Methodology Branch

Time Series Research and Analysis Division

Direction de la méthodologie

Division de la recherche et de l'analyse des chroniques

ADJUSTMENT FOR HOLIDAY EFFECTS DURING REFERENCE WEEK OF VARIABLE DATES IN THE ACTUAL HOURS WORKED SERIES

by

Marietta Morry

Statistics Canada Statistique Canada

Canada

WORKING PAPER NO. TSRA 93-001E

TIME SERIES RESEARCH & ANALYSIS DIVISION
METHODOLOGY BRANCH

ADJUSTMENT FOR HOLIDAY EFFECTS DURING REFERENCE WEEK OF VARIABLE DATES IN THE ACTUAL HOURS WORKED SERIES

by

Marietta Morry

Abstract

This paper investigates the impact of moving holidays during reference weeks of variable dates on the Actual Hours Worked series for the Canada Total, nine one-digit industrial classification groups and four aggregates of these groups. The actual effect of Easter, Thanksgiving, Remembrance Day and July vacations is estimated by applying linear regression models to the irregulars obtained from the X11ARIMA/88 seasonal adjustment. The estimated effect of the different holidays is then analyzed by industry.

Résumé

Ce document examine l'effet des jours chômés à date mobile sur les séries d'Heures travaillées, pour l'ensemble du Canada, pour neuf groupes industriels et pour quatre agrégations de ces groupes. Les effets de Pâques, de l'Action de grâce, du jour du Souvenir et des vacances de juillet sont estimés en appliquant des modèles de régression linéaire aux valeurs de la composante accidentelle, produite par la méthode de désaisonnalisation X-11-ARIMA/88. L'effet des divers jours chômés est ensuite analysé par groupe industriel.

1.0 INTRODUCTION

The Canadian Labour Force survey is conducted once a month and it collects information,among other things, about the actual hours worked during the reference week, which normally includes the fifteenth of the month.This definition of the reference week implies that the actual dates of the reference week vary from year to year.This variability combined with the presence of fixed (Remembrance Day) or moving holidays (Easter, Thanksgiving) during the reference week in some years but not in others introduces significant fluctuations in the data that is not removed by the usual seasonal adjustment procedures, such as X11ARIMA.The presence of these fluctuations can greatly distort the month-to-month movements of the seasonally adjusted series making trend analysis a rather difficult task.

This paper examines the combined effect of variable reference week and holiday dates in April, when Good Friday, Easter Monday or neither can be part of the reference week; in October, when the reference week might include the Thanksgiving holiday, as is the case in most years; and similarly in November, when Remembrance Day is included in the reference week in most years but not all.

The impact of the holiday effect is estimated from the residuals obtained through a preliminary seasonal adjustment step using the X11ARIMA method (Dagum, 1988).Permanent prior adjustment factors are then derived from these residuals via regression and applied for the months of April, October and November during a second seasonal adjustment step using X11ARIMA.

After having adjusted for the holidays the series were smoothed out considerably with the exception of a few minor peaks and troughs in July, which again turned out to be related to the reference week dates, specifically to the phenomenon that vacation time taken by employees was more likely to coincide with a reference week that fell on later dates in July than on earlier dates. In effect, this inverse relationship between hours worked and reference week dates was found to be statistically significant and therefore possible to adjust for in a manner similar to the one described for the other three holidays.

The paper is organized in the following fashion; Section 2.1 discusses the problem with the occasional presence of Easter holidays during reference week, Section 2.2 deals with the issue of Thanksgiving holiday in October and Remembrance Day holiday in November.The timing of summer vacations and their effect on the Actual Hours Worked series constitutes the topic of Section 2.3 Section 3.0 analyzes the results in some detail. Finally, Section 4.0 concludes the study.

2.1 THE EFFECT OF EASTER HOLIDAYS DURING THE APRIL REFERENCE WEEK

It has become evident over the years that the seasonally adjusted 'Actual Hours Worked in Canada' series displays very large fluctuations in some years, especially around April and October (see Figure 1.). It has also been suspected that these fluctuations are somehow related to the presence and absence of holidays during the reference week.

Since large month-to-month fluctuations in the seasonally adjusted data are normally associated with the irregular component rather than the trend, a reasonable next step was to analyze the irregular component obtained through X11ARIMA, the official seasonal adjustment method used in Statistics Canada to produce the published 'Actual Hours Worked' series. An examination of the magnitude of the irregular in relation to Table 1, which lists the occurrence of Easter in the April reference week throughout the years, reveals that low irregular values, for example, occur in those years when Good Friday is part of the reference week.

Table 1. Easter Holidays during the Reference Week in April

Year	Reference Week Dates in April	Easter Dates	Presence of Easter Holiday in Reference Week
76	11-18	April 16-19	Good Friday
77	10-17	April 8-11	Easter Monday
78	9-16	March 24-27	none
79	15-22	April 13-16	Easter Monday
80	13-20	April 4- 7	none
81	12-19	April 17-20	Good Friday
82	11-18	April 9-12	Easter Monday
83	10-17	April 1- 4	none
84	15-22	April 20-23	Good Friday
85	14-21	April 5- 8	none
86	13-20	March 28-31	none
87	12-19	April 17-20	Good Friday
88	10-17	April 1- 4	none
89	9-16	March 24-27	none
90	15-22	April 13-16	Easter Monday

Table 1. Easter Holidays during the Reference Week in April
(continued)

Year	Reference Week Dates in April	Easter Dates	Presence of Easter Holiday in Reference Week
91	14-21	March 29-April 1	none
92	12-19	April 17-20	Good Friday
93	11-18	April 9-12	Easter Monday

In effect, Figure 2 clearly shows that there is a significant difference in the size of the irregulars during the three types of years, those with Good Friday in the reference week are lowest followed by those with Easter Monday in the reference week, while years in which no Easter holiday occurs during the reference week have the largest irregulars, indicating that in those years more 'actual hours worked' are registered in April than in an average April month.Incidentally, although Easter can also occur in March, it never has any effect on the March figures since it always falls after the 22 of March and thus never becomes part of the reference week which ends on March 22 as the latest possible date.

In order to estimate the systematic part in the irregular component due to the variation in the presence of holidays, the irregulars were regressed against three dummy variables (equation(1)) each representing one of the three types of Easter occurrences.

$$I_j = \sum_{i=1}^{3} \beta_i X_{ij} + e_j \qquad (1)$$

where

$X_{1j} = 1, X_{2j} = 0, X_{3j} = 0$ if in year j Good Friday is in the Reference Week.

$X_{1j} = 0, X_{2j} = 1, X_{3j} = 0$ if in year j Easter Monday is in the Reference Week.

$X_{1j} = 0, X_{2j} = 0, X_{3j} = 1$ if in year j there is no Easter in the Reference Week.

e_j = error term in year j.

The estimated coefficients β_1, β_2 *and* β_3 give the systematic part of the irregular component related to the three types of Easter occurrences.

Actual Hours Worked in All Industries
Original and Seasonally Adjusted Series

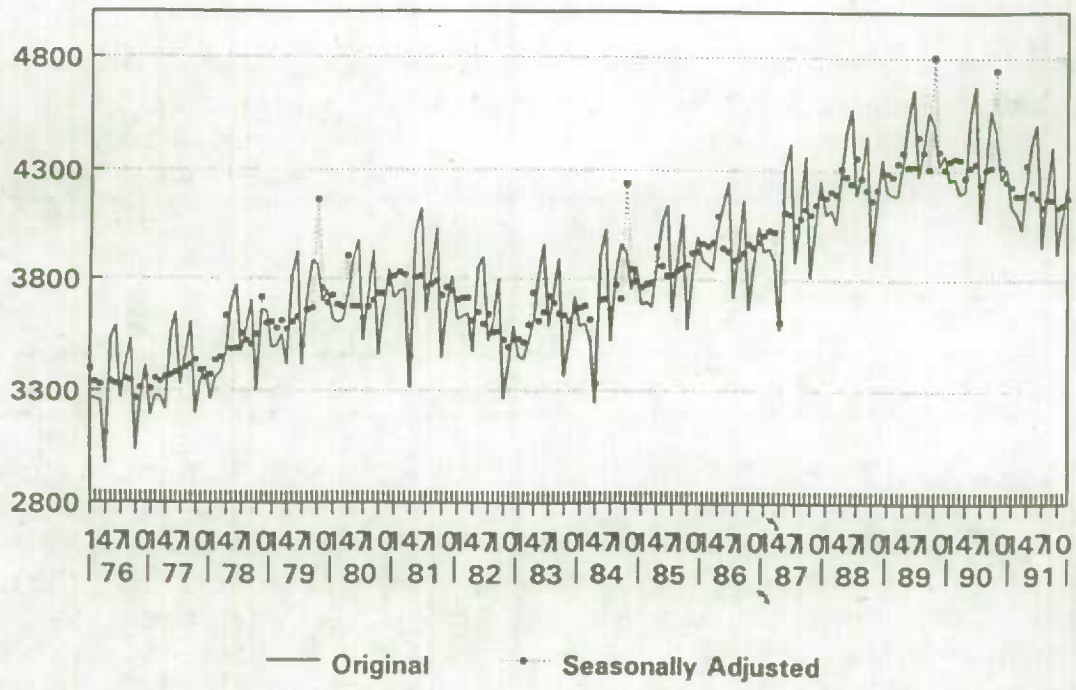

Figure 1.

