EMBARGO: 11:30AM (CANBERRA TIME) DECEMBER 1999

Australian Bureau of Statistics

Working Paper No. 99/3

Adjusting for an Easter Proximity Effect

Caleb Leung, Craig H. McLaren, and Xichuan (Mark) Zhang

December 1999

This Working Paper Series is intended to make the results of current research within the Australian Bureau of Statistics available to other interested parties. The aim is to present accounts of developments and research or analysis of an experimental nature so as to encourage discussion and comment.
Views expressed in this paper are those of the author and do not necessarily represent those of the Australian Bureau of Statistics. Where quoted or used, they should be attributed clearly to the author.

ABS Catalogue No. 1351.0

Produced by the Australian Bureau of Statistics

Contents

INQUIRIES

- *For further information about the contents of this Working Paper, please contact the author:* Mark Zhang — *telephone:* Canberra 026252 5132 *or facsimile:* Canberra 0262528015 *or e-mail:*mark.zhang@abs.gov.au

- *For information on the* Working Papers in Econometrics and Applied Statistics, *please contact the Managing Editor:*
John De Maio — *telephone:* Canberra 02 6252 6804 *or facsimile:* Canberra 02 6252 5251 *or e-mail:* john.de.maio@abs.gov.au *or write* c/- Analytical Services Section (W21b), Australian Bureau of Statistics, PO Box 10, Belconnen, ACT 2616, Australia.

- *For information about other ABS statistics and services,* please refer to the back cover of this publication.

Abstract

The timing of Easter Sunday varies from one year to the next and can affect activity in time series data. To reveal the underlying movement of a time series, the date of Easter's occurrence and its impact on the time series have to be taken into account. New approaches are developed to model and remove the impact of Easter. The monthly Australian Total Retail Turnover series is used to illustrate the effectiveness of the modelling approaches.

Keywords: Easter proximity effect, regression-ARIMA, X11

1 Introduction

The observance of Australia's Easter holiday period, from Good Friday to Easter Monday, usually occurs completely within March or completely in April. On occasions, the Easter holiday period may start at the end of March and finish in the start of April. The effect of this movement of the Easter holiday period can directly impact on time series data aggregated on a regular calendar basis because of the variations in activity associated with Easter. For example, monthly or quarterly retail trade activity is likely to vary from its usual pattern around Easter. This effect is referred to as an Easter proximity effect.

As a calendar event, movement of the Easter holiday period needs to be taken into account in the seasonal adjustment process to avoid biased seasonally adjusted and trend estimates. Biased estimates can lead to misleading commentary and decisions by users and policy makers. To illustrate, a 2.3% increase in the seasonally adjusted Australian Total Retail Turnover series was reported for March 1999. Graphical evidence of the series (ABS, 1999a) suggested existence of an Easter Proximity effect of at most 1.5%. This suggests that the true underlying movement of the seasonally adjusted series was around 0.8%, not 2.3% as reported.

The Australian Bureau of Statistics (ABS) seasonally adjusts series using the commonly used seasonal adjustment procedure X11 (Shiskin *et. al*, 1967). The X11 procedure and its initial extension, X11ARIMA (Dagum, 1980), do not explicitly correct for an Easter proximity effect. X11ARIMA/88 (Dagum, 1988) and X12ARIMA (Findley *et. al*, 1998) do explicitly include a correction for the Easter proximity effect, however the correction is based on a North American Easter holiday period, not an Australian Easter holiday period. The correction methods used in X11ARIMA/88 and X12ARIMA are therefore not suitable for Australian time series data.

In this paper, we present approaches to correcting for an Easter proximity effect in the seasonal adjustment process that can be applied to Australian time series data. An approach is chosen from a regression method with an appropriate regressor. Our diagnostics indicate the proposed approaches are effective. We also present an approach that can be implemented into seasonal adjustment packages which do not use ARIMA extensions.

This paper is organised as follows. Section 2 describes the Easter proximity effect in detail. Regression based methods to estimating an Easter proximity effect correction are described in Section 3. Various regressors and their rationales are discussed in Section 4. Section 6 contains a test for the existence of an Easter proximity effect. The last section summarises our findings. The new approaches, together with that of the Statistics Canada and US Bureau of the Census X12ARIMA approaches, are evaluated in Section 7.

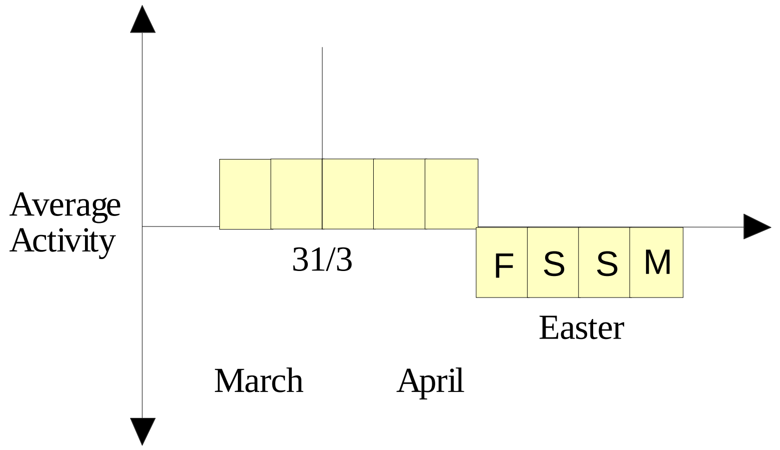
2 What is an Easter Proximity Effect?

Special dates in a year may have an impact on certain activities in a time series. For example, the Easter holiday period may have an effect on retail trade figures if monthly retail activity rises as Easter is approaching, activity falls when Easter arrives, and activity returns to normal once Easter has finished. The figures of monthly or quarterly aggregated data will be affected by the occurrence of the Easter holiday period close to or on the boundary of March and April. When time series data is affected by Easter starting late in March or early in April, the effect is referred to as an Easter proximity effect.

When the period before Easter and the Easter holiday period fall into the same month, high daily activity prior to the Easter holiday period may cancel the low activity observed during the Easter holiday period. No noticeable effect may exist on the month's aggregated data or its subsequent seasonally adjusted and trend estimates (see Section 6.2). However, when the period before Easter occurs in late March and the Easter holiday period falls in early April, the March aggregated data will not include the low activity during the Easter holiday period and so may be inflated. Similarly, the April aggregated data will not include the high activity prior to the Easter holiday period and so may be lower than expected. Without a correction, the seasonal adjustment process will produce biased estimates of the March and April seasonally adjusted estimates.

Figures 1 and 2 illustrate this concept. In figure 1, we assume that daily activity is constant. As Easter approaches, daily activity is increased by a constant daily amount. During the Easter holiday period, daily activity is reduced, also by a constant daily amount.

Figure 1: **Constant activity per day**



Note: The boundary between March and April is highlighted by 31/3. F=Good Friday, S=Saturday, S=Sunday, M=Easter Monday.

Figure 1 can also be thought of as having a monthly linear increase in activity before the Easter holiday period. That is, the constant activity prior to the Easter holiday period is cumulative for each day. In figure 2, daily activity is assumed to be linearly increasing, which is equivalent to a monthly quadratic increase in activity.

