Seasonal Adjustment Review of Construction (Output) Time Series for the Northern Ireland Statistics and Research Agency

Final Report

Gary Brown and Kevin Moore
Methodology Consultancy Service
Office for National Statistics
July 2009
Tel: 01633 45 5649
Email: gary.brown@ons.gov.uk, kevin.moore@ons.gov.uk

1. Introduction

This report provides details of the first seasonal adjustment review of Construction (Output) time series for Northern Ireland. The reviewers were Gary Brown and Kevin Moore of the UK Office for National Statistics (ONS).

The review was performed in accordance to the detailed specification agreed with the Northern Ireland Statistics and Research Agency (NISRA) and fulfils the contract signed between the two organisations.

The report is structured as follows.

- Section 2 - data supplied and software used
- Section 3 - objectives of review
- Section 4 - summary of analysis
- Section 5 - results of review
- Section 6 - recommendations for implementation and for the future

2. Data and software

In total, 17 series from quarter 1 2000 – quarter 4 2008 (based on constant 2005 prices) were reviewed. Each series was saved as an individual "Text (Tab delimited) (*.txt)" file with a ".dat" file extension prior to reading into X-12-ARIMA.

The software used for seasonal adjustment was X-12-ARIMA version 0.3 build 177. Microsoft Excel 2002 and R version 2.3.1 were used for graphics.

3. Objectives

The objectives for this review, as agreed in "Seasonal Adjustment Review of NISRA Construction (Output) series.doc", included the following.

1. Full review of all Construction series, leading to recommendations on the four key seasonal adjustment options.
    a) Decomposition
    b) ARIMA models
    c) Seasonal and Henderson (trend) moving averages (SMA & TMA)
    d) Prior adjustments
2. Provision of X-12-ARIMA specification files for NISRA to run seasonal adjustment.
3. A report explaining processes, and illustrating results, from the review.

The review will not include:

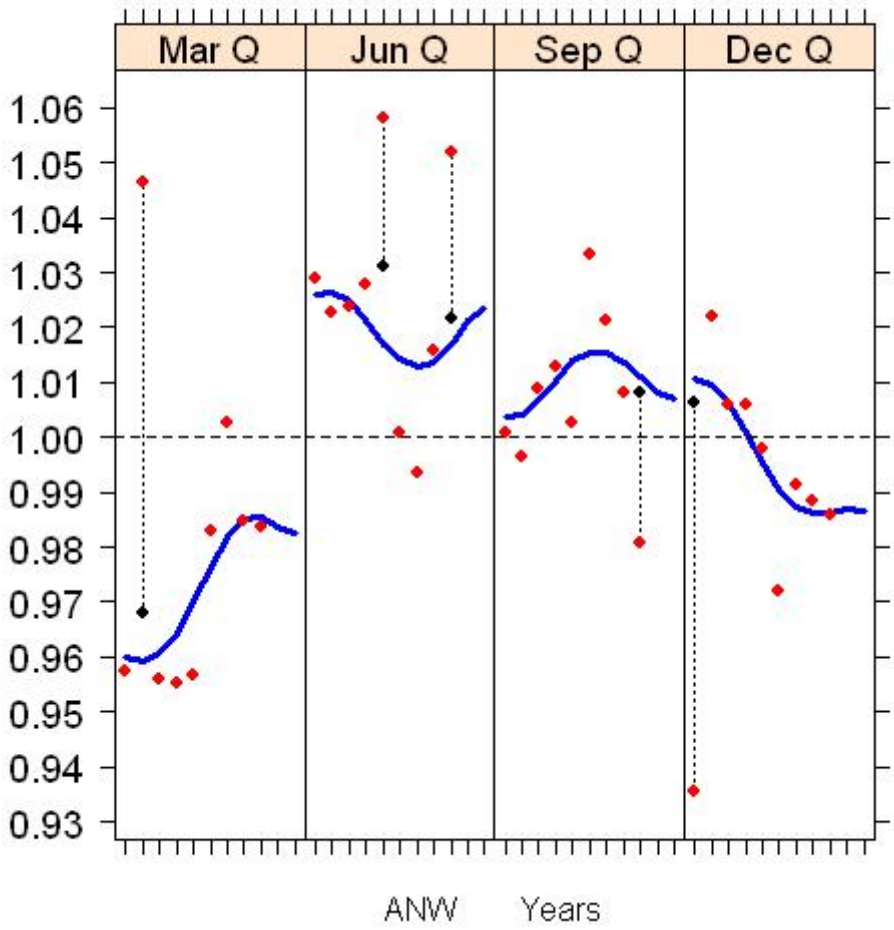
1. Forcing additivity in aggregation structures. There is no statistical reason for additivity to hold, and it's unlikely by chance. For example, the "All New Work" unadjusted series is the sum of six component unadjusted series. When all of these series have been seasonally adjusted, the "All New Work" seasonally adjusted series is not guaranteed to equal the sum of its seasonally adjusted components.