Actual hours worked in All Industries
X11-ARIMA Irregulars in April

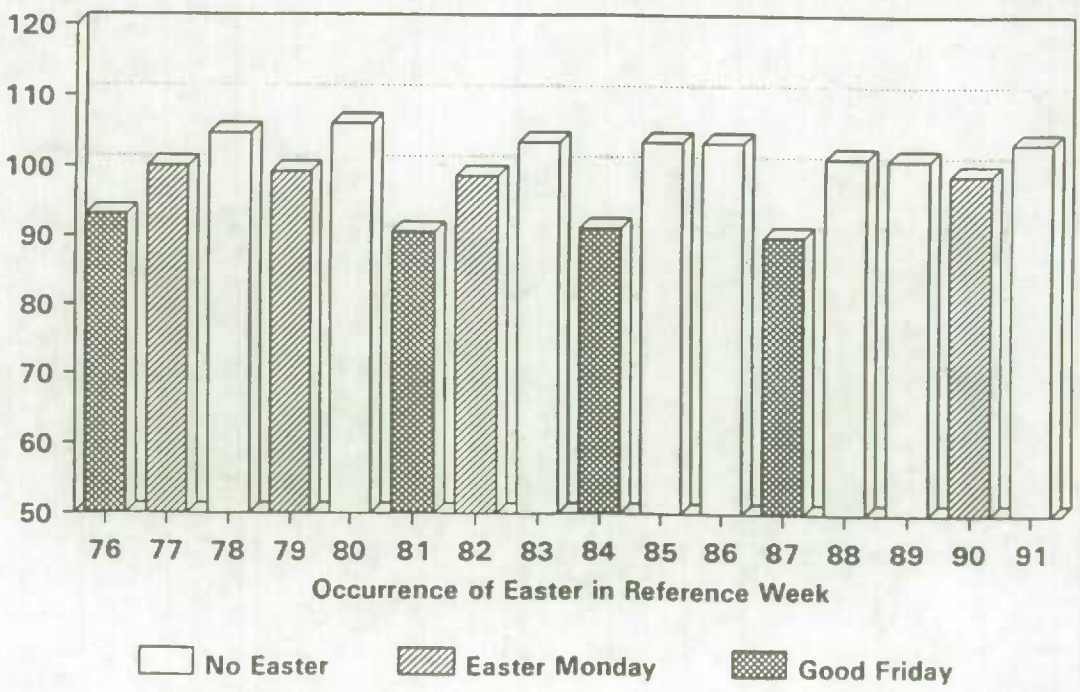

Figure 2.

The formulation of model (1) is basically equivalent to the approach used by the U.S. Bureau of Labor Statistics as described in McIntire(1990). There the holiday effect is calculated by taking averages of the X11ARIMA irregulars within subsets defined based on ranges of possible dates for the holiday.

In the case of the 'Actual Hours Worked in Canada' series, the estimated effects were 91.17, 99.26 and 103.40, respectively. The F-value associated with the analysis of variance test, indicating how much variability is explained by the model (1), was highly significant.

The estimated effects mean that in those years when Good Friday was in the April Reference Week the actual hours registered were about 8% less than average (100.00 being average) while there were almost 3.5% more than average hours reported in April for those years when there was no Easter holiday in the Reference Week. Removing these percentages from the original series, using the permanent prior option available in X11ARIMA, before processing it through a second seasonal adjustment managed to get rid of the wide fluctuations in the April figures and yielded a much smoother seasonally adjusted series around that month.

2.2 THE EFFECT OF THANKSGIVING AND REMEMBRANCE DAY HOLIDAYS DURING THE OCTOBER AND NOVEMBER REFERENCE WEEKS

If the reader turns his attention back to Figure 1. it can be observed that in certain years high peaks appear not only in April but also in October and in one instance although not as pronounced as in the other months, even in November. Here again these high irregular values happen to be associated with those years where the timing of the holidays is different from the average occurrence, i.e. normally the reference week includes the Thanksgiving or Remembrance Day holiday but not in the years where the peaks can be found.

Examining the reference week dates in October it turns out that with the exception of the years 79, 84, 89 and 90, the Thanksgiving holiday always falls in the reference week. Thus, the relatively low figures reported for October in most years because of the presence of the holiday are interpreted as part of the seasonal behaviour for this month and is corrected for through seasonal adjustment. Consequently, the years that stand out after seasonally adjusting the series are the ones in which the hours reported correspond to a full working week, namely 79, 84, 89 and 90.

Similarly, in November the only year that displays a peak is 78, is the one which coincides with the only time when Remembrance Day is not celebrated during the reference week within the time span studied, i.e. from 1976 to 1992.

Figure 3. and Figure 4. give a good idea about the magnitude of the holiday effect in October and November respectively.In order to quantify the impact of the absence of holidays during the reference week the same type of model is applied as in Section 2.1, except here equation (1) only has two dummy variables ,one for the years with holidays in the reference week and one for the years without holidays.

The variation explained by the model, as attested by the analysis of variance F-value, was again highly significant for both October and November. The estimated regression coefficients for the October model were 99.66 (i. e. basically an average value) for the years with holidays and 110.71 for the years without holidays. The corresponding figures for November were 99.99 and 104.10 respectively.

The procedure applied for correcting for these effects was similar to the one used for April, except here only those years were modified through the use of permanent prior factors where no holiday occurred in the reference week since the other years showed no effect (their regression coefficient was practically 100). In practice this means that for the other years the permanent prior factors applied have a value of 100.0, since it is necessary to provide a factor for every point of the series under this option.

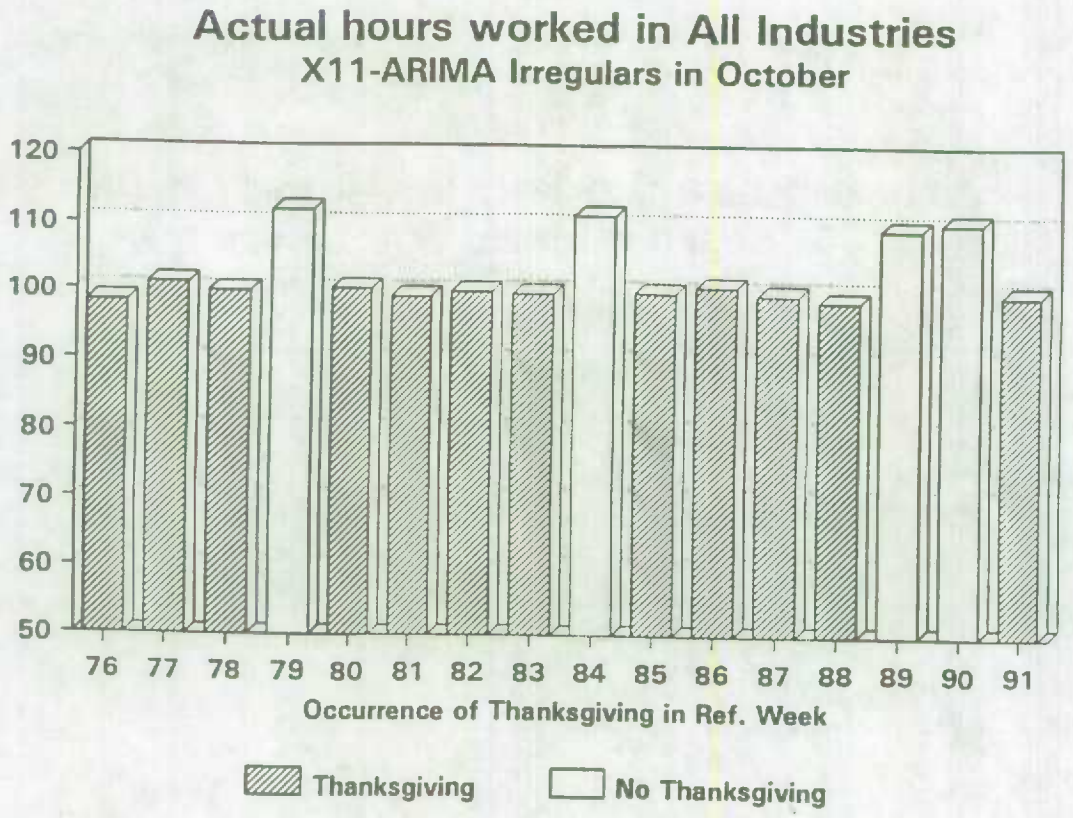

Figure 3.

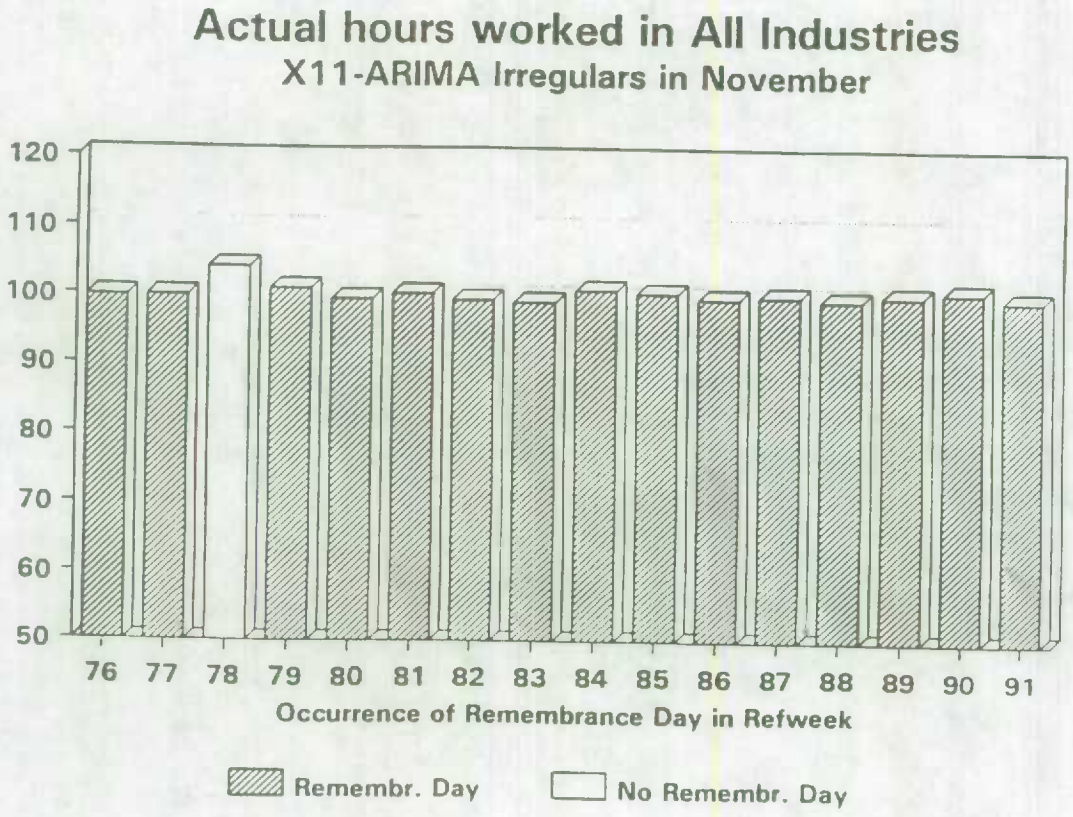

Figure 4.