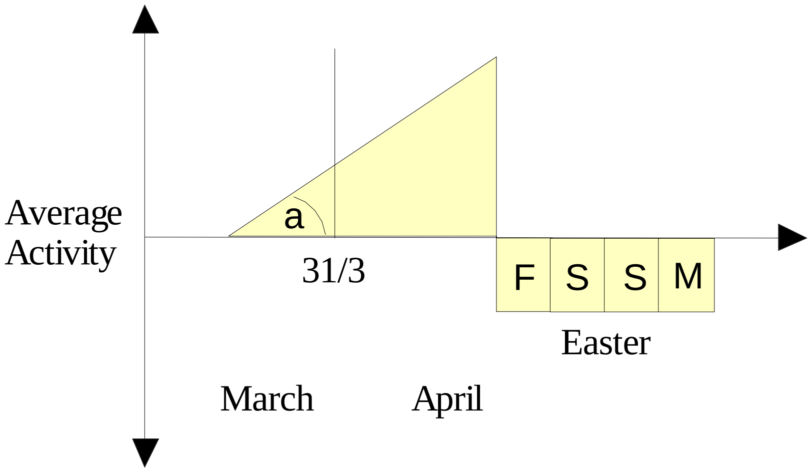


Figure 2: Linearly increasing activity per day, up to Easter
Note: The boundary between March and April is highlighted by 31/3. F=Good Friday, S=Saturday, S=Sunday, M=Easter Monday. The angle a represents the angle of the linear increase.

An Easter proximity effect is only likely to exist in years when Easter falls late in March or early in April. For the years 1962 to 2010, the years of interest are 1972, 1991 and 2002 when Easter occurs late in March and 1983, 1988, 1994, 1999 and 2010 when Easter starts early in April. In these years, an increase in activity in March and a decrease in April are expected resulting in seasonally adjusted and trend estimates for March and April that do not reflect the true underlying activity of a series. Most concern lies with misleading measures of growth at the current end of the series. It is also important to historically correct for an Easter proximity effect as the seasonally adjusted series may be used as input for economic modelling, or more generally, for studying relationships between variables over time.

Different countries observe different Easter holiday periods. Easter Sunday is the only observed Easter holiday in the United States (US) while Australia observes Easter from Good Friday through to Easter Monday. In the US, increased retail activity associated with Easter ends on the Saturday preceding Easter Sunday while in Australia, our results using the Australian Total Retail Turnover series suggest that increased retail activity associated with Easter ends on the Thursday before Good Friday.

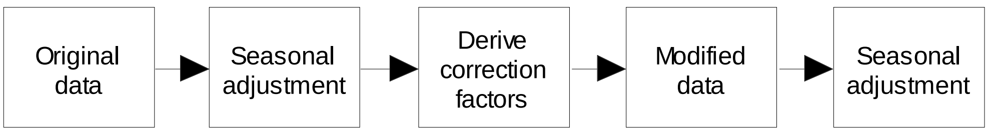
3 Regression Based Methods

Regression methods are widely used by national statistical organisations to estimate and remove the Easter proximity effect before producing seasonally adjusted and trend estimates. They can be classified into two different methods. The first method is a recursive method based on the use of irregular values obtained after performing an initial seasonal adjustment. It is expected that the Easter proximity effect resides in the irregular series and this series is used to derive correction factors. The original data are modified using these derived factors before another seasonal adjustment run is undertaken. Hence, the Easter proximity effect correction is made after the first seasonal adjustment run of X11. This method is referred to as the D13 method. The second method is a simultaneous estimate method that makes a correction to the original data before any seasonal adjustment is undertaken. This method uses a regression model with an ARIMA error term to derive the correction factors and adjust the original data. It is referred to as the regression-ARIMA method. More details are given in Sections 3.1 and 3.2.

3.1 D13 Method

The D13 method uses the irregular values of March and April obtained from X11 (Table D13 from X11 output) to identify the Easter proximity effect. If an Easter proximity effect exists, the irregulars in the affected years will deviate from one, the neutral line of the irregulars. Therefore, the deviations can be used to estimate the Easter proximity effect. A drawback of this method is that the seasonally adjusted estimates may already be distorted by a Easter proximity effect so the irregulars derived from the seasonally adjusted estimates would also be distorted by the seasonal adjustment process.

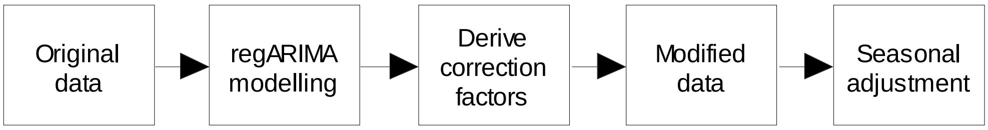
Diagrammatically, the D13 method can be illustrated as follows:



3.2 Regression-ARIMA Method

The regression-ARIMA method estimates the Easter proximity effect before performing a seasonal adjustment. It does not use the irregulars from a seasonal adjustment to make a correction. This avoids the drawback of the D13 method of using irregulars that are already contaminated by the seasonal adjustment process. The Easter proximity effect can be captured by regressing the original data on a regressor associated with Easter plus an ARIMA model (see Section 4 for more details) which models other sources of variations. This method is used by Bell and Hillmer (1983) in a regression ARIMA framework. This method is also used in X12ARIMA and TRAMO/SEATS (Gomez and Maravall, 1992).

Diagrammatically, the regression-ARIMA method can be illustrated as follows:



Mathematically, the regression-ARIMA method is as follows. A general multiplicative seasonal ARIMA model for a time series, z_t (see for example, Box and Jenkins (1976)) can be written as

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D z_t = \theta(B)\Theta(B^s)a_t \qquad (1)$$

B is the back shift operator, s is the seasonal period, $\phi(B)=(1-\phi_1 B-\ldots-\phi_p B^p)$ is the non-seasonal autoregressive component, $\Phi(B)=(1-\Phi_1 B^s-\ldots-\Phi_P B^{Ps})$ is the seasonal autoregressive component, $\theta(B)=(1-\theta_1 B-\ldots-\theta_q B^q)$ is the non-seasonal moving average component, $\Theta(B)=(1-\Theta_1 B^s-\ldots-\Theta_Q B^{Qs})$ is the seasonal moving average component, and a_t is independent and identically normally distributed with mean 0 and variance σ^2.

Additionally, we assume that a linear regression for a time series y_t can be written as

$$y_t = \sum_i \beta_i x_{it} + z_t \qquad (2)$$

y_t is a dependent time series, x_{it} are regression variables, β_i are regression parameters, and z_t follows the ARIMA model in (1).

The regression-ARIMA model is then found by combining (1) and (2) in a single equation

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D\left(y_t-\sum_i \beta_i x_{it}\right) = \theta(B)\Theta(B^s)a_t$$

4 Easter Regressors

A regressor is a predictor variable that explains the variation of a response variable in a regression framework. A regression model can estimate the effect of the regressor on the variation of the response variable. The design of a regressor to measure the Easter effect can range from a simple indicator variable to a more sophisticated one eg. an exponential function. For example, within SEASABS (ABS, 1999b) a simple indicator variable is defined as a regressor E_{reg} as follows:

$$E_{reg} = \begin{cases} 1 \text{ if Easter is wholly in March} \\ 1/2 \text{ if Good Friday is in March and Easter Monday is in April} \\ 0 \text{ if Easter is wholly in April} \end{cases}$$

This regressor can estimate the Easter holiday effect. It will not estimate the variation in activity prior to Easter.

In the following sections, we assume that Easter has an effect of increasing activity for w days before Easter. The period of w days can be thought of as a window. The number of the w days, n , that fall into March and/or April can be used to create regressors.

4.1 TRAMO

The TRAMO regressor reflects how March and April share the w days. For months other than March and April, the regressor is zero. The TRAMO regressor has values:

$$E_{reg} = \begin{cases} n/w \text{ in March} \\ 1 - n/w \text{ in April} \\ 0 \text{ otherwise} \end{cases}$$

4.2 US Bureau of the Census

The US Bureau of the Census (USBC) uses a similar regressor to the TRAMO regressor. Its values depend on the monthly proportions of the span of the w days before Easter which fall in February, March and April, after subtracting the monthly means of the proportions for February, March and April respectively. The monthly means are estimated by calculating the sample means of the monthly proportions over many calendar years. The rationale is that the Easter proximity effect should be balanced between the affected months of February, March and April. By subtracting the monthly mean, the regressor is symmetric for the pair of months February and March as well as March and April. The USBC regressor has values:

$$E_{reg} = \begin{cases} (\text{number of the } w \text{ days before Easter falling in month } t)/w - E(\text{month } t) \\ 0 \text{ except in February, March and April} \end{cases}$$

The monthly means are the long run averages computed over 38,000 years (Findley *et. al*, 1998), although it is not stated what the start and end dates of the long span of years actually are. The X12ARIMA reference manual computes means using years between 1900 and 2100 inclusive.

