Statistics Canada Statistique Canada

Methodology Branch

Business Survey Methods Division

Direction de la méthodologie

Division des methodes d'enquêtes-entreprises

Canada

WORKING PAPER NO. BSMD-94-008E

CAHIER DE TRAVAIL NO. BSMD-94-008E

METHODOLOGY BRANCH

DIRECTION DE LA MÉTHODOLOGIE

ESTIMATION OF MOVING TRADING-DAY VARIATIONS WITH APPLICATION TO THE CANADIAN RETAIL TRADE SERIES

by

Pierre A. Cholette and Marietta Morry

September 1994

Estimation of Moving Trading-Day Variations
with Application to the Canadian Retail Trade Series

by
Pierre A. CHOLETTE
and
Marietta MORRY

September 7 1994

Statistics Canada
Time Series Research and Analysis Centre
R.H. Coats Building, 3rd floor "H"
Ottawa, Canada, K1A 0T6
(613) 951-1601

Abstract

Many changes in the opening hours and days of stores have occurred in Canada. Because of this and other reasons explained in the paper, the trading-day variations have evolved considerably during the last two decades. However, the methods used to calculate those variations still assume constant relative daily weights. As a result, the monthly seasonally adjusted Retail Trade series have recently displayed residual variations, which are likely attributable to an inadequate estimation of the trading-day component.

The paper presents the solution adopted to solve the problem. The paper also presents an alternative method, based on regression with stochastic parameters.

Key words: Time Series, Seasonal adjustment, Trading-day variations, Regression analysis with stochastic parameters

Résumé

Le Canada a connu de nombreux changements dans les heures et les jours d'ouverture des magasins. À cause de cela et pour d'autres raisons expliquées dans le document, les variations attribuables à la rotation des jours ont considérablement évolué durant les deux dernières décennies. Cependant les méthodes utilisées pour calculer ces variations supposent toujours des poids relatifs des jours constants. Ainsi, les séries mensuelles de Commerce de détail désaisonnalisés ont récemment affiché des variations résiduelles, qui sont vraisemblablement attribuables à une estimation inadéquate de la composante de rotation des jours.

Le document présente la solution adoptée pour résoudre le problème. Il présente également une autre méthode, fondée sur la régression à paramètres stochastiques.

Mots clés: séries chronologiques, désaisonnalisation, variations dues à la rotation des jours, analyse par régression à paramètres stochastique.

1. Introduction

An important source of the month-to-month variations in many monthly economic time series is trading-day variations. In many activities such as production, sales, shipments, trade, the monthly value recorded depends on the number of "trading" days in the month, or more generally speaking on the number of times each of the seven days of the week appears in the month. A classical example is Retail Trade, where more sales are made on Fridays and Saturdays than on other days; as a result a 30-day month containing five such days tends (everything else being equal) to register higher sales than any other 30-day month containing four such days. The phenomenon also affects social and demographical series, such as marriages, immigration - and even births because of the practices of induced labour and caesarean section, which are not uniformally distributed throughout the week (Dumas, 1992).

Thus trading-day variations in monthly values originate from the existence of an underlying daily pattern of activity defined over the week. That daily pattern states the relative importance of the days in the week; just as the seasonal pattern states the relative importance of the months in the year. As Young (1965) pointed out however, the daily pattern may reflect recording and reporting practices, for instance the sales of one day may systematically be recorded the following day.

Like seasonal variations, trading-day variations obscure the fundamental trend-cyclical (as opposed to transient) movements present in the series. For that reason, the *seasonally adjusted series published by statistical agencies exclude both seasonal and trading-day variations.* The trading-day estimation method in the X-11-ARIMA (Dagum 1980) and the Census Method II X-11 (Shiskin, Young and Musgrave, 1967) seasonal adjustment procedures, used by many statistical agencies, is based on Young's method (1965).

In the past, the daily trading pattern was considered stable over several years. This assumption is currently no longer tenable. In the past decade or so, Canada has witnessed dramatic changes in the opening hours of retail outlets: stores were allowed to open on Thursday evenings, then on Wednesday evenings - and more recently on Sundays. The computerization of inventory control, by means of sophisticated cash registers, probably affects reporting practices (presumably making them more timely) and thus the daily pattern. The changing composition of a series also causes evolution: For instance if the computer sales by Furniture and Appliance stores (say) have their own daily pattern and gain relative importance in the Furniture and Appliance series, this gradually alters the daily pattern of the series. For all these reasons, subject matter experts of Retail Trade strongly suspect that some unexpected transient month-to-month fluctuations in the seasonally adjusted series in the recent past are attributable to the inadequate estimation of the trading-day component over the last years. Indeed the X-11-ARIMA seasonal adjustment procedure used in this case assumes constant daily pattern for the whole series.

There are methods in the statistical literature to estimate evolving daily patterns of activity. However, to our knowledge, they have not yet been incorporated in seasonal adjustment procedures used by statistical agencies. Dagum and Quenneville (1988) and Dagum *et al.* (1992) proposed approaches, based on state-space modelling (Kalman, 1962), in which the daily weights change from month to month. They applied such methods to some 40 Canadian Retail Trade series ranging from 1977 to 1986. They found that the change in the daily weights was significant only in the case of Total Retail Sales for the province of Nova Scotia. This paper presents an alternative model, based on stochastic regression analysis, which assumes the daily weights change between trading-day *regimes*, covering a number of years (e.g. 4), instead of between months. The results obtained with this model also suggests no evolution of the daily pattern. The experiment does shed some light as to the nature of the problem: under both methods, the signal-to-noise ratio, which governs the rate of evolution of the daily weights, seems to be largely underestimated. Perhaps, for real series, the presence of outliers causes an inadequate estimation of that ratio.

An experimental version of the X-11-ARIMA seasonal adjustment procedure was developed to allow two independent trading-day regimes, the date of the regime change being carefully selected for each series. In other words, the method of Young, which assumes constant daily pattern, is applied on two separate series intervals, and the other components of the series are estimated over the whole series as before. The main practical advantage of this solution is that the method is embedded in the other calculations of X-11-ARIMA. This is crucial in the case of iterative time series decomposition methods, like X-11 and X-11-ARIMA, where the improved estimation (and re-estimation) of each component leads to improved (re-)estimation of the other components.

Section 2 summarizes the method of Young (1965), as implemented in X-11-ARIMA. Section 3 presents the results obtained with this approach applied on two trading-day regimes. Section 4 presents the aforementioned stochastic regression model. Section 5 compares the performance of the three different methods, on the Canadian Department Store Series: the Young method applied for the whole series, the Young methods applied to two independent trading-day regimes and the stochastic regression method of Section 4. Section 6 concludes with a discussion of this moving trading-day variation estimation exercise.

2. Modification to the X-11-ARIMA Seasonal Adjustment Procedure

The X-11-ARIMA seasonal adjustment procedure was changed to allow two trading-day regimes. The break-point between the two regimes can be supplied by the user, based on subject matter expertise for instance. For example if the series covers the period from January 1981 to December 1993 and the user knows there was a change in shopping patterns starting in October 1991, the first regime would extend from January 1981 to September 1991; and the second, from October 1991 to December 1993. The Young method, already in X-11-ARIMA, is thus applied to the two series intervals separately.

This section describes briefly the trading-day estimation method of Young (1965). The method actually combines two methods: (1) a smoothing method, which directly estimates the monthly trading-day factors from the *trading-day-irregular residuals* (this is the series obtained by removing the seasonal and the trend-cycle estimates from the original series), and (2) a regression method, which estimates the daily pattern from these residuals and then derives the corresponding monthly trading-day factors. First the notion of monthly trading-day factors is formalized.

2.1 The Monthly Trading-Day Factors

Under the multiplicative seasonal adjustment model, $z_t = c_t \times s_t \times M_t \times i_t$, where z_t is the observed monthly series, c_t the trend-cycle component, s_t the seasonal component, i_t the irregular, the monthly trading-day factor M_t relates to the daily weights as follows:

$$M_t = [\sum_{j=1}^{7} n_{t,j}(1+\beta_j)]/N_t^* \equiv [28 + \sum_{j:\,5\,times} (1+\beta_j)]/N_t^*, \tag{2.1}$$

where N_t is the actual number of days in each month, $N_t^* = N_t$ and $N_t^* = 28.25$ for Februaries, $n_{t,j}$ is the number of days j in month t (4 or 5) and β_j, j=1,...,7 are the daily weights (from Monday to Sunday) summing to 0. A weight $\beta_6 = 0.50$ means that 50% more activity takes place on Saturday than on an average day of the week; a weight $\beta_1 = -0.50$, that 50% less activity takes place on Monday than on an average day of the week.