The combined result of adjusting for the April, October and November holiday effects can be seen in Figure 5.

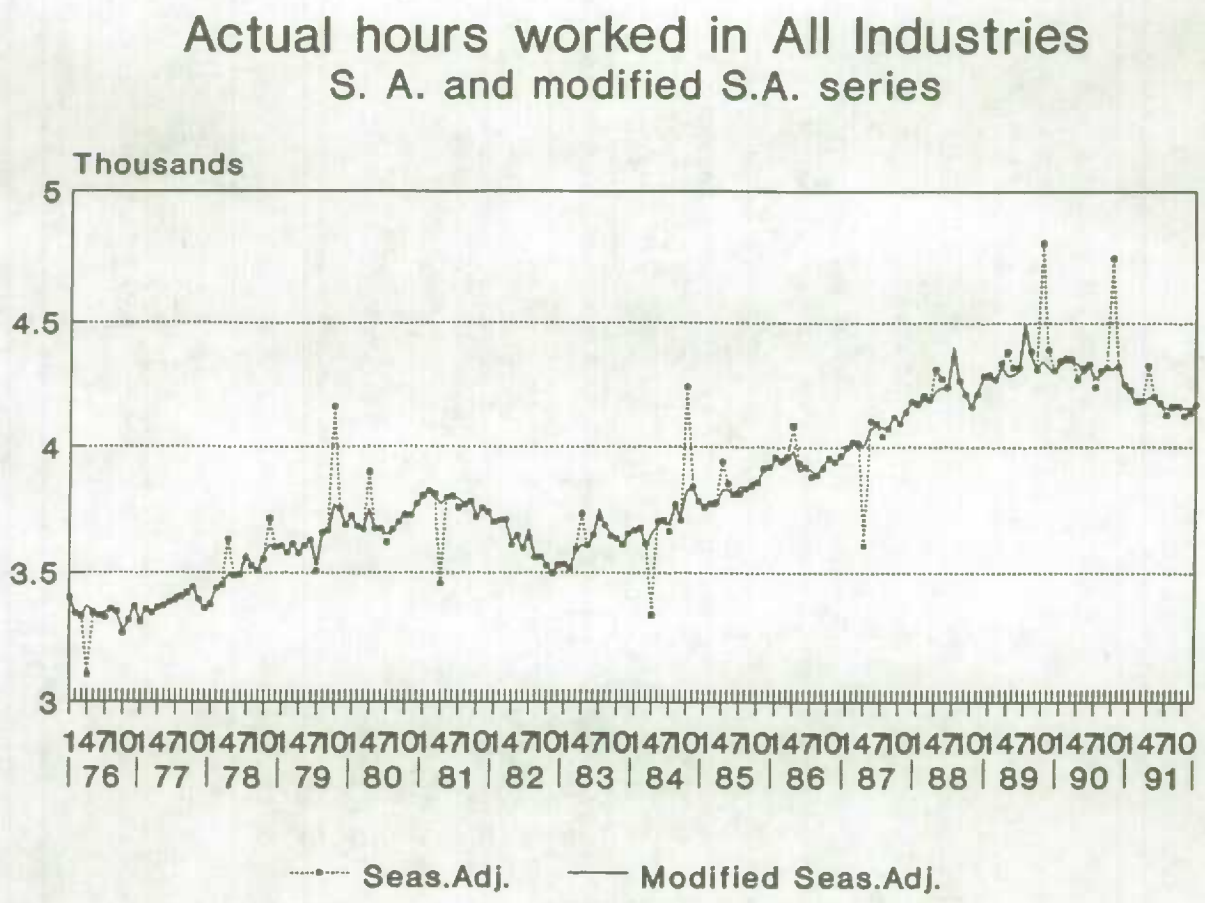

Figure 5.

It is evident that the modified seasonally adjusted series has a much smoother appearance than the non-modified one. However, the 'purist' might still detect some minor peaks and troughs left in the series in certain years around the month of July. The investigation as to the origin of these fluctuations will constitute the topic of the next section.

2.3 THE EFFECT OF VACATIONS DURING THE JULY REFERENCE WEEK

In order to determine if there is a connection between the reported actual hours worked and the timing of the reference week in July we ,once again, examine the irregulars identified by the X11ARIMA seasonal adjustment process for July. Figure 6. plots the irregulars and the date of reference week Wednesdays.

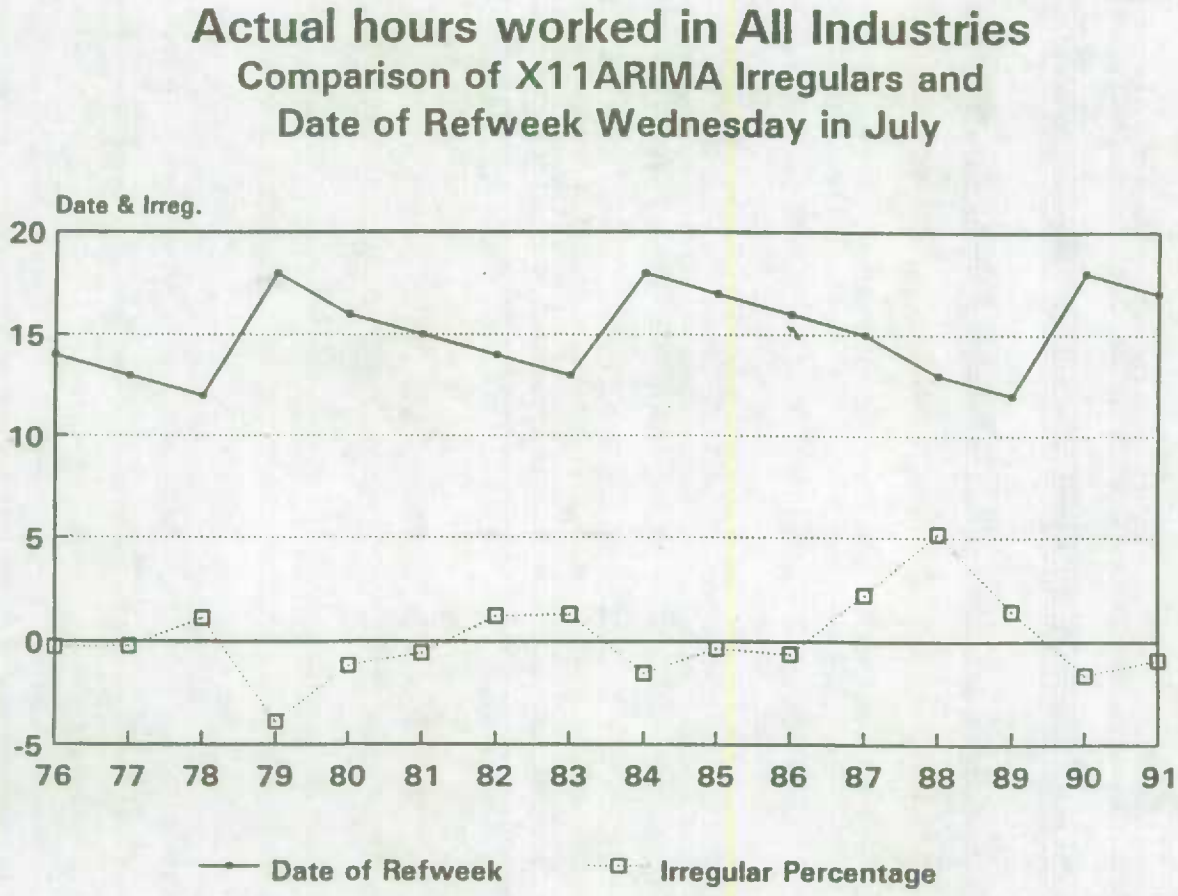

Figure 6.

According to Figure 6. there seems to be an inverse relationship between the two lines, i. e. whenever the reference week date gets higher the irregular values get lower and vice versa.To see just how strong this inverse relationship was the X11ARIMA irregulars for July were regressed against the date of reference week Wednesdays in each year. The resulting fit yielded a highly significant F-value of 32.63. The fitted regression line along with the original irregulars can be seen in Figure 7. There is a more than three percentage point difference in the fitted value corresponding to the earliest and latest reference week date.