4.3 Statistics Canada

The Statistics Canada (StatCan) regressor is a simplified version of the USBC regressor. It keeps the symmetric property for the regressor but avoids estimating the mean of each month. This is achieved by forcing April to have a negative effect of the value calculated for March. The StatCan regressor has values:

$$E_{reg} = \begin{array}{l} n/w \text{ in March} \\ -n/w \text{ in April} \\ 0 \text{ otherwise} \end{array}$$

Both the USBC and StatCan regressors are built into the X12ARIMA seasonal adjustment package.

All three regressors use the same idea but have a very different philosophy in relation to the seasonal factors. For the TRAMO regressor, the regression model will remove all of the Easter proximity effect, in other words, the seasonal factors for March and April do not contain any Easter proximity effect. For the USBC regressor, the regression model removes part of Easter proximity effect which deviates from the average Easter proximity effect. For the StatCan regressor, all of the Easter proximity effect is put in to the April seasonal factor with the assumption that the whole period days w before Easter, falls into April. An adjustment is made if part (or whole) of the w days falls into March. Although the seasonal factors may be different as a result of the different adjustments based on the three different regressors, the seasonally adjusted estimates should be the same because both the estimated Easter proximity effect and the corresponding seasonal factors are removed. ie. the net adjustment for the Easter proximity effect and seasonal factors should be the same for the three different regressors used to produce seasonally adjusted estimates.

4.4 Alternative Regressors

In this section, new regressors are proposed that reflect the characteristics of the Australian Easter holiday pattern. These include the use of a quadratic regressor as well as a regressor which accounts for a decrease in activity during the Easter holiday period.

Current investigations on the Australian Total Retail Turnover series have shown that Easter has an approximate seven-day effect on retail activity. For this study, we use a window of length seven in all calculations. The window length for the Easter holiday period is taken to be four as this is the length of the holiday in Australia. However, the window length of the Easter holiday period can be changed to reflect the holiday period in other countries.

We also have to determine when the pre-Easter rising period ends. Based on the analysis of the Australian Total Retail Turnover series, we found that the rising period ends one day before the start of the traditional Easter holiday, ie. Thursday. In comparison, US and Canada end the rising period on Saturday.

Pre-Easter regressor

The TRAMO, USBC and StatCan regressors all assume a constant extra activity per day in the w days before Easter (i.e. assume constant increase in activity over the window w). In practice, this assumption may not be appropriate. We introduce a regressor that assumes a linear increase in extra daily activity.

For a linear increase in daily activity in the pre-Easter period as shown in Figure 2, consider the following. Let the linear increase have slope $\tan(a)$. The extra activity is equal to the triangle area within the w days. This gives $w^2 \tan(a)/2$ being the total extra activity. Activity belonging to March is $n^2 \tan(a)/2$. A regressor is constructed using the proportion of activity in a month from the extra activity. For example, the proportion in March is $(n^2 \tan(a)/2)/(w^2 \tan(a)/2) = (n/w)^2$.

Therefore the TRAMO style regressor will have values:

$$E_{reg} = \begin{array}{l} (n/w)^2 \text{ in March} \\ 1-(n/w)^2 \text{ in April} \\ 0 \text{ otherwise} \end{array}$$

The USBC style regressor will take values:

$$E_{reg} = \begin{array}{l} [(\text{number of the } w \text{ days before Easter falling in month } t)/w]^2 - E(\text{month } t) \\ 0 \text{ except in February, March and April} \end{array}$$

The StatCan style regressor will take values:

$$E_{reg} = \begin{array}{l} (n/w)^2 \text{ in March} \\ -(n/w)^2 \text{ in April} \\ 0 \text{ otherwise} \end{array}$$

These regressors are referred to as quadratic regressors for their quadratic power. The original TRAMO, USBC and StatCan regressors are referred to as linear regressors.

During Easter regressor

Another approach is to model separately the activity prior to Easter and the activity during the Easter holiday period. An additional window is added to reflect activity during the Easter holiday period. Construction of this additional regressor assumes the activity during the Easter holiday period, Good Friday to Easter Monday, is constant. This is similar to the TRAMO, USBC and StatCan regressor assumptions.

Linear regressors take the following form. The TRAMO style regressor has values:

$$_dE_{reg} = \begin{array}{l} m/4 \text{ in March} \\ 1-m/4 \text{ in April} \\ 0 \text{ otherwise} \end{array}$$

where 4 is used as this is the length of the holiday period, and m is the number of the 4 days falling in March.

The USBC style regressor will have values:

$$ {}_{d}E_{reg} = \begin{cases} (\text{number of the 4 days before Easter falling in month } t)/4 - E(\text{month } t) \\ 0 \text{ except in February, March and April} \end{cases} $$

The StatCan style regressor will have values:

$$ {}_{d}E_{reg} = \begin{cases} m/4 \text{ in March} \\ -m/4 \text{ in April} \\ 0 \text{ otherwise} \end{cases} $$

Quadratic regressors with a linear decrease in daily activity during the Easter holiday period can be constructed in a similar fashion.

5 Example: Australian Total Retail Turnover

An exploratory investigation into the likely magnitude of an Easter proximity effect in Australian Total Retail Turnover and the component series was conducted by the Australian Bureau of Statistics in April 1999 (ABS, 1999a). It was found that the magnitude of the Easter proximity effect for March and April seasonally adjusted estimates was no greater than 1.5%. To illustrate this a concurrent seasonal adjustment was performed using SEASABS on the original Australia Total Retail Turnover series for the span April 1962 to April 1999. Figure 3 shows how the irregulars from this concurrent seasonal adjustment for both March and April are distributed as the date of Easter Sunday changes. The labels indicate which year the irregular was observed.

From Figure 3, when Easter Sunday falls on or after 5th of April there is no graphical evidence of an Easter proximity effect. When Easter Sunday has started on the 3rd and 4th of April, the irregulars for March and April are consistently above and below one respectively. Detecting a definite cutoff for an Easter proximity effect is difficult as there are only a few observations of Easter Sunday in the first week of April.

The original Australian Total Retail Turnover series has recorded monthly data since April 1962. When an Easter proximity effect exists, it may evolve over the years. To minimise any risk of using a data span that may be contaminated by an evolving Easter Proximity Effect, a data span that displays homogenous seasonality is desirable. We can use seasonal factors, produced without an Easter proximity effect correction, as an indicator of possible changes in the Easter proximity effect. Figure 4 shows that the series before 1980 has a different seasonal pattern to the remaining years in the series. After 1980, the March and April seasonal factors are almost parallel. Prior to 1980, this was not the case. This may be due to, in part, retail purchasing patterns changing over recent years with the deregulation of shopping hours.

As a result of this investigation, a truncated series from 1980 onwards is used in the evaluation of the regression methods (note: different Easter proximity effect estimates for the different data spans can be found in Appendix 10.1). Figure 5 shows for the truncated data span how the irregulars from a concurrent seasonal adjustment for both March and April are distributed about one as the date of Easter Sunday changes. The labels indicate which year the irregular was observed.