Specification (2.1) assigns the length-of-month allowance to the seasonal factors. The fact that March (say) has 31 days is captured by a higher March seasonal factor; and similarly the fact that April has 30 days, by a lower seasonal factor. If the length-of-month allowance is to be captured by the trading-day factors, then N_t^* is set to the average number of days in a month: 30.4375 = 365.25/12.

2.2 The Smoothing Method

Young (1965) first observes that any series has at most 22 *types of months*, depending on the length of the month and on the first day of the month; and, that each type of month has the same distribution of days. For instance all 31-day months starting with a Monday have the same distribution of Mondays, Sundays, etc.; all 31-day months starting with a Tuesday; etc. There are thus 7 possible distributions for 31-day months, 7 for 30-day months, 7 for 29-day months and 1 for 28-day months, for a total of 22 types of months. Furthermore, if the daily pattern is the same for all the months of a series (interval) considered, there are (according to (2.1)) only 22 possible values for the monthly trading-day factors. The smoothing approach simply consists of setting each of the 22 possible values of M_t equal to the *group average* of the trading-day-irregular residuals $I_t = M_t \times i_t$ pertaining to the same type of month.

The problem with this smoothing technique, is that some of the 22 types of months occur very seldom (which would translate into unreliable estimates) or not at all in a given series. Indeed 28 years of monthly data are needed to ensure that all types of month occur at least once. The smoothing method thus cannot be used for the estimation of the trading-day factors. It is used, however, to locate outlier trading-day-irregular residuals (for the months occurring more than once at least) which will be excluded from the regression used to estimate the daily weights. Basically (the procedure is more intricate than this), a residual I_t is ruled outlier if it departs from its group average by a number (usually 2) of standard deviations; the latter is the square root of the sum of squared deviations of the residuals from their group average, divided by the number of observations, T.

2.3 The Regression Method

The regression method consists of estimating the 7 daily weights by regressing the trading-day-irregular residuals I_t without outliers on the number of Mondays, Tuesdays, etc., occurring in each month; and, of calculating all the monthly trading-day factors M_t from the estimated daily weights by means of (2.1). Cholette and Quenneville (1994) show that, for the multiplicative seasonal adjustment model, the appropriate specification is

$$y_t = \sum_{j=1}^{6} (z_{t,j} \bar{N} / N_t^{*}) \beta_j + e_t, \; t=1,...,T, \tag{2.2}$$

where $y_t = (I_t - N_t / N_t^{*}) \bar{N}$ is the regressand and $(z_{t,j} \bar{N} / N_t^{*}), j=1,...,6$ are the regressors with $z_{t,j} = (n_{t,j} - n_{t,7}), j=1,...,6$ being the number of times day j appears in month t, minus the number of Sundays in the same month. $\bar{N}$ is the average of N_t over time.

3. Results Obtained with the Modified X-11-ARIMA Procedure

The experimental version of X-11-ARIMA described in section 2 was applied to the officially published system of Canadian Retail Trade Series which consists of 18 Trade Group Totals and 11 Provincial Totals and the grand Canada Total. Each series was processed with several break-points defining the trading-day regimes (the break-point being the starting date of the second regime). Because, the biggest shopping pattern changes occurred recently, January 1991 was chosen as the earliest break-point. Since large month-to-month fluctuations in the seasonally adjusted series can be indicative of having removed the wrong trading-day component from the data, statistics on month-to-month percentage change in the seasonally adjusted series were calculated. The break-point yielding the lowest values for the statistics, i.e. the smoothest series, would be optimal. However, subject matter knowledge was also taken into consideration in the final selection of the break-points.

Table 1 displays the statistics on change in the seasonally adjusted Department Stores series, for the last six years ("88-93" in table), for years 1988 to 1990 ("88-90") and for the last three years ("91-93"); for each break-point considered. The statistics for the last three years are more important, because of the role of these years for current economic analysis. The first three rows of the table display the statistics obtained under only one trading-day regime ("No Break"), i.e. under the status quo; these statistics are repeated for each break-point (in columns 4, 6 and 8 under heading "No Break"). The other sets of three rows (rows 4, 5, 6; 7, 8 9, etc.) display the statistics obtained under the various break-points, ("9101" standing for January 1991, etc.). The most relevant statistic used in determining the optimal break was

the Mean Absolute Percent month-to-month Change, especially for the last three years. In the case of Department Stores, the minimum occurs around June and July 1992. Table 2 presents the same statistics calculated for each break-point and each trading-day regime. The minimum also occurs around June and July 1992. The July date was ruled preferable on the basis of subject matter knowledge and the lower Number of Changes of Direction of the series (in Table 1). Table 2 also displays the daily weights estimated on each regime. The steady progression of the Sunday weight and decline of the Monday weight of the second regime, as the break-point becomes more recent, is quite remarkable: Sunday triples in importance compared to the status quo value (in first row), while Monday shrinks by half. Also noteworthy is the systematically widening gap between the same-day weights on the first and second regimes. These observations confirm subject matter expectation.

Table 1 Analysis of the Changes in the Seasonally Adjusted Department Store Series, by Break-Point and Year

Break Point	Years	Mean Abs. % Change	Mean Abs. % Change No Break	S.D. of % Change	S.D. of % Change No Break	Max. Abs. % Change	Max. Abs. % Change No Break	No. Changes of Direction
0	88-93	1.39	1.39	2.06	2.06	7.22	7.22	47
0	88-90	1.41	1.41	1.93	1.93	6.51	6.51	25
0	91-93	1.36	1.36	2.16	2.16	7.22	7.22	22
9101	88-93	1.32	1.39	1.82	2.06	6.82	7.22	47
9101	88-90	1.43	1.41	1.95	1.93	6.82	6.51	25
9101	91-93	1.20	1.36	1.66	2.16	4.72	7.22	22
9104	88-93	1.33	1.39	1.82	2.06	6.77	7.22	49
9104	88-90	1.42	1.41	1.93	1.93	6.77	6.51	25
9104	91-93	1.23	1.36	1.68	2.16	4.84	7.22	24
9107	88-93	1.34	1.39	1.90	2.06	6.74	7.22	49
9107	88-90	1.40	1.41	1.91	1.93	6.74	6.51	25
9107	91-93	1.27	1.36	1.86	2.16	5.22	7.22	24
9110	88-93	1.32	1.39	1.91	2.06	6.25	7.22	47
9110	88-90	1.37	1.41	1.87	1.93	6.25	6.51	23
9110	91-93	1.27	1.36	1.95	2.16	5.53	7.22	24
9201	88-93	1.33	1.39	1.90	2.06	6.63	7.22	47
9201	88-90	1.42	1.41	1.92	1.93	6.63	6.51	25
9201	91-93	1.24	1.36	1.87	2.16	4.86	7.22	22
9204	88-93	1.36	1.39	1.94	2.06	6.67	7.22	44
9204	88-90	1.45	1.41	1.96	1.93	6.67	6.51	23
9204	91-93	1.27	1.36	1.91	2.16	5.60	7.22	21
9205	88-93	1.38	1.39	1.97	2.06	6.62	7.22	44
9205	88-90	1.43	1.41	1.93	1.93	6.62	6.51	23
9205	91-93	1.32	1.36	2.00	2.16	5.52	7.22	21
9206	88-93	1.33	1.39	1.93	2.06	6.55	7.22	46
9206	88-90	1.43	1.41	1.94	1.93	6.55	6.51	25
9206	91-93	1.23	1.36	1.90	2.16	5.29	7.22	21
9207	88-93	1.33	1.39	1.93	2.06	6.49	7.22	44
9207	88-90	1.43	1.41	1.96	1.93	6.49	6.51	23
9207	91-93	1.24	1.36	1.89	2.16	5.29	7.22	21
9208	88-93	1.34	1.39	1.94	2.06	6.50	7.22	44
9208	88-90	1.42	1.41	1.95	1.93	6.50	6.51	23
9208	91-93	1.25	1.36	1.91	2.16	5.36	7.22	21
9209	88-93	1.34	1.39	1.94	2.06	6.57	7.22	48
9209	88-90	1.44	1.41	1.97	1.93	6.57	6.51	27
9209	91-93	1.24	1.36	1.90	2.16	5.30	7.22	21
9210	88-93	1.33	1.39	1.93	2.06	6.56	7.22	44
9210	88-90	1.43	1.41	1.97	1.93	6.56	6.51	23
9210	91-93	1.23	1.36	1.89	2.16	5.24	7.22	21
9211	88-93	1.34	1.39	1.94	2.06	6.58	7.22	44
9211	88-90	1.40	1.41	1.94	1.93	6.58	6.51	23
9211	91-93	1.28	1.36	1.92	2.16	5.42	7.22	21
9212	88-93	1.34	1.39	1.94	2.06	6.58	7.22	44
9212	88-90	1.40	1.41	1.94	1.93	6.58	6.51	23
9212	91-93	1.27	1.36	1.91	2.16	5.40	7.22	21