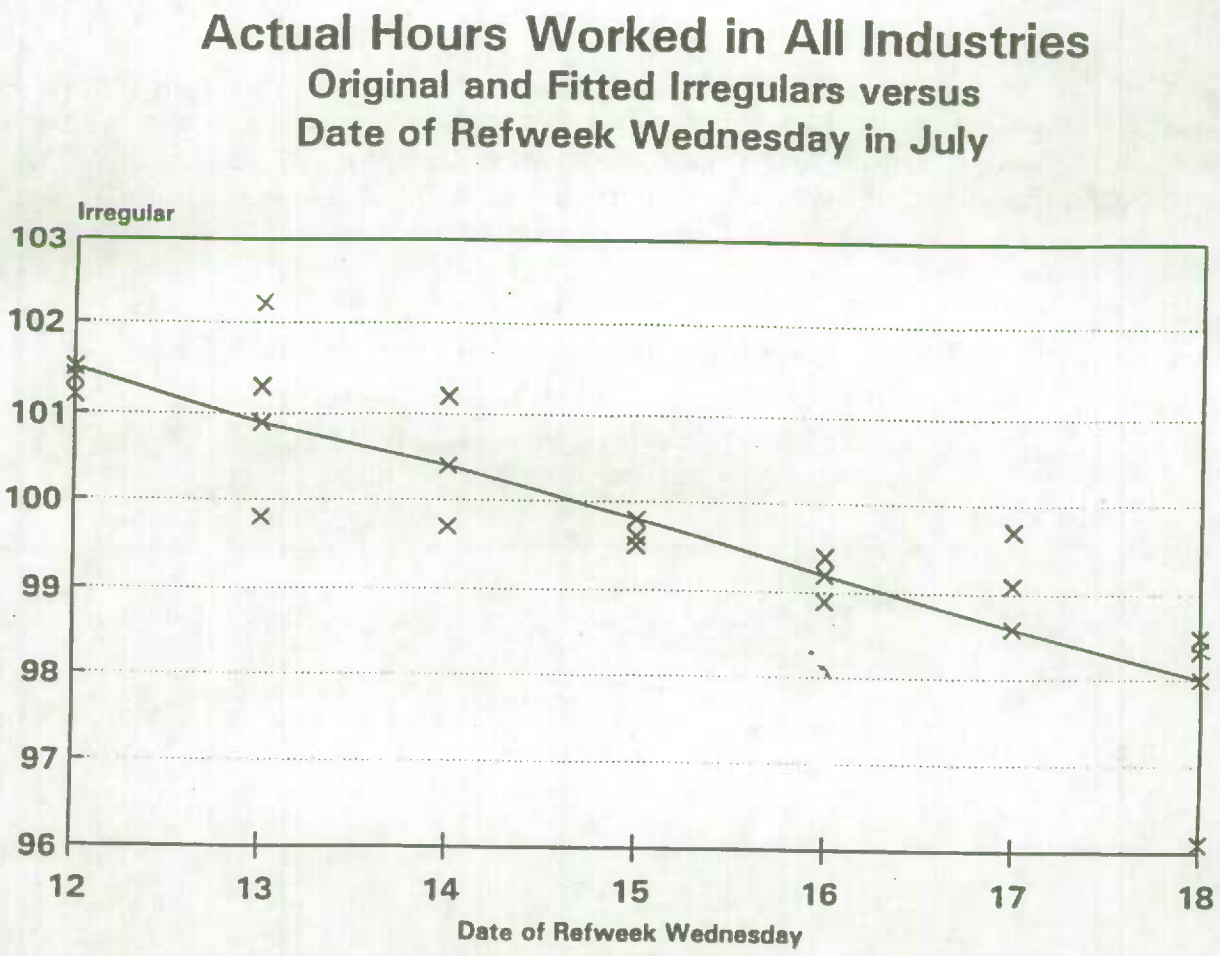

Figure 7.

The linear relationship between the two variables is an indication of the tendency among workers to have their summer vacation later in July (resulting in less hours worked when the reference week occurs later). Of course, the timing of summer vacations is quite often not the choice of the employee but rather it reflects a decision by management to close down the whole plant for a few weeks because it is operationally more desirable to grant holidays in this fashion.These practices are especially prevalent in certain industries, such as manufacturing.

Using the regression estimates corresponding to the seven different reference week dates permanent-prior factors ranging from 98.04 to 101.52 were established for July of each year in the series. Thus the final permanent factor table included values other than 100. (which implies no adjustment for that observation) for April, July, October and November. Applying this final set of factors the series was processed through X11ARIMA to yield the final modified seasonally adjusted to be found in Figure 8.

Actual hours worked in All Industries
S. A. and modified S.A. series
adjusted for July vacations

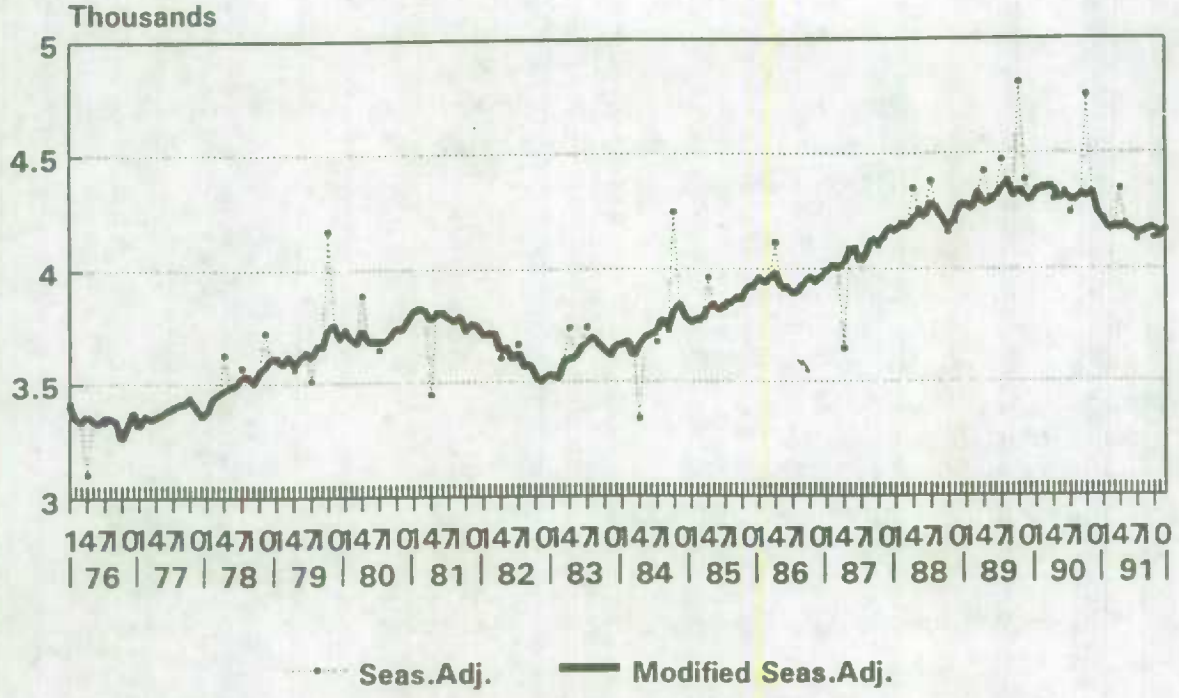

Figure 8.

The advantage of this modified seasonally adjusted series over the seasonally adjusted values obtained in the first X11ARIMA round is that here the month-to-month movements around April, July, October and November are basically the result of changes in the trend-cycle rather than the consequence of the relative timing of holidays and reference weeks, i.e., the modified series lends itself readily for trend-cycle analysis.

So far we have only concerned ourselves with the effect of holidays on hours worked in the Canadian economy as a whole. The next section will focus on a more detailed description of how different industries vary with regard to these holiday effects.

3.0 THE EFFECT OF HOLIDAYS DURING THE REFERENCE WEEK IN DIFFERENT SECTORS OF THE CANADIAN ECONOMY

The previous sections gave a composite view of the effects of holidays on the Total Actual Hours Worked series in Canada. However, not every industry is influenced to the same degree by the presence of holidays during the reference week; i. e. Easter might not play the same role in the hours worked in the primary as in the secondary or tertiary industries, plant closures are more common in manufacturing than anywhere else, Remembrance Day is not observed in all sectors of the economy, etc.

In order to gain a greater insight into the holiday effects in different industries, the Actual Hours Worked series was analyzed according to the following industrial classifications at the Canada level:

Agriculture (AGRIC)
Construction (CONC)
Community, business and personal services (CSBPC)
Finance and real estate (FIREC)
Manufacturing (MANUC)
Other primary industries (OTHPC)
Public administration (PUBLC)
Transportation, communication (TCOUC)
Trade (TRADC)

Apart from these nine major components the following four aggregates were also included in this study:

Commercial industries (COMMC)
Goods producing industries (GOODC)
Non-commercial industries (NCOMC)
Service producing industries (SERVC)

After having processed all the above series through a first round of seasonal adjustment the majority of the series appeared to have peaks and troughs at the same positons as the ones observed in the case of the Total Actual Hours Worked series suggesting the presence of holiday effects. For the removal of these effects, as before, the analysis focussed on the irregulars identified by the X11ARIMA. Four groups were separated out of each set of irregulars; i. e. those corresponding to April, July, October and November. The appropriate linear regression model, as described in Sections 2.1, 2.2 and 2.3 was fitted to each group of irregulars from each of the thirteen series.

Most of the linear fits proved to be significant as attested by the high F-values shown in Table 2. Of all the series included in this study the only one that did not display any of the four types of holiday effects was the series from the Agriculture sector. Statutory holidays evidently are not observed the same way here as in other industries

and thus they don't have an impact on the number of hours worked, rather in agriculture the larger fluctuations are the consequence of highly favourable or adverse weather conditions. For example, in April the irregulars don't display the typical low values in years when the reference week includes Good Friday nor do they show high values in years with no Easter holiday in the reference week, instead the April irregulars are characterized by a rather random behaviour, as can be seen from Figure 9, which also explains why the analysis of variance F-test did not yield a significant value.