Figure 3: Easter Proximity chart for original Australia Total Retail Turnover - full span: April 1962 to April 1999 inclusive

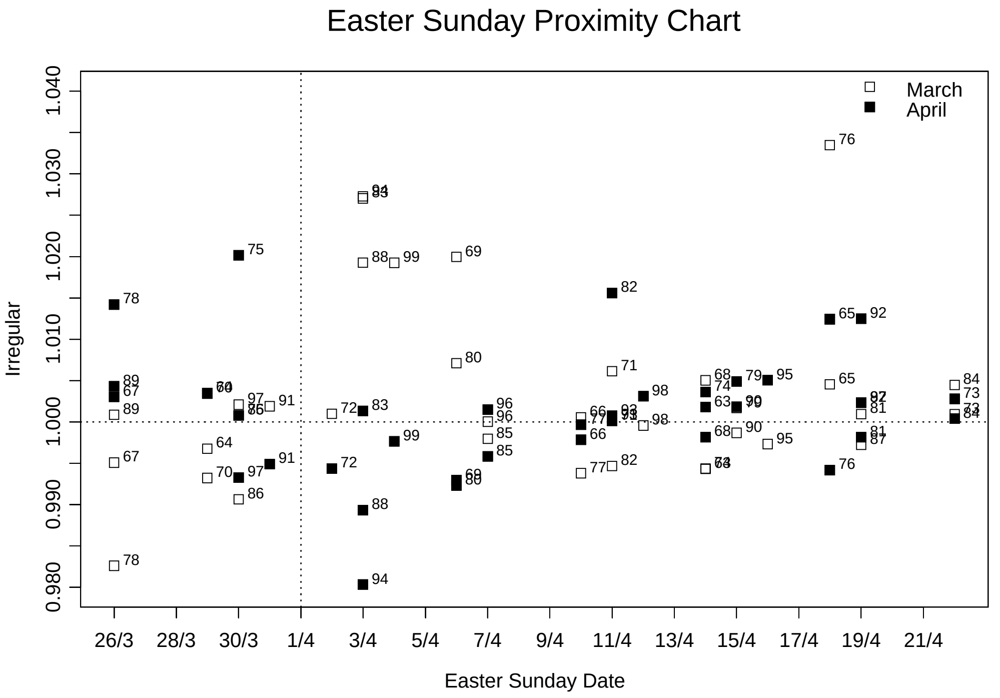


Figure 4: Estimated Seasonal Factors for March and April for Australia Total Retail Turnover - full span: April 1962 to April 1999 inclusive

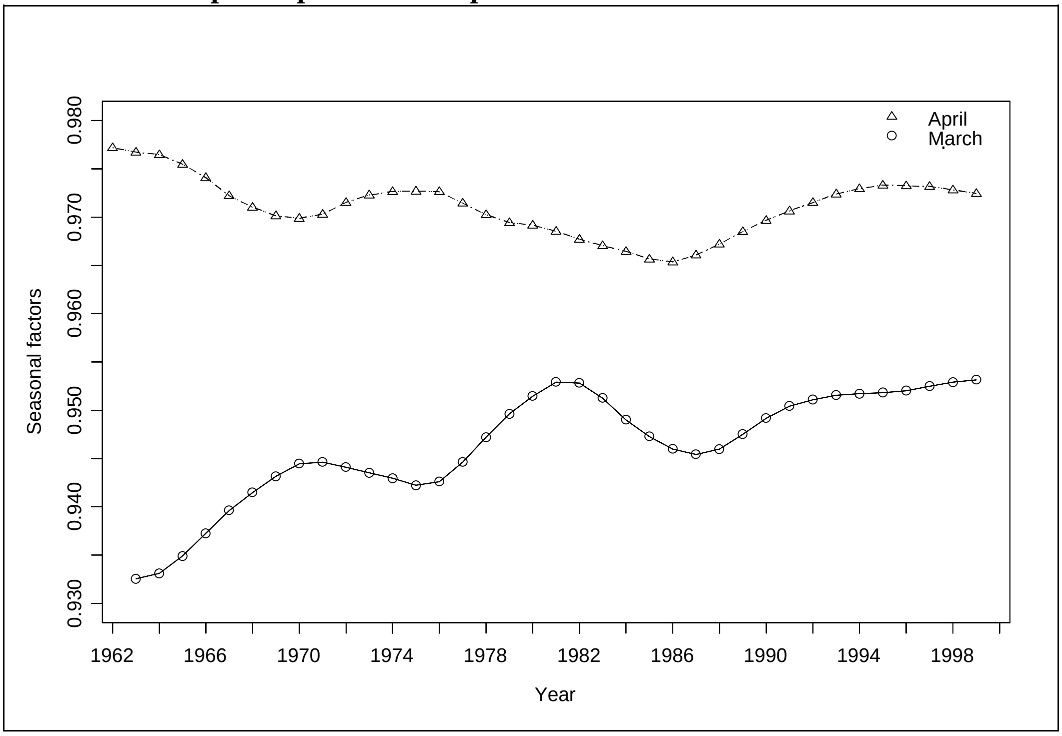
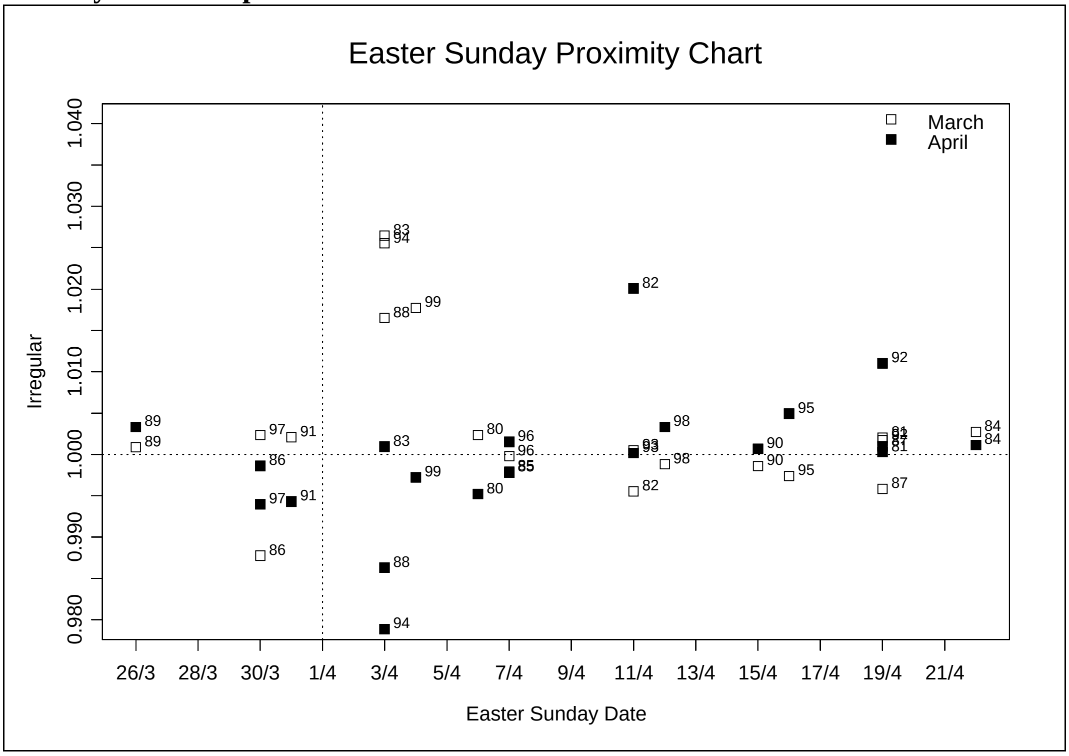


Figure 5: Easter Proximity chart for original Australia Total Retail Turnover - truncated span: January 1980 to April 1999 inclusive



6 Hypothesis testing

6.1 Testing for the existence of an Easter proximity effect

Before correcting for an Easter proximity effect it is important to determine if such an effect exists in the series.