Table 2 Analysis of the Changes in the Seasonally Adjusted Department Store Series, by Break-Point and by Regime for the Last Six Years of the Series

Break Point	Regime	Mean Abs. % Change	Mean Abs. % Change No Break	S.D. % Change	S.D. % Change No Break	Max. Abs. % Change	Max. Abs. % Change No Break	Estimated Daily Weights Mon.	Tues.	Wed.	Thur.	Fri.	Sat.	Sun.
none	1st	1.39	1.39	2.06	2.06	7.22	7.22	0.997	0.742	1.014	1.337	1.083	1.515	0.312
9101	Both	1.32	1.39	1.82	2.06	6.82	7.22							
9101	1st	1.43	1.41	1.95	1.93	6.82	6.51	1.092	0.802	0.934	1.361	1.091	1.505	0.215
9101	2nd	1.20	1.36	1.66	2.16	4.72	7.22	0.729	0.686	1.169	1.258	0.886	1.775	0.496
9104	Both	1.33	1.39	1.82	2.06	6.77	7.22							
9104	1st	1.40	1.41	1.91	1.91	6.77	6.51	1.083	0.790	0.945	1.365	1.087	1.503	0.227
9104	2nd	1.24	1.37	1.70	2.18	4.84	7.22	0.752	0.625	1.228	1.236	0.887	1.834	0.437
9107	Both	1.34	1.39	1.90	2.06	6.74	7.22							
9107	1st	1.56	1.52	2.17	2.07	6.74	6.51	1.073	0.787	0.953	1.355	1.098	1.495	0.240
9107	2nd	1.05	1.22	1.46	2.02	4.17	7.22	0.690	0.690	1.159	1.271	0.768	1.933	0.488
9110	Both	1.32	1.39	1.91	2.06	6.25	7.22							
9110	1st	1.54	1.49	2.14	2.05	6.25	6.51	1.026	0.828	0.937	1.408	1.052	1.531	0.219
9110	2nd	0.98	1.24	1.48	2.06	4.21	7.22	0.704	0.611	1.147	1.314	0.753	1.843	0.629
9201	Both	1.33	1.39	1.90	2.06	6.63	7.22							
9201	1st	1.49	1.45	2.04	1.99	6.63	6.51	1.046	0.781	0.961	1.359	1.099	1.517	0.236
9201	2nd	1.03	1.28	1.59	2.16	4.73	7.22	0.651	0.625	1.149	1.468	0.674	1.781	0.653
9204	Both	1.36	1.39	1.94	2.06	6.67	7.22							
9204	1st	1.66	1.67	2.27	2.37	6.67	7.22	1.052	0.790	0.934	1.375	1.087	1.551	0.211
9204	2nd	0.68	0.76	0.81	0.99	1.71	2.29	0.667	0.620	1.115	1.535	0.830	1.354	0.878
9205	Both	1.38	1.39	1.97	2.06	6.62	7.22							
9205	1st	1.69	1.67	2.28	2.36	6.62	7.22	1.049	0.778	0.946	1.388	1.073	1.552	0.214
9205	2nd	0.62	0.69	0.72	0.89	1.27	2.29	0.674	0.619	1.121	1.467	0.849	1.418	0.852
9206	Both	1.33	1.39	1.93	2.06	6.55	7.22							
9206	1st	1.63	1.66	2.22	2.35	6.55	7.22	1.057	0.778	0.943	1.394	1.067	1.554	0.207
9206	2nd	0.54	0.67	0.65	0.89	1.33	2.29	0.578	0.658	1.101	1.380	0.950	1.415	0.918
9207	Both	1.33	1.39	1.93	2.06	6.49	7.22							
9207	1st	1.60	1.63	2.19	2.32	6.49	7.22	1.041	0.767	0.956	1.393	1.062	1.557	0.224
9207	2nd	0.59	0.70	0.79	0.91	2.16	2.29	0.554	0.653	1.154	1.336	0.966	1.415	0.922
9208	Both	1.34	1.39	1.94	2.06	6.50	7.22							
9208	1st	1.61	1.65	2.19	2.32	6.50	7.22	1.041	0.760	0.962	1.393	1.069	1.549	0.226
9208	2nd	0.49	0.60	0.58	0.73	1.34	1.77	0.494	0.739	1.074	1.386	0.856	1.480	0.971
9209	Both	1.34	1.39	1.94	2.06	6.57	7.22							
9209	1st	1.59	1.62	2.17	2.30	6.57	7.22	1.052	0.749	0.963	1.396	1.056	1.560	0.224
9209	2nd	0.50	0.62	0.63	0.75	1.32	1.77	0.447	0.787	1.067	1.384	0.951	1.383	0.981
9210	Both	1.33	1.39	1.93	2.06	6.56	7.22							
9210	1st	1.57	1.61	2.16	2.28	6.56	7.22	1.051	0.752	0.962	1.396	1.056	1.560	0.223
9210	2nd	0.47	0.60	0.59	0.76	1.26	1.77	0.420	0.821	1.085	1.341	0.978	1.378	0.978
9211	Both	1.34	1.39	1.94	2.06	6.58	7.22							
9211	1st	1.55	1.60	2.15	2.27	6.58	7.22	1.048	0.746	0.998	1.361	1.063	1.532	0.252
9211	2nd	0.50	0.57	0.59	0.75	1.15	1.77	0.454	0.796	1.070	1.400	0.935	1.397	0.949
9212	Both	1.34	1.39	1.94	2.06	6.58	7.22							
9212	1st	1.54	1.57	2.13	2.25	6.58	7.22	1.048	0.746	0.998	1.362	1.062	1.531	0.253
9212	2nd	0.47	0.60	0.58	0.77	1.13	1.77	0.474	0.742	1.096	1.399	0.946	1.348	0.996