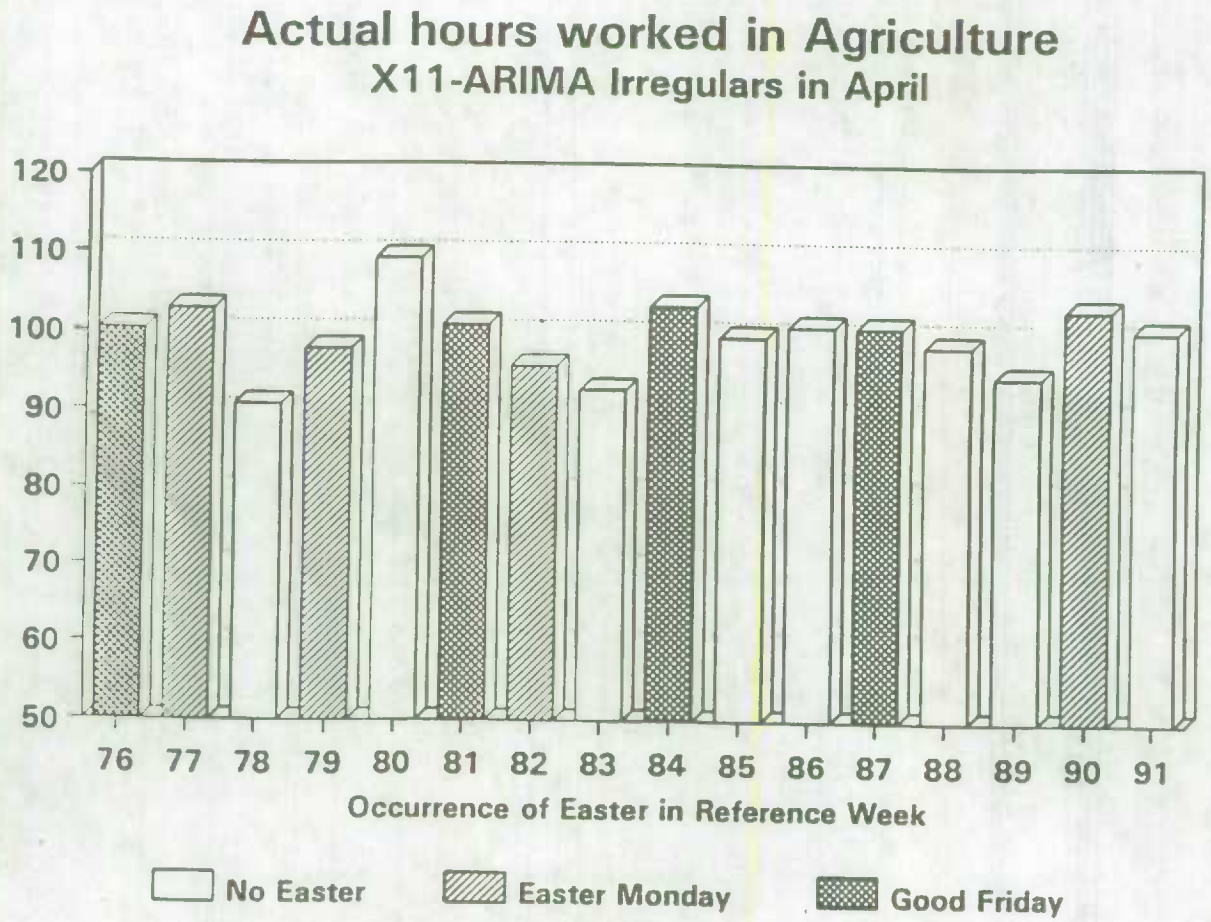

Figure 9.

The rest of the series are all effected by the presence of holidays during the reference week at least in some of the four months. The bar charts of these irregulars exhibit the same type of systematic behaviour as seen in Figures 2, 3 and 4. Of course, the impact of holidays is not as clear in some industries as in others. It depends on how erratic the series is in general and how large is the percentage fluctuation in hours due to the holiday. The larger is the percentage fluctuation around holidays and the less erratic the series is during the rest of the year the more the holiday effect will stand out and a more significant F-value will result.

TABLE 2. ANALYSIS OF VARIANCE F-VALUES ASSOCIATED WITH FITTED REGRESSION MODEL BY INDUSTRY

IDENT	EASTER	JULY VACATION	THANKSGIVING	REMEMBERANCE DAY
AGRIC	N.S.*	N.S.	N.S.	N.S.
CONSC	36.55	9.07	219.93	20.16
CSBPC	264.97	9.68	607.83	9.30
FIREC	152.48	N.S.	967.46	122.70
MANUC	620.96	21.62	257.60	N.S.
OTHPC	15.01	N.S.	59.92	8.19
PUBLC	51.51	4.06	267.87	140.42
TCOUC	49.15	8.70	192.35	70.60
TRADC	191.77	12.37	655.29	18.90
COMMC	120.44	8.97	276.78	N.S.
GOODC	100.36	20.28	117.32	6.66
NCOMC	107.86	6.10	403.63	11.66
SERVC	401.85	19.25	1293.48	41.23
TOTAL	124.00	32.63	459.65	48.92

* N.S. - not significant at the 1% level

According to Table 2, Easter effects are most significant in the Manufacturing industry, followed by the Service Producing sector and they show up the least in the Other Primary industries, although they are still significant.

July vacations don't seem to have significant impact in the Finance, Real Estate and in the Other Primary industries, while Manufacturing, Goods Producing and Service Producing industries are very strongly influenced by them.

Thanksgiving holidays play significant role in all but the Agriculture sector with the Service Producing industries registering the highest F-values.

Finally, with the exception of Agriculture, Manufacturing and the Commercial Industry aggregate the absence of Remembrance Day during the reference week can be felt significantly in all other industries, with Public Administration showing the strongest impact.

Table 3. lists the actual size of the estimated effect obtained through the regression for the different types of holiday occurrences. Values below 100 mean that the actual hours worked due to that type of holiday occurrence were below average (i.e. the holiday occurrence pushed down the value below average for that month). Values above 100, on the other hand, indicate that that particular holiday occurrence inflated the value above average for that month. Using these estimated effects as permanent-prior adjustment factors, as was recommended in the previous sections for the removal of the holiday effect, implies dividing the original figure by one hundredth of the value listed in Table 3. and thus raising the below average values to average or lowering the above average value to average, whatever the case may be.

Here again there are large differences among industries as to the magnitude of estimated effect. In April, the Public sector shows the widest variation in hours depending on the timing of the Easter holiday vis-à-vis the reference week, ranging from 92.1% to 108.4% of the average value.

The closing down of plants for summer vacation is responsible for the Manufacturing industry displaying the largest effect in July , the hours registering 7.2% above average when the reference week Wednesday falls on the 12-th, all the way down to 4.6% below average when that day falls on the 18-th of July. The practice of forced July vacations must be present to some extent in the Construction industry judging by the second largest effect of this type among the different sectors (103.9%-97.2%).

The holiday that appears to have basically the same impact across all the industries is Thanksgiving. In the few years when Thanksgiving Monday is not part of the reference week the hours reported seem to be about 10% higher than the average October figure with only slight variations among the industries (Finance registering the highest values at 115.6%).

TABLE 3. ESTIMATED ADJUSTMENT FACTORS CORRESPONDING TO THE OCCURRENCE OF HOLIDAYS DURING REFERENCE WEEK BY INDUSTRY

IDENT	APRIL			JULY Reference Week Wednesday			OCTOBER Thanksgiving		NOVEMBER Rememberance Day	
	Good Friday	Easter Monday	No Easter	12	-	18	Holiday	No Holiday	Holiday	No Holiday
AGRIC		-------			-------			-------		-------
CONSC	89.8	99.5	102.7	103.9	-	97.2	99.6	112.8	100.0	106.2
CSBPC	93.3	99.7	105.9	100.9	-	98.5	99.8	119.2	100.0	103.0
FIREC	87.2	99.3	103.4		-------		99.8	115.6	100.0	108.6
MANUC	87.2	99.1	101.1	107.2	-	95.4	99.2	114.8	-------	
OTHPC	92.2	97.9	101.4		-------		98.8	109.9	99.9	105.0
PUBLC	92.1	97.1	108.4	100.5	-	99.4	99.8	113.3	99.9	111.7
TCOUC	91.7	99.6	103.9	101.2	-	98.9	99.1	112.2	99.8	107.4
TRADC	91.3	99.2	103.4	100.5	-	97.2	99.8	111.4	100.0	104.5
COMMC	91.7	99.7	102.3	100.5	-	96.8	99.8	108.2	-------	
GOODC	90.5	99.7	101.8	102.5	-	95.1	99.0	110.2	99.9	101.4
NCOMC	95.2	99.7	109.0	101.2	-	92.2	99.7	110.4	99.9	105.7
SERVC	92.4	99.6	105.9	100.8	-	98.8	99.8	111.2	100.0	105.3
TOTAL	91.2	99.3	103.4	101.5	-	98.0	99.7	110.7	100.0	104.1

The effect of Remembrance Day on the other hand, varies greatly among the groups ranging from 111.7% in the Public sector down to only 101.4% in the Goods Producing industries, where Remembrance Day is rarely observed.

To illustrate the importance of adjusting for the holiday effects Appendix A includes a set of graphs showing the seasonally adjusted series before and after modifying for holiday effects.The only exception is the very first graph in the set (Agriculture), where no holiday adjustment was warranted, therefore no modified series appears, instead the seasonally adjusted data is plotted against the original series.

The reader can verify that for the rest of the industries the modified series represent quite an improvement over the unmodified seasonally adjusted data in terms of smoothness, i.e. the implemented holiday adjustment proved effective.

On the basis of these results thefore, it is recommended that this set of series is seasonally adjusted using the **Three-step procedure** described in this paper, i.e.:

1. Process the data through X11ARIMA to obtain the irregulars;
2. Using these irregulars estimate the holiday effects via linear regression;
3. Apply the estimated holiday effects (whenever significant) as the permanent-prior factors for the appropriate year in the month of April, July, October and November in a subsequent seasonal adjustment to obtain the final (modified) seasonally adjusted series.

4.0 CONCLUSIONS

This paper investigated the impact of moving holidays during reference weeks of variable dates on the Actual Hours Worked series for the Canada Total, nine one digit industrial classification groups and four aggregates of these groups.

The actual effect of the holidays in April, July, October and November was estimated through linear regression models applied to the irregulars obtained from a first seasonal adjustment step. After removing these estimated effects from the data using permanent-prior factors in a second seasonal adjustment step we arrived at seasonally adjusted figures that were smoother than before, with the exception of the series for the Agriculture industry where holiday effects did not turn out to be significant.