Let

P_e = coefficient parameter for pre-Easter period, estimated by $\hat{P}_e$

P_d = coefficient parameter for during Easter holiday period, estimate by $\hat{P}_d$

A simple test for the existence of an Easter proximity effect is then to use a *t*-test to determine if the coefficient parameter for the pre-Easter period is zero i.e. our null hypothesis is $H_o : P_e = 0$.

A specific Easter proximity effect could be defined as an increase in activity for the pre-Easter period along with a decrease in activity for the Easter holiday period. This situation is illustrated in Figures 1 and 2. A test for this specific Easter proximity effect could involve testing the null hypothesis $H_o : P_e > 0$ and $P_d < 0$.

Table 1 gives the results of a simple Easter proximity test applied to the Australia Total Retail Turnover series using four different approaches for an Easter regressor. Each approach detects a significant Easter proximity effect in the data.

Table 1: Hypothesis test for the Australian Total Retail Turnover series

Approach	Parameter P_e estimate for pre-Easter period	Hypothesis $H_o : P_e = 0$ t-test statistic (p-value)
Linear-Linear Regressor	0.0181	6.09 (<0.001)
Quadratic-Linear Regressor	0.0196	6.35 (<0.001)
D13 Linear-Linear Regressor	0.0139	6.53 (<0.001)
D13 Quadratic-Linear Regressor	0.016	6.98 (<0.001)

6.2 Testing the net effect of the Easter proximity effect

An assumption often made is that when the whole high activity period before Easter and the Easter holiday period fall into the same month, the net effect of the Easter proximity effect will be zero. The estimated coefficients from a double regressor approach can be used to test whether high activity prior to the Easter holiday period negates low activity during the Easter holiday period within that month.

For example, when both periods are in March, the net effect for March is $P_e \times 1 + P_d \times 1$ and that for April is $P_e \times (-1) + P_d \times (-1)$. When both periods are in April, both net effects are $P_e \times 0 + P_d \times 0 = 0$.

Testing whether the increasing and decreasing effects cancel reduces to the case that both periods fall in March. This is equivalent to testing the null hypothesis $H_o : P_e + P_d = 0$ versus the alternative hypothesis $H_1 : P_e + P_d \neq 0$.

Under the null hypothesis, the test statistic

$$(\widehat{P}_e + \widehat{P}_d)/\sqrt{var(\widehat{P}_e) + var(\widehat{P}_d) + 2cov(\widehat{P}_e, \widehat{P}_d)}$$

is assumed to follow a Normal distribution $N(0, 1)$.

Table 2 shows the results of this test using the four approaches applied to the Australian Total Retail Turnover. (Note that the test for the D13 linear-linear regressor approach is based on one iterative regression). We cannot reject the null hypothesis for the Australian Total Retail Turnover series. The results indicate that when the pre-Easter period and Easter holiday period fall in the same month then the high and low activity balance each other.

Table 2: Hypothesis test for the Australian Total Retail Turnover series

Approach	Parameter P_e estimate for pre-Easter period	Parameter P_c estimate for during Easter holiday period	Hypothesis $H_o : P_e + P_d = 0$ test statistic (p-Value)
Linear-Linear Regressor	0.0181	-0.0156	0.7960 (0.4260)
Quadratic-Linear Regressor	0.0196	-0.0176	0.6166 (0.5375)
D13 Linear-Linear Regressor	0.0139	-0.0145	-0.2564 (0.7976)
D13 Quadratic-Linear Regressor	0.016	-0.0168	-0.3066 (0.7592)

Further investigations into component series of the retail trade series show that this finding does not always hold. Some series do not have their net effects balanced. The Australian food retail series is one of the counter examples. Table 3 lists the hypothesis test results for the four approaches for this series. The results confirm that the null hypothesis cannot be accepted at a probability level of 0.01. That is, the high and low activity does not balance out when the pre-Easter period and Easter holiday period fall within the same month.

Table 3: Hypothesis test for the Australian food retail series

Approach	Parameter P_e estimate for pre-Easter period	Hypothesis $H_o : P_e = 0$ test statistic (p-value)	Parameter P_d estimate for Easter holiday period	Hypothesis $H_o : P_d = 0$ test statistic (p-value)	Hypothesis $H_o : P_e + P_d = 0$ test statistic (p-value)
Linear-Linear Regressor	0.023	8.13 (<0.001)	-0.0075	-2.12 (0.0340)	5.1293 (<0.001)
Quadratic-Linear Regressor	0.024	8.53 (<0.001)	-0.0094	-2.65 (0.0080)	5.0161 (<0.001)
D13 Linear-Linear Regressor	0.0157	7.04 (<0.001)	-0.0078	-2.32 (0.0202)	3.2754 (0.0011)
D13 Quadratic-Linear Regressor	0.0179	7.66 (<0.001)	-0.0101	-2.96 (0.0031)	3.2847 (0.0010)

7 Evaluation

The evaluation consisted of firstly estimating and removing the Easter proximity effect using a specified regression approach, then investigating the irregular series after application of the SEASABS seasonal adjustment package. Double regression methods have been evaluated because of their suitability to estimate the Australian Easter holiday pattern in the retail series. Double regression methods enable both the pre-Easter activity and activity during Easter to be modelled separately.

Quadratic and linear regressors have been evaluated for their suitability for estimating an Easter proximity effect in the pre-Easter period. For the activity during the Easter holiday period we assume constant daily activity and so only evaluate linear regressors. This reduces the scope of our evaluation to considering combinations of quadratic and linear regressors, and linear and linear regressors.

The StatCan style regressors have been evaluated in this study. Our analysis has shown that there is no noticeable difference between the performance of the USBC and StatCan style regressors. The StatCan style regressors are preferred for their ease of derivation. The TRAMO style regressors are not used as they are not balanced across March and April.

Both the modified D13 and regression-ARIMA methods are evaluated to enable comparison between the two regression methods. Table 4 summarises the final four approaches evaluated.

Table 4: New approaches created by combinations of regression methods and regressors

Method	StatCan style regressor for pre-Easter period	StatCan style regressor for during Easter holiday period	Double regression approach
Regression-ARIMA	Linear regressor	Linear regressor	Linear-linear regressor
Regression-ARIMA	Quadratic regressor	Linear regressor	Quadratic-linear regressor
D13	Linear regressor	Linear regressor	D13 linear-linear regressor
D13	Quadratic regressor	Linear regressor	D13 quadratic-linear regressor

Statistical measures are used to quantify the effectiveness of the different approaches. Four statistical measures are used to evaluate the performance of each approach. Table 5 lists the statistical measures and rationales. The assessment for each model under each statistical measure is discussed in Section 7.1 to 7.4.

Table 5: Statistical measurement for model assessment

Statistics	Measure
Average sum of absolute values from D13	Robust measure of closeness of irregular to the neutral line
Average of uncorrected sum of squares from D13	Sensitive measure of closeness of the irregular to the neutral line
ANOVA	Any systematic pattern existing in the irregular can be explained by grouping
Autocorrelation	Any "time" lagged pattern existing in the irregular

7.1 Average Sum of Absolute Values

The average sum of absolute values give a robust overall assessment of the closeness of the irregulars to one, the expected value of the D13 irregulars. Figure 6 shows that all approaches perform significantly better than the X11 run without an Easter proximity correction. The approaches perform almost identically in April. The quadratic-linear regressor approach performs better than both double linear regressor approaches.

7.2 Mean of Uncorrected Sum of Squares

The average uncorrected sum of squares give a sensitive overall measure of the closeness of the irregulars to one, the expected value of the D13 irregulars. This measure is more sensitive to extreme irregulars. Figure 7 shows that all new approaches perform significantly better than the X11 run without the Easter proximity correction. The quadratic-linear regressor approaches using regression-ARIMA and D13 methods have a better performance in March than the double linear regressor approaches, but this is not the case in April. Based on this statistical measure, the use of double linear regressor is preferred as it has a similar performance in both months.