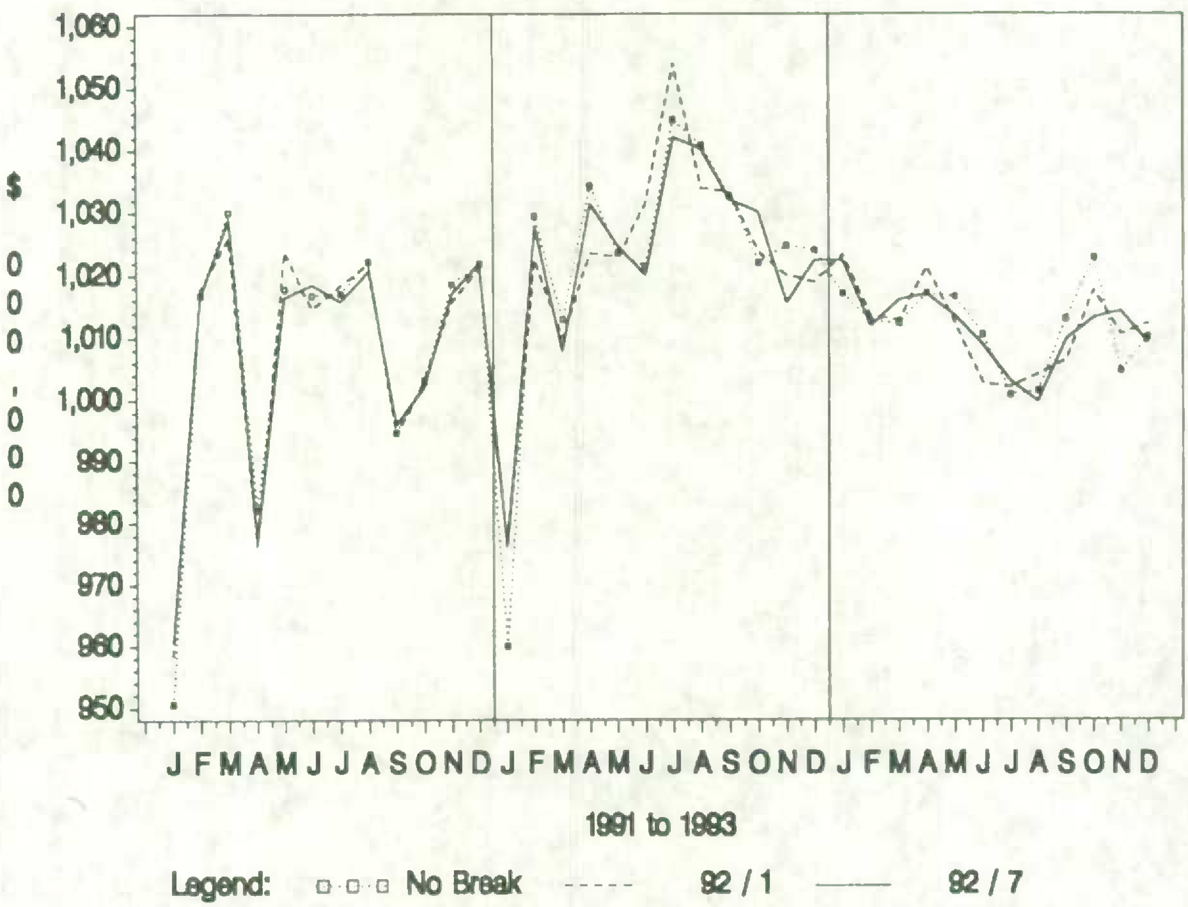


Figure 1 Three seasonally adjusted Department Stores series obtained under three break-points

Figure 1 displays three seasonally adjusted Department Stores series obtained under three break-points: the series with no break point represents the status quo (squares in figure), the two other series with break-points in January 1992 (dashed curve) and in July 1992 (solid) are alternative to the status quo. The July 1992 series is clearly smoother than the other two (especially from July 1992 onwards) and does not contain some of the transient fluctuations present in the status quo series, which were deemed questionable.

Table 3 displays the trading-day regressions, produced the by experimental X-11-ARIMA programme, on the two regimes corresponding to the July 1992 break-point. The other components of the series, namely the trend-cycle and the seasonal are calculated over the whole series.

Table 3 Trading-Day Regressions on the two Trading-Day Regimes Selected for Department Store, Calculated by the Experimental Version of the X-11-ARIMA Programme

Series 131001 DEPARTMENT STORES -- 131 page 5

A07. TRADING-DAY REGRESSION FROM FIRST PASS

REGRESSION on Regime Ranging from 81 1 to 92 6

	Combined Weight	Prior Weight	Regress. Coeff.	Std. Err. Comb. Wt.	t 1.0	t Prior wt.
Monday	1.041	1.000	.041	.071	.572	.572
Tuesday	.767	1.000	-.233	.070	-3.339**	-3.339**
Wednesday	.956	1.000	-.044	.066	-.662	-.662
Thursday	1.393	1.000	.393	.069	5.724***	5.724***
Friday	1.062	1.000	.062	.071	.877	.877
Saturday	1.557	1.000	.557	.068	8.189***	8.189***
Sunday	.224	1.000	-.776	.068	-11.443***	-11.443***

The stars indicate the combined wt. is significantly different from 1 or from the prior weight.
The significance levels are 3 stars (0.1 %), 2 stars (1 %), 1 star (5 %).
No stars indicates not significant at the 5 % level

SOURCE OF VARIANCE:	Sum of Squares	Dgrs.of Freedom	Mean Square	F-Value
Regression	.3464288E+02	6	.5773813E+01	.7407260E+02***
Error	.9509659E+01	122	.7794803E-01	
Total	.4415254E+02	128		

*** Trading-day variation present at the 1 percent level

Standard errors of trading-day adjustment factors derived from regression coefficients

31-day months-	.20
30-day months-	.21
29-day months-	.24
28-day months-	.00

Series 131001 DEPARTMENT STORES PRINT 0 -- 131 page 6

A07. TRADING-DAY REGRESSION FROM FIRST PASS

REGRESSION on Regime Ranging from 92 7 to 93 12

	Combined Weight	Prior Weight	Regress. Coeff.	Std. Err. Comb. Wt.	t 1.0	t Prior wt.
Monday	.554	1.000	-.446	.152	-2.930*	-2.930*
Tuesday	.653	1.000	-.347	.135	-2.566*	-2.566*
Wednesday	1.154	1.000	.154	.137	1.124	1.124
Thursday	1.336	1.000	.336	.153	2.193*	2.193*
Friday	.966	1.000	-.034	.157	-.218	-.218
Saturday	1.415	1.000	.415	.147	2.825*	2.825*
Sunday	.922	1.000	-.078	.167	-.470	-.470

The stars indicate the combined wt. is significantly different from 1 or from the prior weight.
The significance levels are 3 stars (0.1 %), 2 stars (1 %), 1 star (5 %).
No stars indicates not significant at the 5 % level

SOURCE OF VARIANCE:	Sum of Squares	Dgrs.of Freedom	Mean Square	F-Value
Regression	.4083677E+01	6	.6806129E+00	.1508628E+02***
Error	.5413761E+00	12	.4511468E-01	
Total	.4625053E+01	18		

*** Trading-day variation present at the 1 percent level

Standard errors of trading-day adjustment factors derived from regression coefficients

31-day months-	.47
30-day months-	.46
29-day months-	.52
28-day months-	.00

4. Regression Model to Estimate Moving Daily Patterns

This section presents a regression model, estimated by stochastic least squares, in which the daily coefficients change between trading-day regimes. In the models by Dagum *et al.* (1992) the daily weights changed every month; in the model presented here the weights change every regime, and the length of a regime can be specified by the subject matter expert. The model also enables the explicit incorporation of prior information regarding the level of some of the daily weights (e.g. that of Sunday) on some of the regimes and regarding change in the daily weights from one regime to the next. The model could be specified in the state-space framework; in principle the smoothed state-space values would coincide with the results obtained by the regression (e.g. Duncan and Horn, 1972), except for initial conditions which the regression does not have. Since the size of the matrices involved is not large, we use the more familiar regression framework, at this experimental stage at least.

The regression model contains three types of equations: the observation equations, the transition equations (which are notions associated with state-space modelling) and the prior information equations.

4.1 The Observation Equations

The observation equations relate the daily coefficients to the observed data y_t (like in section 2) on each regime r:

$$\underset{T_r\times 1}{\boldsymbol{y}_r} = \underset{T_r\times 6}{\boldsymbol{Z}_r}\ \underset{6\times 1}{\boldsymbol{\beta}_r} + \underset{T_r\times 1}{\boldsymbol{e}_r}, \quad E(e_r)=0,\ E(e_r e_r') = \sigma_e^2 I_{T_r}, \quad r=1,\dots,R \tag{4.1.1}$$

where $\boldsymbol{y}_r = [y_{r,1}\, y_{r,2} \cdots y_{r,T_r}]'$ are the observations on regime r, $\boldsymbol{\beta}_r = [\beta_{r,1}\, \beta_{r,2} \cdots \beta_{r,6}]'$ are the daily weights on trading regime r, R is the number of regimes in the series of $T = T_1 + T_2 + \dots + T_R$ observations and Z_r contains the regressors of equation (2.2).

The observation equations are more written compactly as:

$$\underset{T\times 1}{\boldsymbol{y}} = \underset{T\times 6R}{\boldsymbol{Z}}\ \underset{6R\times 1}{\boldsymbol{\beta}} + \underset{T\times 1}{\boldsymbol{e}}, \tag{4.1.2}$$

where $\boldsymbol{y} = [y'_1\, y'_2 \cdots y'_R]'$, $\boldsymbol{\beta} = [\boldsymbol{\beta}'_1\, \boldsymbol{\beta}'_2 \cdots \boldsymbol{\beta}'_R]'$, $\boldsymbol{e} = [\boldsymbol{e}'_1\, \boldsymbol{e}'_2 \cdots \boldsymbol{e}'_R]'$ and $\mathbf{Z} = block(\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_R)$.