The effect of the different holidays varied a great deal from industry to industry depending on the characteristics of each regarding the observance of holidays and summer vacation practices.

Based on the findings of the study (namely the significant improvement in smoothness of the seasonally adjusted series after the removal of holiday effects) it is stongly recommended that the Actual Hours Worked series be seasonally adjusted using the **Three-step procedure** outlined in Section 3 in order to facilitate short-term trend-cycle analysis.

Since these holiday effects are not expected to vary significantly within a few years the estimated values can serve as the basis for establishing permanent-prior factors in the future for a while but a reestimation of the holiday effects is advisable in about three years time in case the impact of holidays changes during that period.

References

Dagum, E. B. (1988), 'The X11ARIMA/88 Seasonal Adjustment Method- Foundations and User's Manual', **Statistics Canada**.

McIntire, R.J.(1990), 'A procedure to control for moving-holiday effects in seasonally adjusting employment and hours series', Proceedings of the Business and Economic Statistics Section of the American Statistical Association, 443-448.

APPENDIX A

Actual hours worked in Agriculture
Original and Seasonally Adj. Series

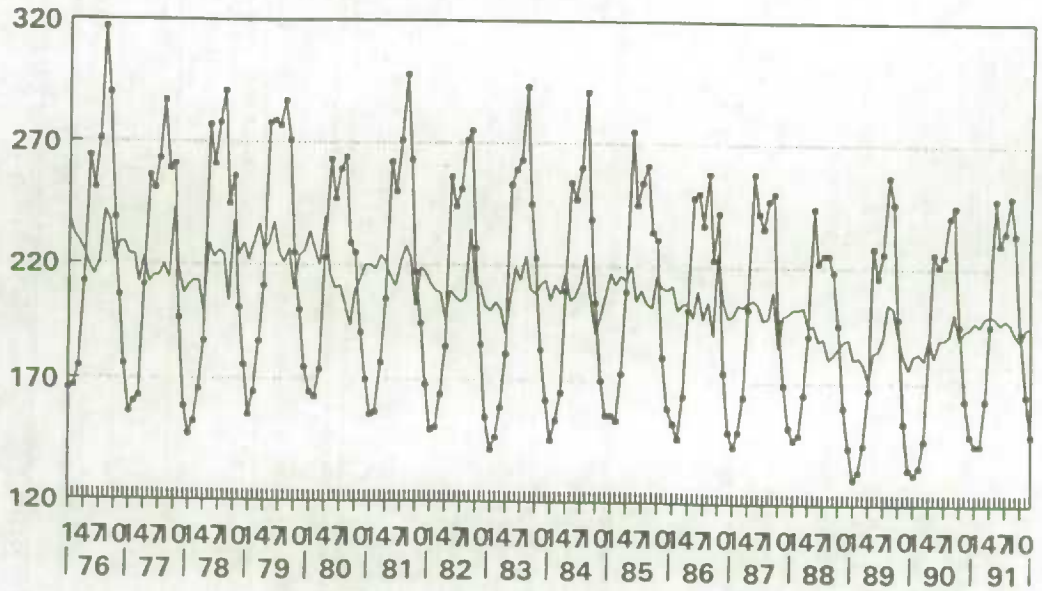

Figure 10.

Actual hours worked in Construction
S. A. and modified S.A. series

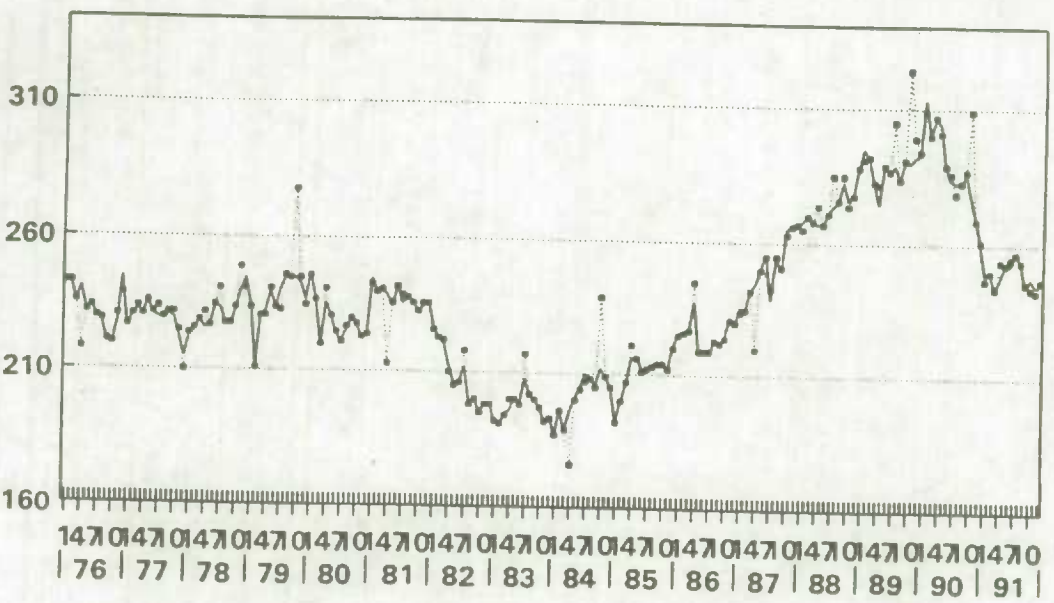

Figure 11.

Actual hours worked in Services
S. A. and modified S.A. series

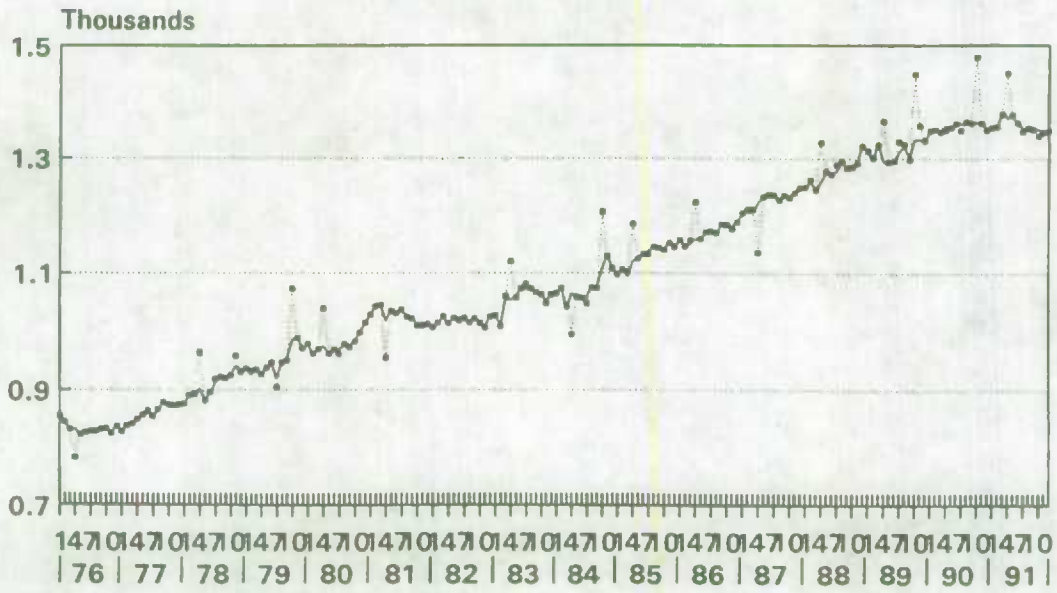

Figure 12.

Actual hours worked in Finance
S. A. and modified S.A. series

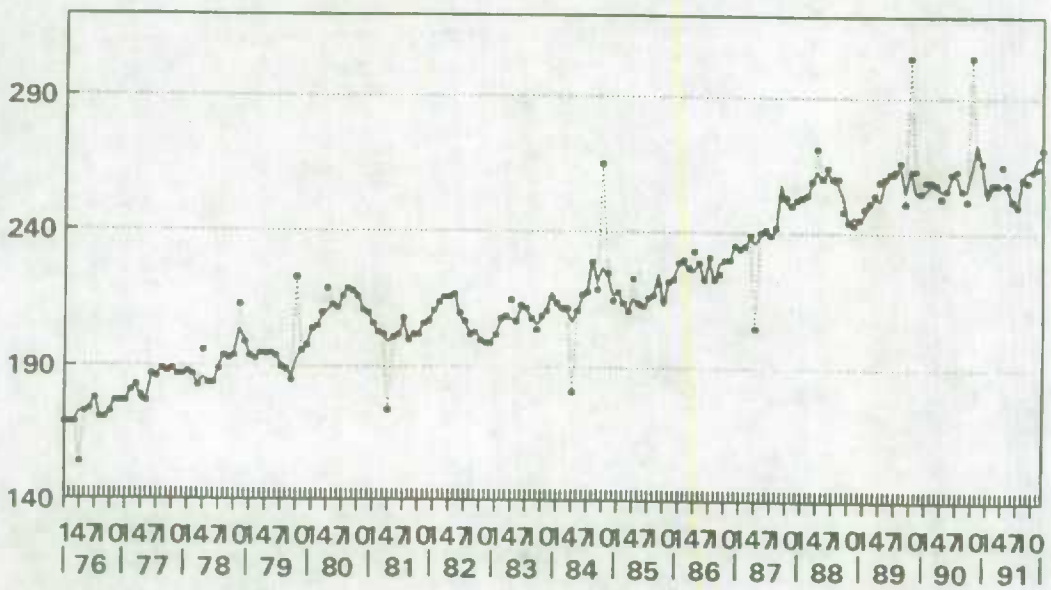

Figure 13.

