate values of pointers.

TABLE C.1: Benchmarking a quarterly series, with the method presented in Appendix A

```
options nocenter;
proc iml;
start Qinvgen;   * subroutine ;
     TT=nrow(X); MM=nrow(perl); Xbar=sum(X)/TT;
     rho=0.999999; denom=xbar#(1-rho##2);
     QinvW=j(TT,MM,0); WQinvW=j(MM,MM,0); rQuinv=j(1,TT,0);
     do r=1 to TT;
        do c=1 to TT; expo=abs(c-r);
           rQuinv(|1,c|)=(rho##expo)#X(|r|)#X(|c|)/denom; end;
        do m=1 to MM; t1=Tau(|m|); t2=Kappa(|m|);
           QinvW(|r,m|)=rQuinv(|t1:t2|)'*W'; end; end;

     do r=1 to MM; t1=Tau(|r|); t2=Kappa(|r|);
        do c=1 to MM;
        WQinvW(|r,c|)=W*QinvW(|t1:t2,c|); end; end;
     finish;

* generation of artificial quarterly data (normally read from files);
     X={100, 150, 125, 175,   200, 225, 200, 250,
        275, 325, 300, 375,   425, 450, 425, 450};

* generation of fiscal year benchmarks (normally read from files);
     Y={693.0, 1028.5, 1534.5};

     * reference periods of benchmarks; Tau=do(1,12,4)'; Kappa=Tau+4;

     * weights with which the benchmarks cover the quarters; W={.2  1  1  1 .8};

* calculation of the benchmarked series, according to (A.4);

     run Qinvgen; b=QinvW*inv(WQinvW)*Y;

* display of data and results;  Cor=b/X; print X b Cor Y Tau Kappa;
quit; run;
```

```
printed output:
 X        COL1  B          COL1  COR        COL1  Y         COL1  TAU        COL1  KAPPA    COL1

 ROW1    100.0  ROW1      110.0  ROW1     1.1000  ROW1     693.0  ROW1     1.0000  ROW1     5.0000
 ROW2    150.0  ROW2      165.0  ROW2     1.1000  ROW2    1028.5  ROW2     5.0000  ROW2     9.0000
 ROW3    125.0  ROW3      137.5  ROW3     1.1000  ROW3    1534.5  ROW3     9.0000  ROW3    13.0000
```

```
ROW4     175.0  ROW4     192.5  ROW4     1.1000
ROW5     200.0  ROW5     220.0  ROW5     1.1000
ROW6     225.0  ROW6     247.5  ROW6     1.1000
ROW7     200.0  ROW7     220.0  ROW7     1.1000
ROW8     250.0  ROW8     275.0  ROW8     1.1000
ROW9     275.0  ROW9     302.5  ROW9     1.1000
ROW10    325.0  ROW10    357.5  ROW10    1.1000
ROW11    300.0  ROW11    330.0  ROW11    1.1000
ROW12    375.0  ROW12    412.5  ROW12    1.1000
ROW13    425.0  ROW13    467.5  ROW13    1.1000
ROW14    450.0  ROW14    495.0  ROW14    1.1000
ROW15    425.0  ROW15    467.5  ROW15    1.1000
ROW16    450.0  ROW16    495.0  ROW16    1.1000
```

TABLE C.2: One-dimensional raking with the method presented in Appendix B

```
proc iml;      * one-dimensional raking;

RR=5; * number of classes;

* generation of alterability coefficient, once for the system;
   A=j(RR,1,1); A(|1|)=0.00;

* calculation of matrix B of equation (B.5) calculated once for the system of series;
    B=j(RR,1,1); B(|1|)=-1;

* generation of artificial data to be raked, normally read from files for each period of time;
     x={   1000.0,   192.0,   144.0,   384.0,   240.0};

 * calculation of the raked series, for each period of time according to equaltion (B.5);
        XAB=X#A#B;
        z=x - (XAB*(B'*x))/(B'*XAB);

 * display of the inputs and of the results for the sake of illustrating the example;
    Cor=(z/x)#100;
    print z(|format=10.2|) x(|format=10.2|)
    A(|format=10.2|) Cor(|format=10.2|);
quit; run;
```

```
printed output:
  Z        - COL1   X         COL1   A         COL1   COR        COL1
  ROW1   1000.00   ROW1   1000.00   ROW1      0.00   ROW1     100.00
  ROW2    200.00   ROW2    192.00   ROW2      1.00   ROW2     104.17
  ROW3    150.00   ROW3    144.00   ROW3      1.00   ROW3     104.17
  ROW4    400.00   ROW4    384.00   ROW4      1.00   ROW4     104.17
  ROW5    250.00   ROW5    240.00   ROW5      1.00   ROW5     104.17
```