4.2 The Transition Equations

The transition equations specify the manner in which the daily weights evolve from regime to regime. A specification often encountered is the first difference (random walk) model, which in this case states that β_{r+1} is basically equal to β_r plus an innovation ϵ_{r+1}:

$$\underset{6\times 1}{c_r} = \underset{6\times 12}{[-I_6\ \ I_6]} \begin{bmatrix} \boldsymbol{\beta}_r \\ \boldsymbol{\beta}_{r+1} \end{bmatrix} + \epsilon_{r+1}, \quad \begin{matrix} E(\epsilon_{r+1}) = \mathbf{0}, \\ E(\epsilon_{r+1}\, \epsilon'_{r+1}) = \sigma_\epsilon^2\, \Omega^*_{r+1} \end{matrix}, \quad r=1,\dots,R-1 \tag{4.2.1}$$

where c_r is the expected change in same-day weights between regimes; $c_r = 0$ specifies gradual transition between regimes. Matrix Ω_r^* specifies that the innovations ϵ_{r+1} are such that β_{r+1} sums to zero (Harvey, 1989, eq. (2.3.54)):

$$\underset{6\times 6}{\Omega_r^*} = \Omega^* = \left[\omega_{ij} = \begin{cases} 6/7, & i=j \\ -1/7, & i\neq j \end{cases} \right]. \tag{4.2.2}$$

Another specification often encountered is the second difference model, which states that the daily weights evolve locally as a straight line.

The transition equations are written more compactly as

$$\underset{6(R-d)\times 1}{c} = \underset{6(R-d)\times 6R}{D} \; \underset{6R\times 1}{\beta} + \epsilon, \quad E(\epsilon) = \mathbf{0}, \quad E(\epsilon\epsilon') = \sigma_\epsilon^2 \Omega, \tag{4.2.3}$$

where d is the degree of differencing (1 or 2) and where

$$c = [c'_1 \dots c'_{R-d}]', \quad \Omega = I_{R-d} \otimes \Omega^* = block(\Omega_1^*, \dots, \Omega_{R-d}^*),$$

$$\underset{6(R-1)\times 6R}{D} = \begin{bmatrix} -I_6 & I_6 & \mathbf{0}_6 & \mathbf{0}_6 & \cdots \\ \mathbf{0}_6 & -I_6 & I_6 & \mathbf{0}_6 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \; d=1, \quad \underset{6(R-2)\times 6R}{D} = \begin{bmatrix} -I_6 & 2I_6 & -I_6 & \mathbf{0}_6 & \cdots \\ \mathbf{0}_6 & -I_6 & 2I_6 & I_6 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \; d=2.$$

4.3 The Prior Information Equations

The prior information on the level of the daily weights is incorporated (Theil, 1971, Ch. 7.8) by means of:

$$\underset{N\times 1}{\beta^*} = \underset{N\times 6R}{P} \; \underset{6R\times 1}{\beta} + \underset{N\times 1}{\alpha}, \quad E(\alpha) = \mathbf{0}, \quad E(\alpha\alpha') = V_\alpha, \tag{4.3.1}$$

where $\beta^* = [\beta^*_{r_1 d_1} \cdots \beta^*_{r_n d_n} \cdots \beta^*_{r_N d_N}]'$ is prior knowledge about some day(s) on some regime(s), $N < 6R$ is the number of priors, $1 \le r_n \le R$, $1 \le d_n \le 7$, V_α reflects confidence in that prior knowledge and P is a design matrix containing 0s, 1s and -1s (for Sundays):

$$\underset{N\times 6R}{P} = \left[\underset{n=1,\dots,N;\; j=1,\dots,6R}{p_{nj}} = \begin{cases} 1, & j-[j/6]6 = d_n, \;\; d_n \neq 7 \\ -1, & [j/6]+1 = r_n, \;\; d_n = 7 \\ 0, & [j/6]+1 \neq r_n \end{cases} \right],$$

where $[x]$ stands for the integer part of x. For example, assuming normality, 4 regimes, a prior weight for Saturday on regime 4 equal to 1.15 ± 0.10 with 95% confidence and a prior

weight for Sunday on regime 4 equal to 1.20 ± 0.05 with 95% confidence, then N=2, R=4, r_1=4, d_1=6, r_2=4, d_2=7, $\boldsymbol{\beta}^* = [0.15\ 0.20]'$,

$$\boldsymbol{V}_{\alpha} = \begin{bmatrix} (0.10/2)^2 & -(0.10/2)(0.05/2)/7 \\ -(0.10/2)(0.05/2)/7 & (0.05/2)^2 \end{bmatrix},$$

$$\underset{2\times 24}{\boldsymbol{P}} = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & \dots & 0 & -1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix}.$$

(The negative covariance in $\boldsymbol{V}_{\alpha}$ is a consequence of (4.2.2) above, under the assumption there is no prior known trade-off between Saturday and Sunday.)

4.4 Overall Model

The observation, transition and prior equations combine into one single equation:

$$\begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{c} \\ \boldsymbol{\beta}^* \end{bmatrix} = \begin{bmatrix} \boldsymbol{Z} \\ \boldsymbol{D} \\ \boldsymbol{P} \end{bmatrix} \boldsymbol{\beta} + \begin{bmatrix} \boldsymbol{e} \\ \boldsymbol{\epsilon} \\ \boldsymbol{\alpha} \end{bmatrix}, \tag{4.4.1}$$

or

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}, \quad E(\boldsymbol{u}) = \boldsymbol{0}, \quad E(\boldsymbol{u}\boldsymbol{u}') = \boldsymbol{V} = \begin{bmatrix} \sigma_e^2 \boldsymbol{I}_T & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_\epsilon^2 \boldsymbol{\Omega} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{V}_\alpha \end{bmatrix}, \tag{4.4.2}$$

where the $6R$+2 parameters to be estimated are $\boldsymbol{\beta}$, σ_e^2 and σ_ϵ^2 and where $\boldsymbol{V}_\alpha$ is known. (One could argue that, if d = 1, there are 8 parameters $\boldsymbol{\beta}_1$, σ_e^2 and σ_ϵ^2, because $\boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_R$, are determined by $\boldsymbol{\beta}_1$, which acts as an initial condition, and by $q = \sigma_\epsilon^2 / \sigma_e^2$; e.g. Newbold and Bos, 1985.) The signal-to-noise ratio: $q = \sigma_\epsilon^2 / \sigma_e^2$ governs how much the daily weights evolve under model (4.2.3). If q is close to 0, the daily weights are close to being constant across regimes when d = 1; and, to being linear when d = 2.

The model is designed to be used as follows:

(1) When there is no discontinuity in the underlying evolution (presumably the general case),
 (a) the division of the series in regimes is designed to approximate the underlying evolution of the daily rates of activity,
 (b) the choice of the regimes is more determined by feasibility considerations, e.g. number of degrees of freedom on one regime, collinearity of the regressors on segments shorter than 4 years (Salinas, 1984),
 (c) when the daily weights are known to change more rapidly on one interval of the series, the regimes can be made shorter in that interval.

(2) When a discontinuity happens in the underlying evolution,
 (a) it is important to start a new regime as close as possible to the point of discontinuity,
 (b) the size of discontinuity should be supplied in the expected change, $\boldsymbol{c}$, for the days involved, or the transition equation should be omitted for those days,

(c) prior information about the new level of those days would be useful, especially if the transition equation were omitted.