Actual hours worked in Manufacturing
S. A. and modified S.A. series

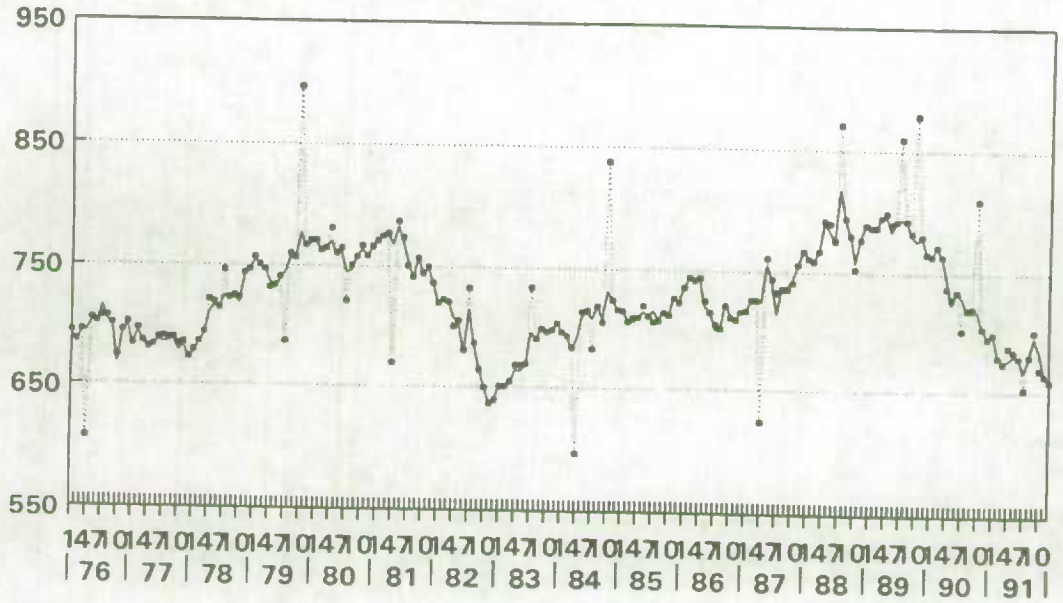

Figure 14.

Actual hours worked in Other Primary Ind
S. A. and modified S.A. series

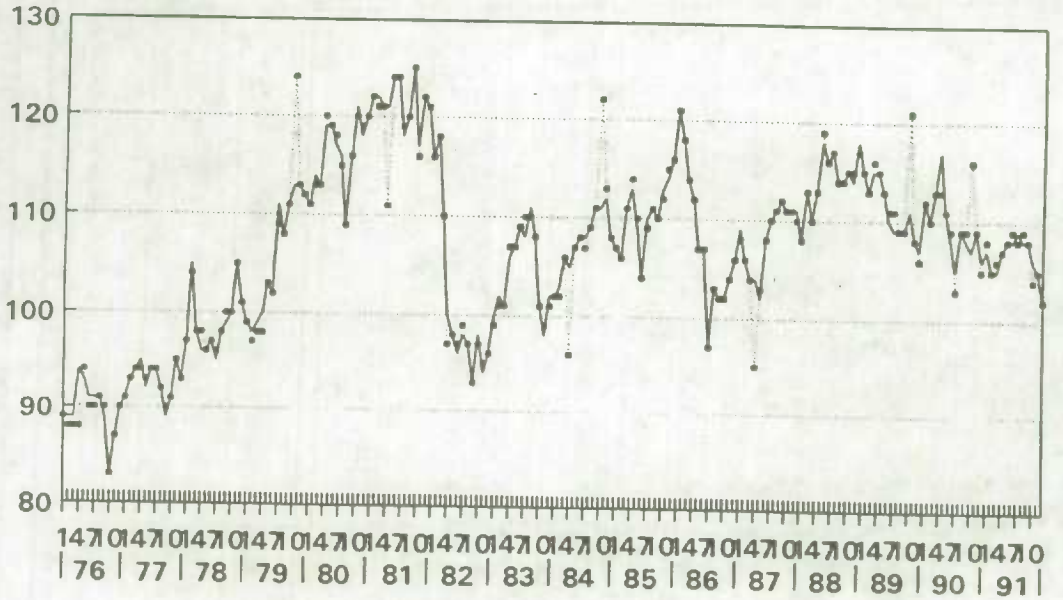

Figure 15.

Actual hours worked in Public Admin.
S. A. and modified S.A. series

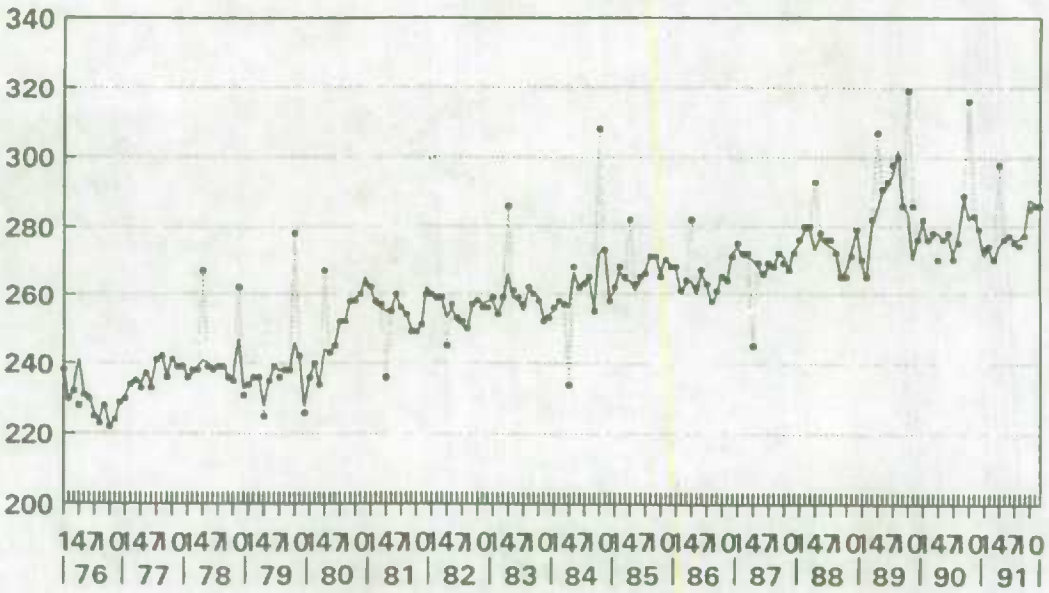

Figure 16.

Actual hours worked in Transport
S. A. and modified S.A. series

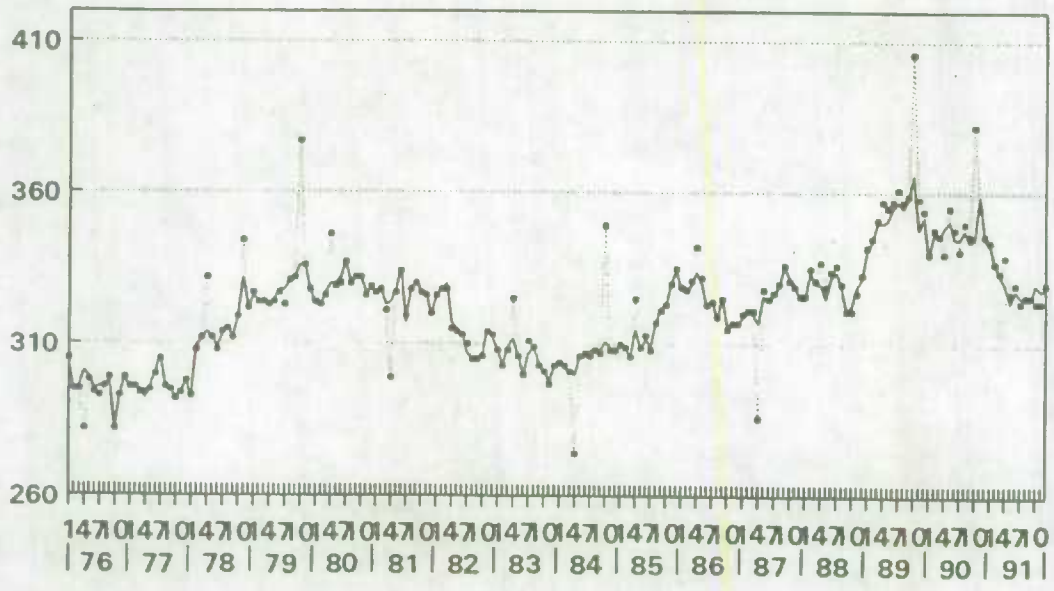

Figure 17.

Actual hours worked in Trade
S. A. and modified S.A. series

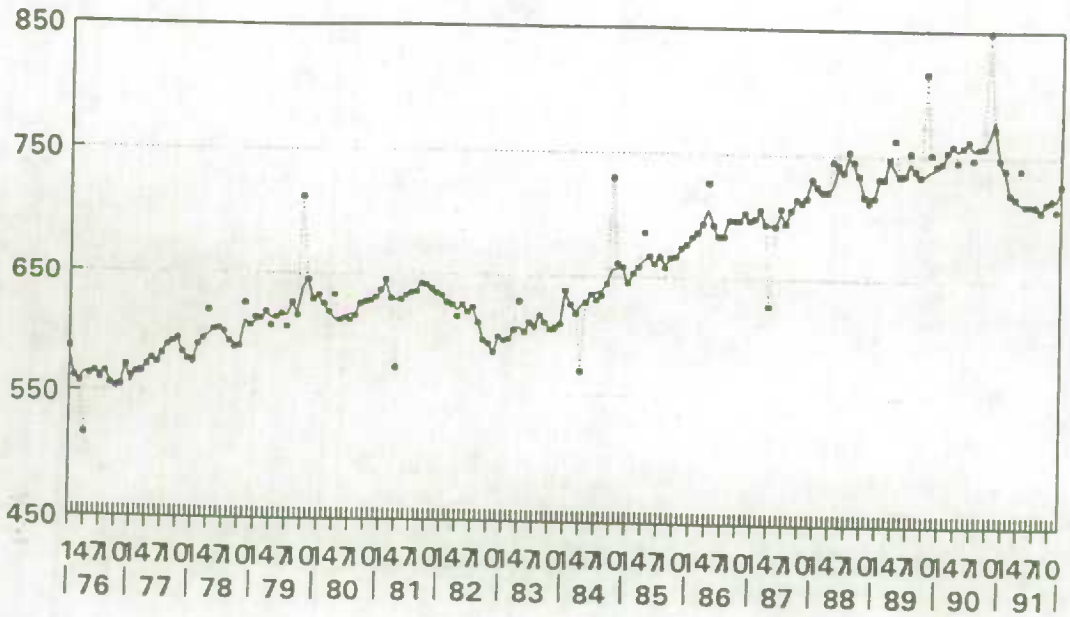

Figure 18.