TABLE C.3: Two-dimensional raking with endogenous totals using the method presented in Appendix B

```
options nocenter;
proc iml;      * two-dimensional raking with endogenous totals;

start rakeEndo;      * subroutine executed for each period of time;
   * initialization of partitions;
     Q11=j(GG,1,0);
     Q22=j(PP-1,PP-1,0); Q12=j(GG,PP-1,0);
     R1=j(GG,1,0); R2=j(PP-1,1,0);

   * diagonal elements of partition 1,1 of Q;
     do g=1 to GG;
     do p=1 to PP; gp=(p-1)#GG+g;
     Q11(|g|)=Q11(|g|)+X(|gp|)#A(|gp|); end; end;

   * partition 2,2 of Q;
     do p=1 to PP-1;
     do g=1 to GG; gp=(p-1)#GG+g;
     Q22(|p,p|)=Q22(|p,p|)+X(|gp|)#A(|gp|); end; end;

   * partition 1,2 of Q;
     do g=1 to GG; do p=1 to PP-1;
     gp=(p-1)#GG+g;
     delta=1; if g=1 & p>1 then delta=-1;
              if p=1 & g>1 then delta=-1;
     Q12(|g,p|)=X(|gp|)#A(|gp|)#delta; end; end;
```

```
   * inversion of Q by parts;
     inPart11=diag(1/Q11);
     M22=inv(Q22-Q12'*inPart11*Q12);
     M12=-inPart11*Q12*M22;
     M11=inPart11-M12*Q12'*inPart11;

   * partition 1 of R;
     do g=1 to GG;
     R1(|g|)=-X(|g|);
     do p=2 to PP; gp=(p-1)#GG+g;
     R1(|g|)=R1(|g|)+X(|gp|); end; end;

   * partition 2 of R;
     do p=1 to PP-1; gp=(p-1)#GG+1;
     R2(|p|)=-X(|gp|);
     do g=2 to GG; gp=(p-1)#GG+g;
     R2(|p|)=R2(|p|)+X(|gp|); end; end;

   * calculation H=inv(Q)*R;
     H1=M11*R1+M12*R2;
     H2=M12'*R1+M22*R2;
     H2=H2//J(1,1,0);

   * calculation of the raked series;
     do g=1 to GG;
     del2=1; if g=1 then del2=-1;
     do p=1 to PP; gp=(p-1)#GG+g;
     del1=1; if p=1 then del1=-1;
     Z(|gp|)=X(|gp|)-X(|gp|)#A(|gp|)#(del1#H1(|g|)+del2#H2(|p|)); end; end;
finish;

GG=5; PP=4; * number of groups and provinces;

* generation of alterability coefficient, once for the system;

     A=j(GG#PP,1,1); A(|1|)=0; A(|2:GG|)=0.001;
                     A(|6|)=0.001; A(|11|)=0.001; A(|16|)=0.001;

* generation of artificial data to be raked
  normally read from SAS datasets as a vector, for each period of time separately;
      X={   1000.0,    192.0,    144.0,    384.0,    240.0,
             441.0,     49.0,     97.0,    144.0,    147.0,
             343.0,     96.2,     47.6,    145.6,     49.0,
             196.0,     50.5,      0.0,     95.2,     49.0};

      Z=j(GG#PP,1,0); * normally initialized in a SAS dataset;

     run rakeEndo;  * calculation of raked series according to solution (B.11'), for each period of time;

  * display of the inputs and of the results in a two-way table,
    for the sake of illustrating the example (normally not done);
     Xtab=shape(X,PP,GG)';
     Ztab=shape(Z,PP,GG)';
     print Xtab(|format=10.2|) Ztab(|format=10.2|);
     Ctab=(Ztab/Xtab)#100;
     Atab=shape(A,PP,GG)';
     print Ctab(|format=10.2|) Atab(|format=10.3|);
quit; run;
```

printed output:

	XTAB	provinces COL1	COL2	COL3	COL4
g	ROW1	1000.00	441.00	343.00	196.00
r	ROW2	192.00	49.00	96.20	50.50
u	ROW3	144.00	97.00	47.60	0.00
p	ROW4	384.00	144.00	145.60	95.20
s	ROW5	240.00	147.00	49.00	49.00