7.3 Analysis of variance

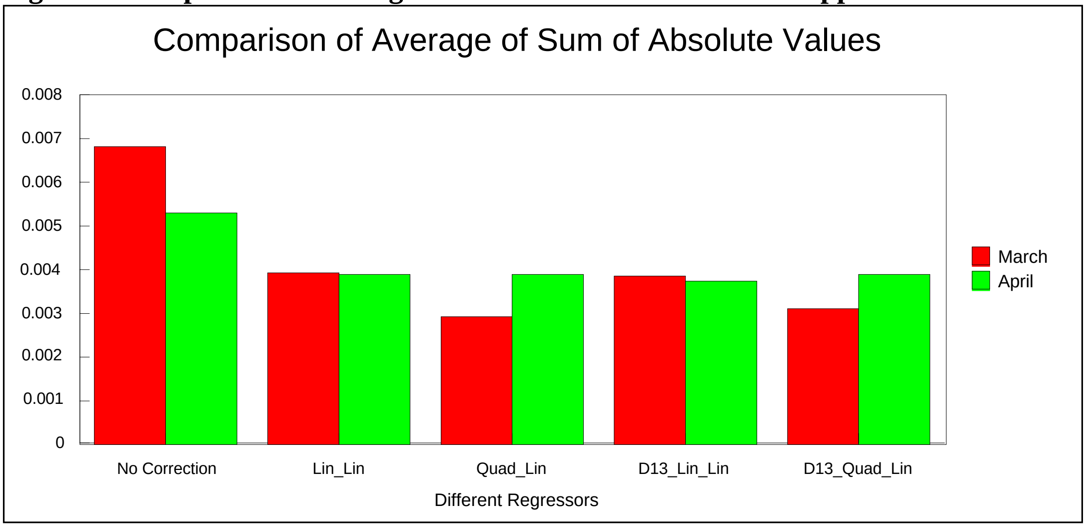
An analysis of variance (ANOVA) is used to test the null hypothesis of no pattern existing in the irregulars after the Easter proximity effect is removed. The rejection of the null hypothesis implies that a pattern exists in the irregulars which may depend on the date of Easter. An ANOVA can only be implemented by grouping calendar dates as the number of observations are very limited on each Easter date.

After applying different correction approaches, an analysis of variance on the irregular values is used to assess their effectiveness. The number of days Good Friday is away from the 31st of March is ordered from the smallest to the largest. The range of these values can be observed in Figure 5. Four groups are formed by grouping the number of days Good Friday is away from 31st March from -7 to 0 days in Group 1, 1 to 7 days in Group 2, 8 to 14 days in Group 3 and 15 to 22 days in Group 4. This resulted in 4, 7, 5 and 4 observations respectively. Table 6 gives the F-statistics and probabilities for detecting any differences between the four groups.

If there is no evidence of an Easter proximity effect, then we would expect no structure in the irregulars as a function of the days from the 31st of March. As a consequence, we would then expect no significant differences between the four groups. Table 6 shows that the no correction approach (X11 approach) clearly has significant differences between the four groups for both March and April. This is also reflected by the correlation coefficient which suggests that for March 46% of the observed structure is explained by the groups. Similarly, April has a correlation coefficient of 40%.

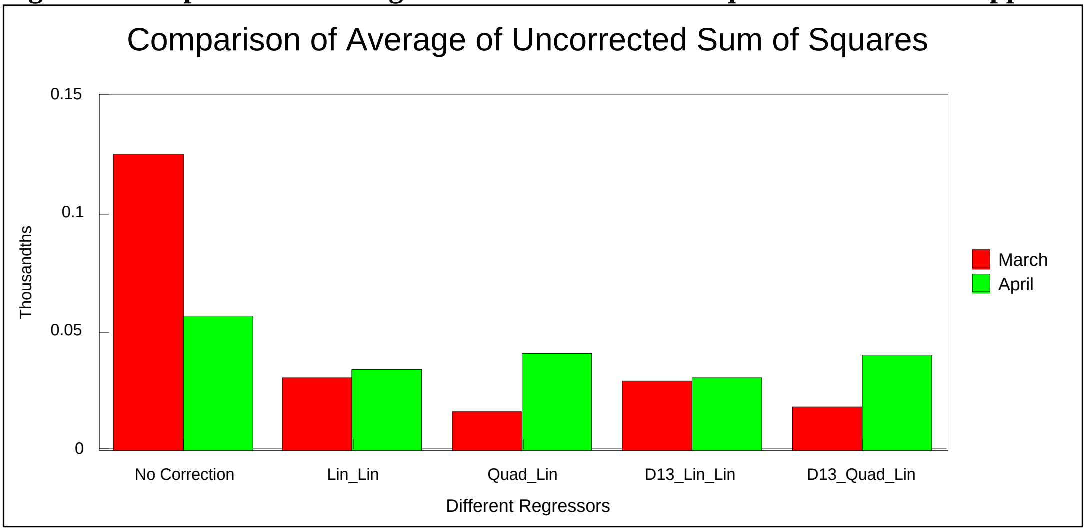
All new approaches performed well for both March and April. The modified D13 approaches perform appreciably better for March than April (March p values are appreciably higher than Aprils) while the regression-ARIMA approaches perform slightly better for April than March (March p values are slightly lower than Aprils).

Figure 6. Comparison of average of absolute value for different approaches



Note: No Correction = X11 approach, Lin_Lin = linear-linear regressor approach, Quad_Lin = quadratic-linear regressor approach, D13_Lin_Lin = D13 linear-linear regressor approach, D13_Quad_Lin = D13 quadratic-linear regressor approach.

Figure 7. Comparison of average of uncorrected sum of squares for different approaches



Note: No Correction = X11 approach, Lin_Lin = linear-linear regressor approach, Quad_Lin = quadratic-linear regressor approach, D13_Lin_Lin = D13 linear-linear regressor approach, D13_Quad_Lin = D13 quadratic-linear regressor approach.

Table 6: F-Statistics for differences between four groups of dates for Good Friday

F-Statistic (probability)		
Approach	March	April
No correction	4.51 (0.0178)	3.53 (0.0389)
Linear-Linear Regressor	0.40 (0.7530)	0.57 (0.6440)
Quadratic-Linear Regressor	0.65 (0.5963)	0.57 (0.6457)
D13 Linear-Linear Regressor	0.21 (0.8866)	0.96 (0.4341)
D13 Quadratic-Linear Regressor	0.51 (0.6834)	0.92 (0.4535)

Note: Number of days Good Friday occurs from the 31st of March where Group 1=[-7,0] days, Group 2=[1,7] days, Group 3=[8,14] days and Group 4=[15,22] days. The respective probabilities at the 5% significance level are given in brackets. Series span: January 1980 to April 1999 inclusive and window length=7.

7.4 Autocorrelation

Since grouping causes a loss of information, an alternative way for testing for a "time" related pattern in the irregulars is to use the autocorrelation function. Autocorrelation indicates whether there is any serial dependence in the irregulars. The rejection of the null hypotheses will indicate that the irregulars are not likely to be a pure random noise.

The aim is to test for any systematic patterns in the D13 irregular series in relation to the boundary between March and April. Since Easter is a moving holiday, the D13 irregular values are an unequally spaced series and can have more than one observation at a particular time point. The D13 irregular values plotted against Good Friday is not a usual representation of a time series as there are multiple observations at different time points. Therefore, the normal procedure for testing autocorrelation of a series is not applicable in our situation.