Actual hours worked in Commercial Ind.
S. A. and modified S.A. series

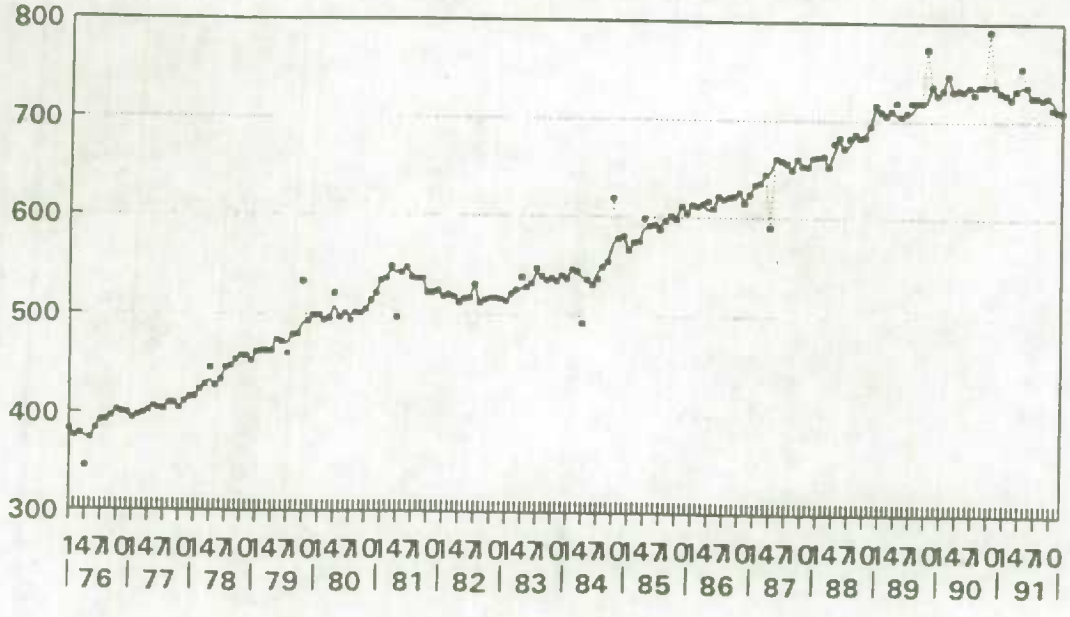

Figure 19.

Actual hours worked in Goods Producing
S. A. and modified S.A. series

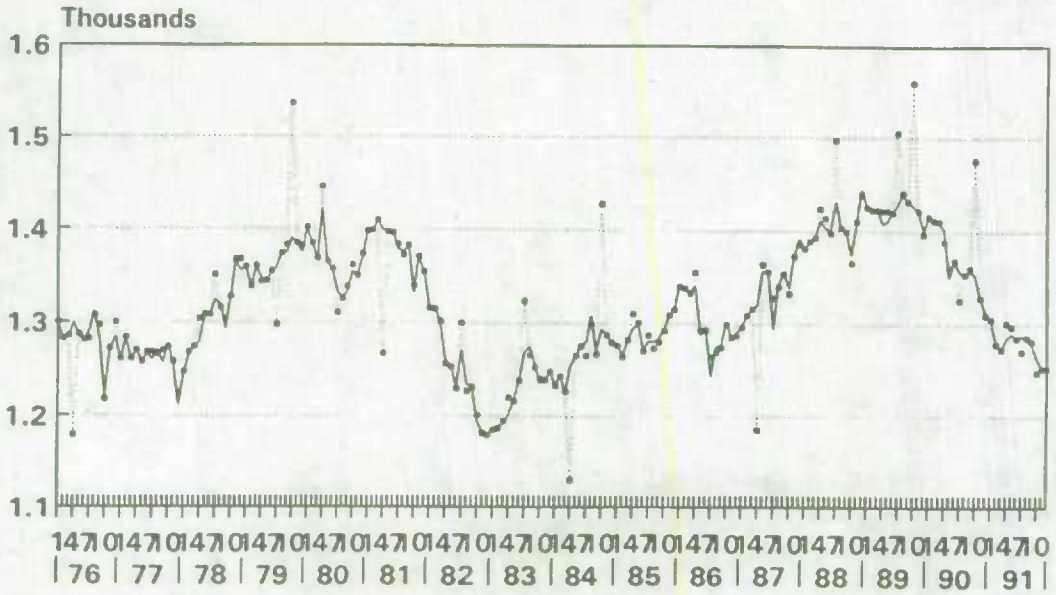

Figure 20.

Actual hours worked in Non-commercial
S. A. and modified S.A. series

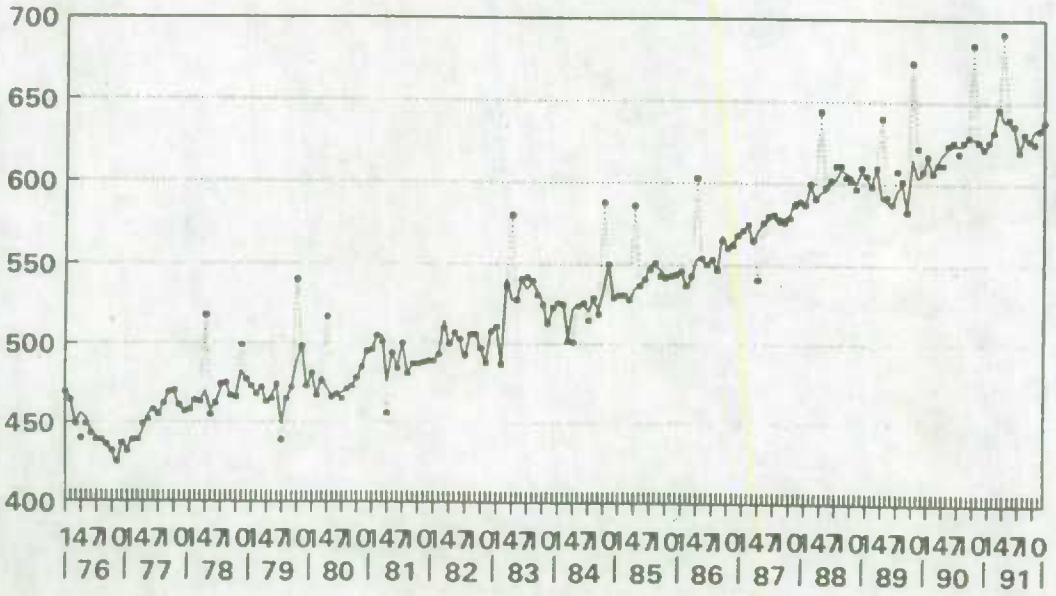

Figure 21.

Actual hours worked in Service Producing
S. A. and modified S.A. series

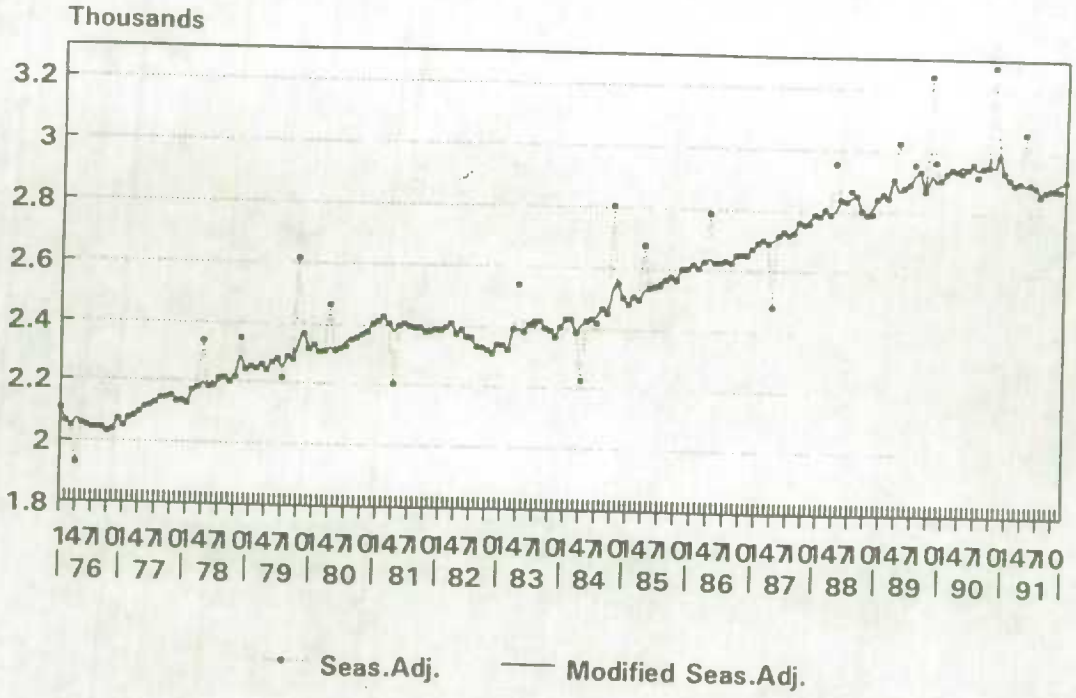

Figure 22.