4.5 Maximum Likelihood Estimation of the Model

The maximum likelihood estimate of $\boldsymbol{\beta}$ conditional on σ_e^2, σ_ϵ^2 (and known V_α) is the generalized least square estimate:

$$\hat{\boldsymbol{\beta}} = (X' V^{-1} X)^{-1} X' V^{-1} Y, \tag{4.5.1}$$

$$V_{\boldsymbol{\beta}} = (X' V^{-1} X)^{-1} \tag{4.5.2}$$

where $(X' V^{-1} X)^{-1} = (\sigma_e^{-2} Z'Z + \sigma_\epsilon^{-2} D' \Omega^{-1} D + P' V_\alpha^{-1} P)^{-1}$ is a $6R \times 6R$ matrix and $X' V^{-1} Y = (\sigma_e^{-2} Z'y + \sigma_\epsilon^{-2} D' \Omega^{-1} c + P' V_\alpha^{-1} \boldsymbol{\beta}^*)$ is $6R \times 1$ vector. The maximum likelihood estimate of σ_e^2 and σ_ϵ^2 conditional on $\boldsymbol{\beta}$ are:

$$\hat{\sigma}_e^2 = [(y - Z\boldsymbol{\beta})'(y - Z\boldsymbol{\beta})] / T, \quad \hat{\sigma}_\epsilon^2 = [(c - D\boldsymbol{\beta})' \Omega^{-1} (c - D\boldsymbol{\beta})] / 6(R-d). \tag{4.5.3}$$

Note that (what we will refer to as) the *joint stochastic estimates* $\hat{\boldsymbol{\beta}}$ can be seen as a fit through the *independent estimates* $\tilde{\boldsymbol{\beta}}$, obtained by ordinary least squares on each regime separately:

$$\tilde{\boldsymbol{\beta}} = V_{\tilde{\beta}} Z'y, \quad V_{\tilde{\beta}} = \sigma_e^2 (Z'Z)^{-1}, \tag{4.5.4}$$

$$\hat{\boldsymbol{\beta}} = \tilde{\boldsymbol{\beta}} + V_{\tilde{\beta}} D' [D V_{\tilde{\beta}} D' + \sigma_\epsilon^2 \Omega]^{-1} [c - D\tilde{\boldsymbol{\beta}}] \quad = W\tilde{\boldsymbol{\beta}}, \tag{4.5.5}$$

$$V_{\hat{\beta}} = V_{\tilde{\beta}} - V_{\tilde{\beta}} D' [D V_{\tilde{\beta}} D' + \sigma_\epsilon^2 \Omega]^{-1} D V_{\tilde{\beta}}, \tag{4.5.6}$$

where (4.5.5) shows $\hat{\boldsymbol{\beta}}$ as a linear combination of $\tilde{\boldsymbol{\beta}}$. This situation also occurs in the repeated survey literature, where the "elementary estimates" (Gurney and Daly, 1965; Jones, 1980) correspond to the independent estimates $\tilde{\boldsymbol{\beta}}$ herein. Also note that the transition equations become constraints $[c - D\tilde{\boldsymbol{\beta}}]$ as σ_ϵ^2 (hence $q = \sigma_\epsilon^2 / \sigma_e^2$) tends to 0 (e.g. Theil, 1971, p. 316).

4.6 Procedure to Estimate the Moving Daily Weights

Equations (4.5.1) to (4.5.3) suggest the following iterative estimation procedure:

(1) Calculate the vectors and matrices which will not change during the procedure: $Z'y$, $Z'Z$, $D'\Omega^{-1}c$, $D'\Omega^{-1}D$, $P'V_\alpha^{-1}\boldsymbol{\beta}^*$, $P'V_\alpha^{-1}P$

(2) Set the initial values of $\boldsymbol{\beta}$ to the independent estimates $\tilde{\boldsymbol{\beta}}$ of (4.5.4).

(3) Calculate the initial values of σ_e^2 and σ_ϵ^2 by substituting $\boldsymbol{\beta} = \tilde{\boldsymbol{\beta}}$ in (4.5.3) .

(4) Calculate joint estimates of $\boldsymbol{\beta}$ by substituting $\sigma_e^2 = \hat{\sigma}_e^2$, $\sigma_\epsilon^2 = \hat{\sigma}_\epsilon^2$ in (4.5.1):

$$V_{\beta} = (\hat{\sigma}_e^{-2}\, Z'Z + \hat{\sigma}_\epsilon^{-2}\, D'\Omega^{-1} D + P' V_\alpha^{-1} P)^{-1} \tag{4.6.1a}$$

$$\hat{\beta} = V_{\beta}\, (\hat{\sigma}_e^{-2}\, Z'y + \hat{\sigma}_\epsilon^{-2}\, D'\Omega^{-1} c + P' V_\alpha^{-1}\, \beta^*). \tag{4.6.1b}$$

(5) Calculate improved estimates for σ_e^2 and σ_ϵ^2 on the basis of the joint estimates, by substituting $\beta = \hat{\beta}$ in (4.5.3)

(6) Repeat steps (4) and (5) until convergence of the estimated parameters.

(7) Generate the sets of weights $\hat{\beta}_r^{\dagger}$ which include Sunday and their covariance matrix by means of:

$$\hat{\beta}_r^{\dagger} = H\,\hat{\beta}_r, \quad V_{\beta_r^{\dagger}} = H\, V_{\beta_r} H', \; r=1,\dots,R \tag{4.6.2}$$

where $\underset{7\times 6}{H} = \begin{bmatrix} I_6 \\ [-1 \dots -1] \end{bmatrix}$.

The estimates of the variances in step (3) are consistent, in the sense that as the number of observations is increased on *each* regime the estimates converge to the true values; and the first estimates of β in step (4) are asymptotically efficient Harvey (1981, p. 192). As a whole, such a procedure leads to the unconditional maximum likelihood estimates (Ibid. p. 136).

5. An Application of the Regression Model

This section applies the regression model to the Department Store series used in section 3. The trading-day-irregular residuals of the series originate from the first table B13 of the X-11-ARIMA seasonal adjustment programme, in which the identification of the extremes described in section 2.2 was suppressed. The calculation described in this section are therefore not embedded in the X-11-ARIMA programme, and the results are not quite comparable with those of section 3.

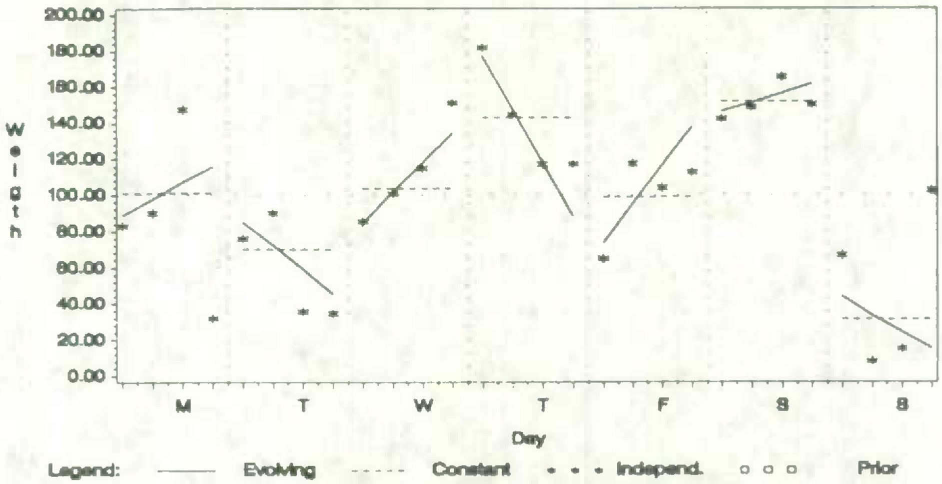
The series of trading-day-irregular residuals was divided in four trading-day regimes, ranging from 1981 to 1984, from 1985 to 1988, from January 1989 to June 1992, and from July 1992 to November 1993, respectively. The daily weights were estimated by three methods. *Method 1* is the that of Young (1965) applied to all regimes jointly (status quo), providing *constant estimates*. *Method 2* is also that of Young but applied to each regime separately, providing the *independent estimates* of (4.5.4). *Method 3* is the regression model presented in section 4, providing the *joint stochastic estimates* of (4.6.1).

Figures 2 and 3 display the estimated daily coefficients by day and regime: all the Mondays on regimes 1 to 4; followed by the Tuesdays on regimes 1 to 4, etc. All panels of Figures 2 and 3 display the constant estimates from method 1 (dotted curves) and the independent estimates (stars) from method 2. The independent estimates suggest evolving daily weights, even in the first regimes. In the case of Wednesday for instance, they almost lie on an upwards-sloping straight line, which confirms subject matter expectations; in the case of Thursday, they lie on a downward-sloping line. For Sunday and Monday, they suggest a discontinuity in behaviour between the third and the fourth regime and a transfer of activity from Monday to Sunday. The constant estimates seem inadequate, at least when seen as a fit through the independent estimates.