The variogram (Diggle, 1990) can handle multiple series and estimate the autocorrelation function provided that the multiple series are stationary. In our case, pseudo multiple time series are created by a two step process. Firstly, the original irregular observation series is treated as an equally spaced series by removing the gaps with no observation. We use the time sequence $\{j : 1, 2, \ldots\}$ to replace the actual date sequence and denote the equally spaced series as $\{G_{qj} : q^{th} \text{ observation at time } j\}$. Secondly, a series i is formed by selecting one observation from each time $\{j : 1, 2, \ldots\}$, ie. a single observation is treated as a repeated observation. The selection process continues until all combinations are found. The number of the series created is the number of all combinations. Let the pseudo multiple series be represented by $\{g_{ij}$: the observation from series i for time $j\}$. We set a null hypothesis that the pseudo multiple series are white noise, ie. at least stationary. The variograms of the pseudo multiple series can then be used to estimate the autocorrelations of the D13 irregular series related to March/April. If the estimated lagged autocorrelations are significantly different from zero, we reject the null hypothesis and conclude that the D13 irregular series does have certain time related patterns.

The autocorrelation function at lag k is given by

$$r(k) = 1 - v(k)/v$$

Where $v(k)$is the mean variogram at lag k and v is the variance of the stationary process $\{G_{qj}\}$. Since all combinations are generated from a single series, only the original D13 irregular values are used for variance estimation. The variance v can be calculated as

$$v = \sum_{l} \sum_{j} (G_{qj} - G)^2 / (\text{total number of } G_{qj})$$

where G is the mean of G_{qj}. The mean variogram $v(k)$ can be evaluated by

$$v(k) = \sum_{i} \sum_{j} u_{ij}(k) / (\text{total number of } u_{ij}(k))$$

where

$$u_{ij}(k) = 0.5(g_{ij} - g_{ih})^2$$

$$h = j + k$$

and $u_{ij}(k)$ is the jth variogram from series i with lag k.

The confidence intervals are given by $(-1.96/\sqrt{(\text{total number of } y_{ij})-k}\,,\,1.96/\sqrt{(\text{total number of } y_{ij})-k}\,)$. Note that the confidence limits are calculated using (total number of $y_{ij}-k$) rather than (total number of y_{ij}) to adjust for the small amount of data available in the series $\{y_{ij}\}$. Each autocorrelation has its own confidence interval. If any autocorrelation lies outside the confidence limits, this is an indication of serial dependence in the D13 irregulars.

The confidence limits only test for the individual autocorrelation at a given time point. To test the autocorrelations as a whole, the portmanteau test (Ljung and Box, 1978) for white noise can be employed. This test statistic is given by

$$Q(m) = n(n+2)\sum_{k=1}^{m}(n-k)^{-1}r(k)^2$$

where n is the number of observation dates. The statistic $Q(m)$ follows a chi-square $\chi^2(m)$ distribution with m degrees of freedom. If $Q(m)$ exceeds the critical value of $\chi^2(m)$, then the white noise assumption for the irregulars is not appropriate. This test would be more reliable when n is much greater than m.

Tables 7 and 8 give the $Q(m)$ statistics for lags of $m = 1, ..., 5$. Estimates of $Q(m)$ for higher lags may be unreliable as n is only 13. From these tables, the Q-statistics reveal that serial dependence exists in the irregulars obtained from X11 up to lag 2 for March and lag 4 for April but not in the other four approaches. This implies that by just using X11 with no correction for an Easter proximity effect, a time related pattern exists in the irregulars.

Table 7: Correlogram calculations using the portmanteau test for March

Correlogram calculations for March - Q-statistics						
m	No correction	Linear-Linear Regressor	Quadratic-Linear Regressor	D13 Linear-Linear Regressor	D13 Quadratic-Linear Regressor	Critical Value: $\chi^2(m)$
1	5.941	2.237	1.247	2.409	1.934	3.841
2	6.134	2.443	1.963	2.586	2.49	5.991
3	6.284	2.565	3.558	2.714	3.459	7.815
4	6.76	5.436	8.76	5.843	8.064	9.488
5	8.664	8.042	13.019	8.513	12.121	11.07

Note: Each value is compared to a $\chi^2(m)$. Series span: January 1980 to April 1999 inclusive and window length=7.

Table 8: Correlogram calculations using the portmanteau test for April

Correlogram calculations for April - Q-statistics						
m	No correction	Linear-Linear Regressor	Quadratic-Linear Regressor	D13 Linear-Linear Regressor	D13 Quadratic-Linear Regressor	Critical Value: $\chi^2(m)$
1	4.097	1.265	1.124	1.265	1.124	3.841
2	6.765	1.713	1.358	2.003	1.528	5.991
3	8.5	3.342	3.418	4.324	3.626	7.815
4	9.616	5.687	8.083	6.933	7.67	9.488
5	9.629	5.728	8.405	6.935	7.858	11.07

Note: Each value is compared to a $\chi^2(m)$. Series span: January 1980 to April 1999 inclusive and window length=7.

8 Conclusion

If the existence of the Easter proximity effect is significant, it is clear that it should be taken into account as part of the seasonal adjustment process.

All the approaches evaluated gave substantial improvements in the seasonally adjusted and trend estimates when compared with the practice of not adjusting for an Easter proximity effect. The new approaches presented provided additional gains over those approaches currently used within X12ARIMA by including an additional regressor to capture the unique characteristics of Australian Easter holiday period.

Based on the performance measures in Section 7, the two regressors linear-linear regressor and quadratic-linear regressor perform equally well. The statistical measures used do not conclusively show that one regressor is better than the other because of limited observations and potential outliers. However, if we exclude the potential outliers, the quadratic-linear regressor appears to perform better than the linear-linear regressor. Easter proximity charts for the quadratic-linear regressor are in Appendix 10.2. These can be compared with Figure 5.

For the pre-Easter effect, we also found that the parameter test statistics from the regression-ARIMA method are more significant than those from the D13 method when the same regressor is used, and the parameter test statistics from the quadratic-linear regressor are more significant than those from the linear-linear regressor.

These findings indicate that the best choice for modelling an Easter proximity effect in the Australian Total Retail Turnover series is a combination of regression-ARIMA and the quadratic-linear regressor. This approach is preferable as it avoids leakages between seasonal factors due to the Easter proximity effect.

Although the D13 method was not as effective as the regression-ARIMA method, it provided an adequate correction for the Easter proximity effect and was only slightly worse than the best approach. This approach can be implemented within seasonal adjustment packages which do not have ARIMA facilities. For the ABS, which does not currently use ARIMA methods for seasonal adjustment, the best choice is the combination of the D13 method with the quadratic-linear regressor.

This evaluation created an input series by dividing the "final" moving trading day factors into the original series. In practice, moving trading days factors are estimated concurrently.Results not presented here show that all four approaches still adequately correct the Easter proximity effect based on all the statistical assessment measures.

We also investigated the application of approaches to different component series of the Australian Retail Trade. For example, Department Stores, Household Good Retailing, Recreational Good Retailing, Food Retailing, Clothing and Soft Good Retailing. Results suggest that for those series where an Easter proximity effect existed the approaches are able to adequately correct for the effect.

The methodology underlying the new approaches could also be applied to other calendar related events. For example, the approaches may be able to calculate correction factors for Fathers Day proximity effect which may occur in some series.

9 References

Australian Bureau of Statistics (1999a). Easter Holiday Effects in Retail Turnover, Australian Economic Indicators, cat. 1350.0, May 1999, Canberra, Australia.

Australian Bureau of Statistics (1999b). SEASABS v2.0. Australian Bureau of Statistics: Knowledge based version of X11.

Bell, W.R., and Hillmer, S.C. (1983), Modeling Time Series with Calender Variation, *Journal of the American Statistical Association, 78, 526-534.*

Box, G.E.P. and Jenkins G.M. (1976). Time Series Analysis: Forecasting and Control, San Francisco: Holden Day.