ZTAB	COL1	COL2	COL3	COL4
ROW1	1000.00	450.00	350.00	200.00
ROW2	200.00	50.01	98.55	51.44
ROW3	150.00	100.50	49.50	0.00
ROW4	400.00	149.55	151.77	98.68
ROW5	250.00	149.95	50.17	49.88

CTAB	COL1	COL2	COL3	COL4
ROW1	100.00	102.04	102.04	102.04
ROW2	104.17	102.06	102.45	101.86
ROW3	104.17	103.61	103.99	0.00
ROW4	104.17	103.85	104.24	103.65
ROW5	104.17	102.00	102.39	101.80

ATAB	COL1	COL2	COL3	COL4
ROW1	0.000	0.001	0.001	0.001
ROW2	0.001	1.000	1.000	1.000
ROW3	0.001	1.000	1.000	1.000
ROW4	0.001	1.000	1.000	1.000
ROW5	0.001	1.000	1.000	1.000

TABLE C.4: Two-dimensional raking with exogenous totals using the method presented in Appendix B

```
options nocenter;
proc iml;      * two-dimensional raking with exogenous totals;
start rakeExo;                     * subroutine;
   * initialization of partitions;
     Q11=j(GG,1,0);
     Q22=j(PP-1,PP-1,0); Q12=j(GG,PP-1,0);
     R1=j(GG,1,0); R2=j(PP-1,1,0);

   * diagonal elements of partition 1,1 of Q;
     do g=1 to GG; do p=1 to PP;
     gp=(p-1)#GG+g;
     Q11(|g|)=Q11(|g|)+X(|gp|)#A(|gp|); end; end;

   * partition 2,2 of Q;
     do p=1 to PP-1; do g=1 to GG;
     gp=(p-1)#GG+g;
     Q22(|p,p|)=Q22(|p,p|)+X(|gp|)#A(|gp|); end; end;

   * partition 1,2 of Q;
     do g=1 to GG; do p=1 to PP-1;
     gp=(p-1)#GG+g;
     Q12(|g,p|)=X(|gp|)#A(|gp|); end; end;

   * inversion of Q by parts;
     inPart11=diag(1/Q11);
     M22=inv(Q22-Q12'*inPart11*Q12);
     M12=-inPart11*Q12*M22;
     M11=inPart11-M12*Q12'*inPart11;

   * partition 1 of R;
     do g=1 to GG;
     R1(|g|)=-Zp(|g|);
     do p=1 to PP; gp=(p-1)#GG+g;
     R1(|g|)=R1(|g|)+X(|gp|); end; end;

   * partition 2 of R;
     do p=1 to PP-1;
     R2(|p|)=-Zp(|GG+p|);
     do g=1 to GG; gp=(p-1)#GG+g;
     R2(|p|)=R2(|p|)+X(|gp|); end; end;

   * calculation H=inv(Q)*R;
     H1=M11*R1+M12*R2;
     H2=M12'*R1+M22*R2;
     H2=H2//J(1,1,0);

   * calculation of the raked series;
     do g=1 to GG; do p=1 to PP;
     gp=(p-1)#GG+g;
     Z(|gp|)=X(|gp|)-X(|gp|)#A(|gp|)#(H1(|g|)+H2(|p|)); end; end;
finish;

GG=4; PP=3; * number of component groups and component provinces;

* generation of alterability coefficient, once for the system;
   A=j(GG#PP,1,1);

* generation of artificial data to be raked,
  normally read from SAS dataset as a vector for each period of time separately;
     X={      49.0,     97.0,    144.0,    147.0,
              96.2,     47.6,    145.6,     49.0,
              50.5,      0.0,     95.2,     49.0};

* generation of the the exogenous totals, normally read from SAS datasets as a vector
  for each period of time separetely;
     Zp={   200.0,    150.0,    400.0,    250.0,
            450.0,    350.0};

     Z=j(GG#PP,1,0); * normally initialized in a dataset;

    run rakeExo; * calculation of the raked series according to solution (B.16') for each period of time;

 * display of the inputs and of the results in a two-way table,
   for the sake of illustrating the example (normally not done);
     Xtab=shape(X,PP,GG)';
     Ztab=shape(Z,PP,GG)';
     print Xtab(|format=10.2|) Ztab(|format=10.2|);
     Ctab=(Ztab/Xtab)#100;
     Atab=shape(A,PP,GG)';
```

```
    print Ctab(|format=10.2|) Atab(|format=10.2|);
quit; run;
```

printed output:

XTAB	COL1	COL2	COL3	ZTAB	COL1	COL2	COL3
ROW1	49.00	96.20	50.50	ROW1	50.01	98.55	51.44
ROW2	97.00	47.60	0.00	ROW2	100.50	49.50	0.00
ROW3	144.00	145.60	95.20	ROW3	149.55	151.77	98.68
ROW4	147.00	49.00	49.00	ROW4	149.94	50.17	49.88

CTAB	COL1	COL2	COL3	ATAB	COL1	COL2	COL3
ROW1	102.06	102.45	101.86	ROW1	1.00	1.00	1.00
ROW2	103.61	103.99	0.00	ROW2	1.00	1.00	1.00
ROW3	103.85	104.24	103.65	ROW3	1.00	1.00	1.00
ROW4	102.00	102.39	101.80	ROW4	1.00	1.00	1.00

References

Bankier, M.D. (1988), "Power Allocation: Determining Sample Sizes for Sub-National Areas", **The American Statistician**, Vol. 42, pp. 174-177.

Bureau of the Census (1986-) **Revised Monthly Retail Sales and Inventories: January 1977 through December 1986**, U.S. Department of Commerce, Cat. No. BR-13-86S

Cholette, P.A. (1984), "Linking Heterogeneous Segments of Seasonal Time Series", **Proceedings of the American Statistical Aswsociation, Business and Economic Statistics**, pp. 178-183

Cholette, P.A. (1988), "Benchmarking Systems of Socio-Economic Time Series", **Statistics Canada**, Time Series Research and Analysis Division, Research Paper No. TSRA-88-017E.

Cholette, P.A., Higginson, J. (1987), "A Note on Aggregating Data Referring to Different Financial Years", **Statistics Canada**, Time Series Research and Analysis Division, Research Paper No. TSRA-87-015E.

Copeland, K.R., Peitzmeier, F.K. and Hoy, C.F. (1987), "An Alternative Method of controlling Current Population Survey Estimates to Population counts", **Survey Methodology**, Vol. 13, No. 2, pp. 173-181.

Dagum, E. Bee (1988), **The X11ARIMA/88 Seasonal Adjustment Method - Foundations and User's Manual**, Statistics Canada, Cat. No. 12-564E

Deming, W.E., Stephan, F.F. (1940), "On a Least Squares Adjustment of a Sampled Frequencuy Table When the Expected Marginal Totals are Known", **Annals of Mathematical Statistics**, Vol. 34, pp. 911-934.

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**, Vol. 66, No. 333, pp. 99-102.

Federal Reserve Board (1962), "Flow of Funds Seasonally Adjusted", **Federal Reserve Bulletin**, Nov. 1962, pp. 1393-1407.

Helfand, S.D., Monsour, J.J., Trager, M.L., (1977), "Historical Revision of Current Business Survey Estimates", **Proceedings of the Business and Economic Statistics Section, American Statistical Association**, pp. 246-250.

Luery, D. (1986). "Weighting Sample Survey Data under Linear Constraints on the weights", **Proceedings of the Section on Social Statistics, American Statistical Association**, pp. 325-350.

Monsour, N.J. Trager, M.L. (1979) "Revision and Benchmarking of Business Time Series", **Proceedings of the Business and Economic Statistics Section, American Statistical Association**, pp. 333-337.

Searle, S.R., Hausman, W.H. (1970), **Matrix Algebra for Business and Economics**, Wiley-Interscience.

SAS Institute Inc. (1985), **SAS/IML ™ User's Guide, Version 6 Edition**, Cary, NC: SAS Institute Inc.

Statistics Canada (1985-), **Wholesale Trade Statistics**, Cat. 63-226, Annual.

Statistics Canada (1987-), **Retail Chain and Department Stores**, Cat. 63-210, Annual.

Statistics Canada (1990a-), **Retail Trade**, January 1990, Cat. 63-005, Monthly.

Statistics Canada (1990b-), **Wholesale Trade**, January 1990, Cat. 63-008, Monthly.

Taylor, S. (1963), "Balancing a Seasonally Adjusted Matrix, Using a Constrained Minimum Sum", **Federal Reserve Board**, Flow of Funds Section, Internal document dated Nov. 13, 1963.