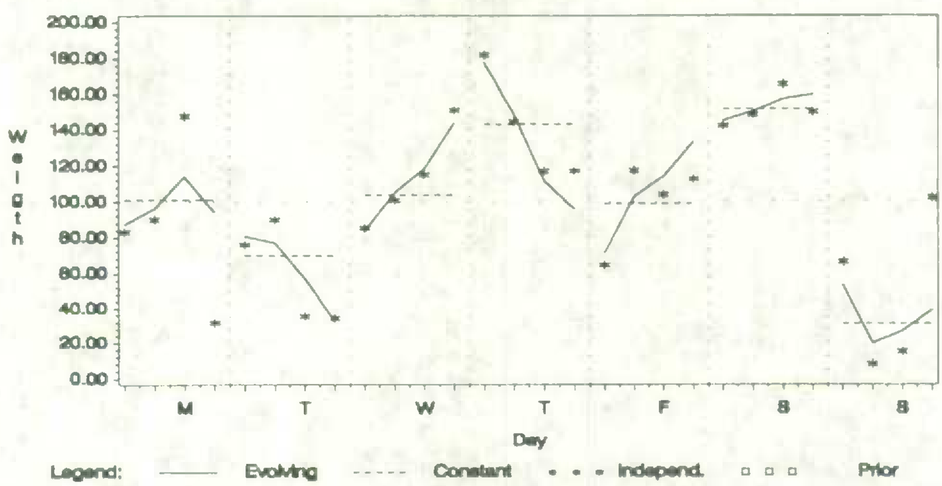
The various panels of Figures 2 and 3 correspond to different variants of method 3, the other two curves remaining the same. Method 3 was applied according to the procedure described in section 4.5, except that the signal-to-noise ratio (**S/N**) was exogenously set to some arbitrary value. This was done because the procedure lead to a ratio of 0.0, implying that the coefficients do not change (despite the strong evidence provided by Section 3). As mentioned earlier, the procedure works for artificial series, i.e. converges to the "true" S/N ratio used in generating the data. We suspect the presence of outliers in the data to be the cause of the problem, which will be discussed later.

In Figure 2, the transition equations are second differences, which specify a linear behaviour of the joint estimates (solid line). Panel (a) illustrates the effect of a S/N ratio close to 0 (0.01). As mentioned in section 4, the transition equations almost become constraints and nearly impose a linear behaviour (in this case) to the joint estimates. In the example, this translates into a poor fit through the independent estimates of Monday and Sunday (e.g. not reflecting the discontinuity between regimes 3 and 4). In panel (b) on the other hand, the ratio is 0.50, the joint estimates tend to be behave linearly, but better track the local independent estimates. This translates into an improved fit for Monday and Sunday. (The independent estimates of the shorter regimes 3 - and 4 especially - receive less weight that the longer regimes 1 and 2.) In panel (c), the ratio is still 0.50, but prior information was supplied for the level of Sunday (square in fig.) on the fourth regime and for the second differences (change) of Monday and Sunday between regimes 2, 3 and 4 (not indicated in the figure). This translates into a much improved fit for those two days. The prior information on the level (1.10 ± 0.15 with 95% confidence) would normally reflect subject matter expertise (we chose those values, which are likely, to illustrate the use of prior information). The values on change (-0.867 for Monday and 0.402 for Sunday) were half the second differences observed in the independent estimates; this is referred to as empirical Bayes estimation.

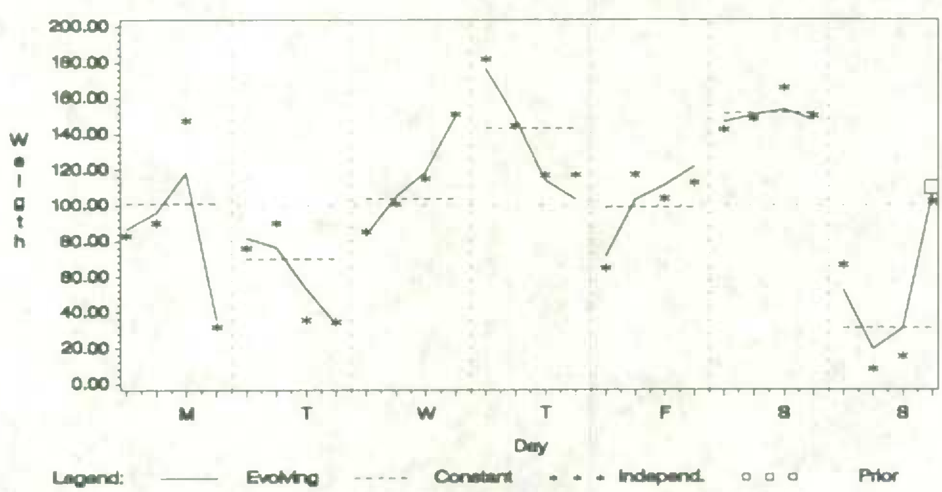
In Figure 3, the transition equations are first differences, which specify constant behaviour for the joint estimates. In panel (a), the S/N ratio is 0.01, which nearly imposes constant behaviour of the joint estimates. As a result these almost coincide with the constant estimates of method 1 and fit the independent estimates poorly for all days. In panel (b), the S/N ratio is 0.50 and the joint estimates tend to be constant but to some extent track the independent estimates of method 2. In panel (c), the ratio is still 0.50, but prior information was supplied for the level of Sunday (square in fig.) on the fourth regime and for the first differences (change) of Monday and Sunday between regimes 3 and 4 (not indicated in the figure). This translates into an even better fit for those two days. The prior information on the level of Sunday is still the same as in Figure 2 (c) (1.10 ± 0.15 with 95% confidence). The values on change (-0.58 and 0.44) were half the first differences observed in the independent estimates between the third to the fourth regime for Monday and Sunday.