Dagum, E.B. (1980). The X11ARIMA Seasonal Adjustment Method. Ottawa: Statistics Canada, cat. 12-564E.

Dagum, E.B. (1988). The X11ARIMA/88 Seasonal Adjustment Method. Ottawa: Statistics Canada, cat. K1A 0T6.

Diggle, P.J. (1990). Time Series: A Biostatistical Introduction. Oxford University Press, Oxford.

Findley, D.F., Monsell, B.C., Bell, W.R., Otto, M.C. and Chen B. (1998). New Capabilities and Methods of the X-12-ARIMA Seasonal-Adjustment Program. *Journal of Business and Economic Statistics*, Vol. 16, No. 2., 127-177.

Gomez, V. and Maravall, A. (1992). Time Series Regression with ARIMA Noise and Missing Observations - Program TRAMO, EUI Working paper ECO NO. 92/81, Department of Economics, European University Institute.

Ljung, G.M. and Box, G.E.P. (1978). On a Measure of Lack of Fit in Time Series Models. *Biometrika*, 65, 297-303.

Shiskin, J., Young, A.H. and Musgrave, J.C. (1967). The X-11 Variant of the Census Method II Seasonal Adjustment Program. Technical Paper 15, Bureau of the Census, U.S. Department of Commerce, Washington, D.C.

10 Appendix

10.1 Evidence of Evolving Easter Proximity Effect

In Section 5, we used seasonal factor estimates without an Easter Proximity effect correction to indicate the possibility of the evolving nature of an Easter Proximity effect. To reduce the risk of a biased evaluation we only used the Australian Total Retail Turnover data after 1980. Now, we apply the four different approaches evaluated in Section 7 on two other data spans: (1) data span from April 1962 to December 1979 and (2) full data span from April 1962 to April 1999. Table A.1 shows the estimated parameters of the Easter proximity effect for all three data spans evaluated.

Table A.1 Estimated coefficients of Easter proximity effect for three different spans

Span/Approach	Linear-Linear Regressor		Quadratic-Linear Regressor	
	P_e estimate for pre-Easter period	P_d estimate for during Easter holiday period	P_e estimate for pre-Easter period	P_d estimate for during Easter holiday period
April 1962 - December 1979	0.0117	-0.022	0.011	-0.0216
April 1962 - April 1999	0.0158	-0.0202	0.0168	-0.0217
January 1980 - April 1999	0.0181	-0.0156	0.0196	-0.0176
Span/Approach	D13 Linear-Linear Regressor		D13 Quadratic-Linear Regressor	
	P_e estimate for pre-Easter period	P_d estimate for during Easter holiday period	P_e estimate for pre-Easter period	P_d estimate for during Easter holiday period
April 1962 - December 1979	0.0139	-0.0221	0.0133	-0.0216
April 1962 - April 1999	0.0142	-0.0188	0.0158	-0.0204
January 1980 - April 1999	0.0139	-0.0145	0.016	-0.0168

For each approach, the values of the estimated parameters for the three data spans are evolving. For example, for regARIMA with a linear-linear approach, the estimated value (column 2) of pre-Easter effect P_e increases with year while the estimated value (column 3) of Easter holiday period effect P_d decreases. For the full span data, both the pre-Easter and during Easter holiday period effects are between the pre-Easter and during Easter holiday period effects from the pre and post 1980 spans respectively. This pattern is consistent for the estimated parameters over the four different approaches. These consistent patterns of the estimated coefficients variations indicate the evolving nature of the Easter proximity effect over the years.

Table A.2 lists the hypothesis test results for the four approaches over three different data spans.

Table A.2 Hypothesis test for three different spans

Span/Approach	Linear-Linear Regressor	Quadratic-Linear Regressor
	$H_o : P_e + P_d = 0$ (p-value)	$H_o : P_e + P_d = 0$ (p-value)
April 1962 - December 1979	-0.0102 (<0.001)	-0.0106 (<0.001)
April 1962 - April 1999	-0.0044 (0.0337)	-0.0049 (0.0173)
January 1980 - April 1999	0.0025 (0.4260)	0.0019 (0.5375)

Span/Approach	D13 Linear-Linear Regressor	D13 Quadratic-Linear Regressor
	$H_o : P_e + P_d = 0$ (p-value)	$H_o : P_e + P_d = 0$ (p-value)
April 1962 - December 1979	-0.0083 (<0.001)	-0.0082 (<0.001)
April 1962 - April 1999	-0.0046 (0.0042)	-0.0047 (0.0036)
January 1980 - April 1999	0.0012 (0.6074)	0.0010 (0.6639)

The hypothesis tests show that the net effects of the pre-Easter increase and Easter holiday decrease are significantly negative if the whole period of pre-Easter and Easter holiday fall into a same month when data span is from April 1962 to December 1979. The net effects are not significantly different from zero when the data span is from January 1980 to April 1999.

The evidence from Tables A.1 and A.2 confirm that the estimated Easter proximity effect does not reflect more recent years if the full data span of Australian Total Retail Turnover data is used. This justifies the use of a truncated series from 1980 onwards when the proposed approaches were under investigation.

10.2 Easter proximity charts: Quadratic-linear regressor

Figure A.3: **D13 quadratic-linear approach with two iterations. Easter Proximity chart for original Australia Total Retail Turnover - truncated span: January 1980 to April 1999 inclusive**

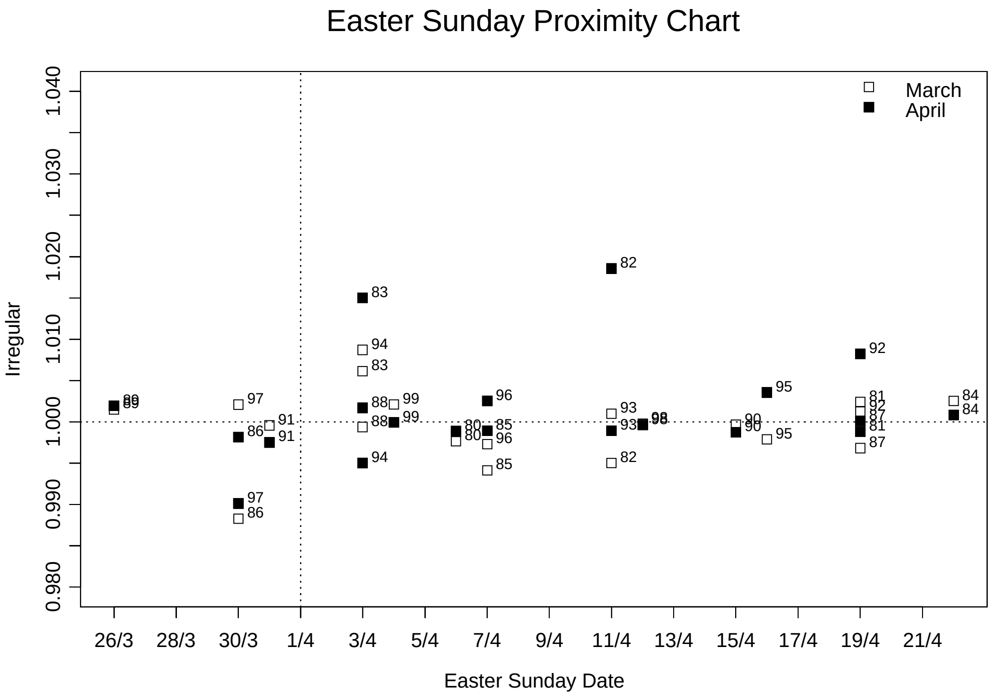


Figure A.4: **Reg-arima quadratic-linear approach. Easter Proximity chart for original Australia Total Retail Turnover - truncated span: January 1980 to April 1999 inclusive**