(a) Without Prior Information, S/N set to 0.01

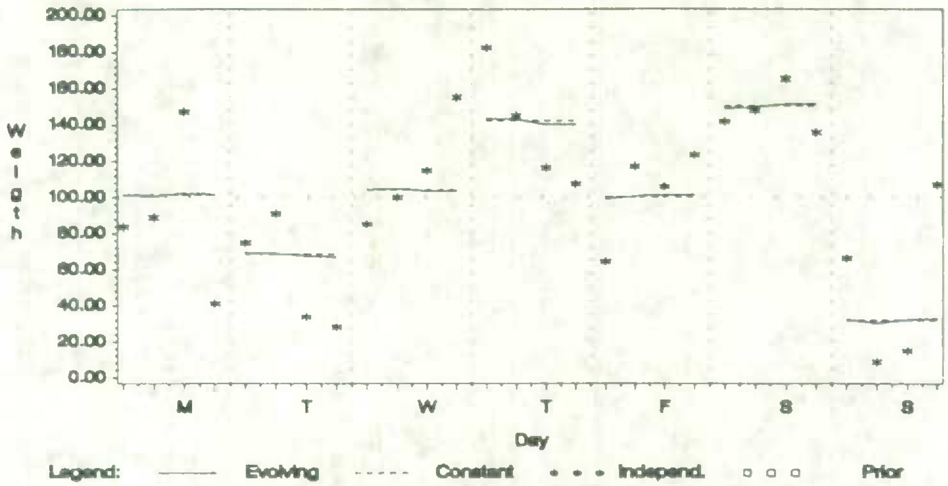


(b) Without Prior Information, S/N set to 0.50

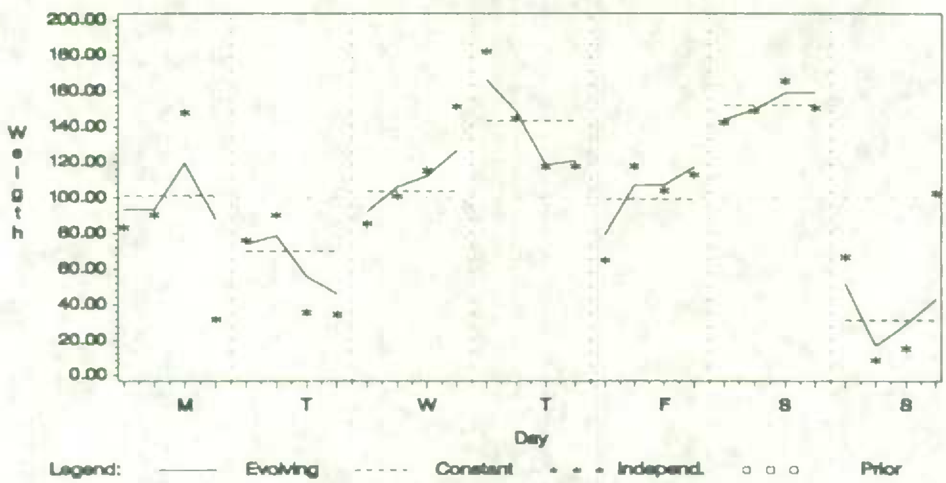


(c) With Prior Information, S/N set to 0.50

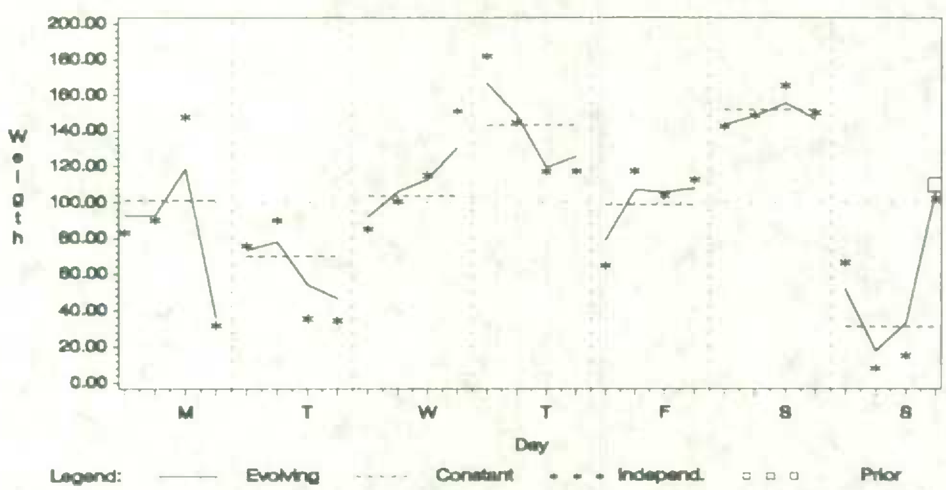
Figure 2: Constant, Independent and Joint (evolving) Daily Coefficients Estimated for the Department Stores series on 4 Trading-Day Regimes, ranging from 1981 to 1984, 1985 to 1988, January 1989 to June 1992 and July 1992 to November 1993, under Second Difference Transition equations



(a) Without Prior Information, S/N set to 0.01



(b) Without Prior Information, S/N set to 0.50



(c) With Prior Information, S/N set to 0.50

Figure 3: Constant, Independent and Joint (evolving) Daily Coefficients Estimated for the Department Stores series on 4 Trading-Day Regimes, ranging from 1981 to 1984, 1985 to 1988, January 1989 to June 1992 and July 1992 to November 1993, under First Difference Transition Equations

6. Discussion

In order to adjust the Canadian Retail Trade series for moving trading-day variations, we resorted to the classical method of Young (1965), but applied it to two trading-day regimes. The reasons for doing so were mainly practical. The Young method was already incorporated in the official seasonal adjustment programme used by Statistics Canada, X-11-ARIMA, and we did not have enough time to incorporate any other method. In an iterative estimation procedure (such as X-11-ARIMA), it is intrinsically crucial that all "methods", namely the method to estimate seasonality, the method to estimate trading-day variations, etc., be part of the procedure. Indeed each estimation (and re-estimation) of one component by one method improves the (re-)estimation of the other components by the other methods. Furthermore, our subject matter client needed an improvement over the current practice (which left them dissatisfied) by the time of the annual revision, even if the solution was not the best way of handling the problem. The weights estimated on second regime may be subject to substantial revision, however they will stabilise as the second regime incorporates more data points.

In the longer run a variant of the stochastic regression approach or of the state-space approach (Dagum *et al.*, 1992) could be incorporated in X-11-ARIMA. However the signal-to-noise ratio (**S/N**) problem has to be solved. As mentioned earlier, for the Retail Trade Series at least, that ratio seems to be systematically underestimated, both for the state-space approach and for the regression model presented in section 3 (where S/N converges to 0). Panels (a) of Figure 2 and 3 illustrate that this translates into poor fit: the independent estimates suggest a smooth monotonic behaviour of the daily weights from regime to regime (except for Sunday and Monday perhaps), and the joint stochastic estimates are rigidly linear in the case of second difference transition equations and constant in the case of first difference. Perhaps a strategy to detect and accommodate outliers would solve or alleviate the problem. The state space model by Durbin and Cordero (1994), which allows a mixture probability distribution, could be another solution.

Perhaps a more pragmatic solution would be to set S/N to some value, which would depend on the series and strike a better trade-off between fitting and smoothing (Brown, 1963). More specifically, for a time series, five (say) values of S/N could be attempted. The lowest value could be such as to yield nearly constant (smooth) and *possibly* poorly fitting estimates; and the highest, yielding possibly erratic but closely fitting estimates. The value of S/N yielding daily estimates which change the most but monotonically - between same-month especially - would be chosen. The rationale for such a scheme is that the monotonic evolution of the estimated daily weights, the independent ones of Figure 2 for instance, can not be due to chance. The programme could offer the option of trying all 5 values automatically or of letting the user choose one of them. The initial conditions could be provided by an independent regression on the first few years of the series.

This exogenous choice of S/N has some advantages. Revisions to past estimates are minimized, compared to a situation where S/N is estimated by the programme and substantially revised as the series gets longer as in Dagum and Quenneville (1988). The amount of calculations is reduced. Indeed, the expensive part of the state-space approach, both in terms of computing time and algorithms required, is precisely the estimation of S/N. Such an approach would be rather consistent with the way the X-11-ARIMA chooses the seasonal moving average, based on the "I/S" ratio, and the trend-cyclical Henderson moving average, based on the "I/C" ratios.

We are currently working on these problems and on a stochastic regression approach, where the daily weights change every year, instead of every regime, in order to reduce reliance on subject matter knowledge.

References

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

Cholette, P.A. Quenneville, B. (1994), "A Methodological Note on the Estimation of Trading-Day Variations Time Series", **Statistics Canada**, Time Series Research and Analysis Center, Research Paper.

Dagum, E. (1980), The X-11-ARIMA Seasonal Adjustment Programme, **Statistics Canada**, Cat. 12-564E.

Dagum, E. Bee, Quenneville, B., Sutradhar, B. (1992), "Trading-Day Variations Multiple Regression Models with Random Parameters", **International Statistical Review**, Vol. 60, pp. 57-73.

Dagum, E. Bee, Quenneville, B. (1988), "Deterministic and Stochastic Models for the Estimation of Trading-Day Variations", **Proceedings of the Fourth Annual Research Conference**, Bureau of the Census, Washington, pp. 569-590.

Dumas, J., Cholette, P. (1992), "L'impact des politiques pronatalistes du Québec sur sa fécondité", Document présenté au Congrès de l'A.I.D.E.L.F. à Delphes, **Statistique Canada**, Division de la démographie.

Duncan, D.B., Horn, S.D. (1972), "Linear Dynamic Regression from the Viewpoint of Regression Analysis", **Journal of the American Statistical Association**, Vol. 67, pp. 815-821.

Durbin, J., Cordero, M. (1994), "Handling Structural Shifts, Outliers and Heavy Tailed Distributions in State Space Time Series Models", to appear (work done in the Statistics Research Division of the U.S. **Bureau of the Census**.

Gurney, M. and Daly, J.F. (1965), "A Multivariate Approach to Estimation in Periodic Sample Surveys", **Proceedings of the Section on Social Statistics, American Statistical Association**, pp. 242-257.

Harvey, A.C. (1981), **The Econometric Analysis of Time Series**, Philip Allan

Harvey, A.C. (1989), **Forecasting, Structural Time Series Models and the Kalman Filter**, Cambridge University Press.

Jones, R.G. (1980), "Best Linear Unbiased Estimators for Repeated Surveys", **Journal of the Royal Statistical Society**, Series B, Vol. 42, No. 2, pp. 221-226

Newbold, P., Bos T. (1985), **Stochastic Parameter Regression Models**, Sage University Paper Series on Quantitative Applications in the Social Sciences Number 07-051, Beverly Hills and London: Sage Publications.

Salinas, T.S and Hillmer, S.C. (1987), "Multicollinearity Problems in Modeling Time Series With Trading-Day Variation", **Journal of Business and Economic Statistics**, Vol. 5, No. 3, pp. 431-436.

Theil, H. (1971), **Principle of Econometrics**, North-Holland.

Young, A.H. (1965), "Estimating Trading-Day Variation in Monthly Economic Time Series", **U.S. Bureau of the Census**, Technical Paper No. 12.