4. Analysis

All New Work (ANW)

The default[1] run on this series indicated seasonality present at the 1% level, but the combined test indicated none[2]. However, the "SI ratio" chart[3] in Figure 1 indicates seasonality as values are all above 1 in some quarters and below in others. Figure 1 also shows a sharp rise in quarter 1. This is sometimes an indication of a seasonal break[4]. However, when tested, the break was found to be not statistically significant. The proximity of final M7 quality measure (0.845)[5] to 1 also gives some indication that the series is only slightly seasonal.

Figure 1: SI chart of series ANW

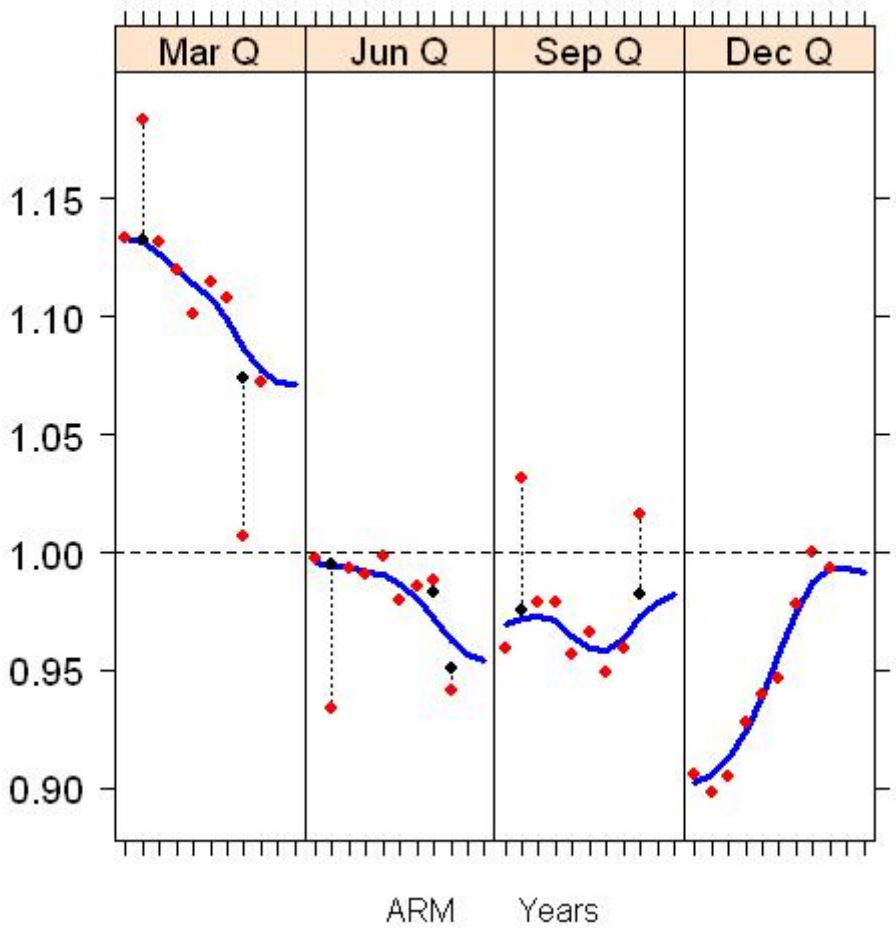


[1] The "default run" applies a standard X-12-ARIMA specification file which automatically selects the best model on statistical principles only. This selection is then subject to expert quality assurance.
[2] X-12-ARIMA automatically runs three tests for seasonality – two assuming stability (parametric and non-parametric), and one for moving seasonality. An individual test may indicate seasonality whilst the combination of the three may not – in this case we always investigate further.
[3] The "SI ratio" chart is a misnomer. It is actually the S*I series (or S+I for an additive series) split by period. However, the convention of calling it the "SI ratio" is followed here.
[4] A seasonal break is a sudden and sustained change in the seasonal pattern in a series.
[5] M-statistics range between 0 and 3. Values below 1 are deemed a "pass".

All Repair and Maintenance (ARM)

The unadjusted series (not shown) looks seasonal, with a potential seasonal break in recent years. The default run indicated seasonality with good M-statistics. An examination of the SI chart (Figure 2) shows seasonality changing in quarters 1 and 4. However, the moving averages (blue line) seem to be coping with the change, and as values are on or close to the line we decided not to adjust for a seasonal break.

Figure 2: SI chart of series ARM



All Work (AW)

The default run indicated seasonality and identified an additive outlier (t-value 4.40) in January 2001. Normally, an outlier would not be adjusted for unless a real-world reason was provided[6]. However, the outlier is large relative to the movements in the series and not adjusting for it makes the quality of the adjustment considerably worse.

Index of Housing (IH)

The default run produced good M-statistics except the last - M10 just failed and M11 just passed. High values for M10 and M11 sometimes point to the possibility of a

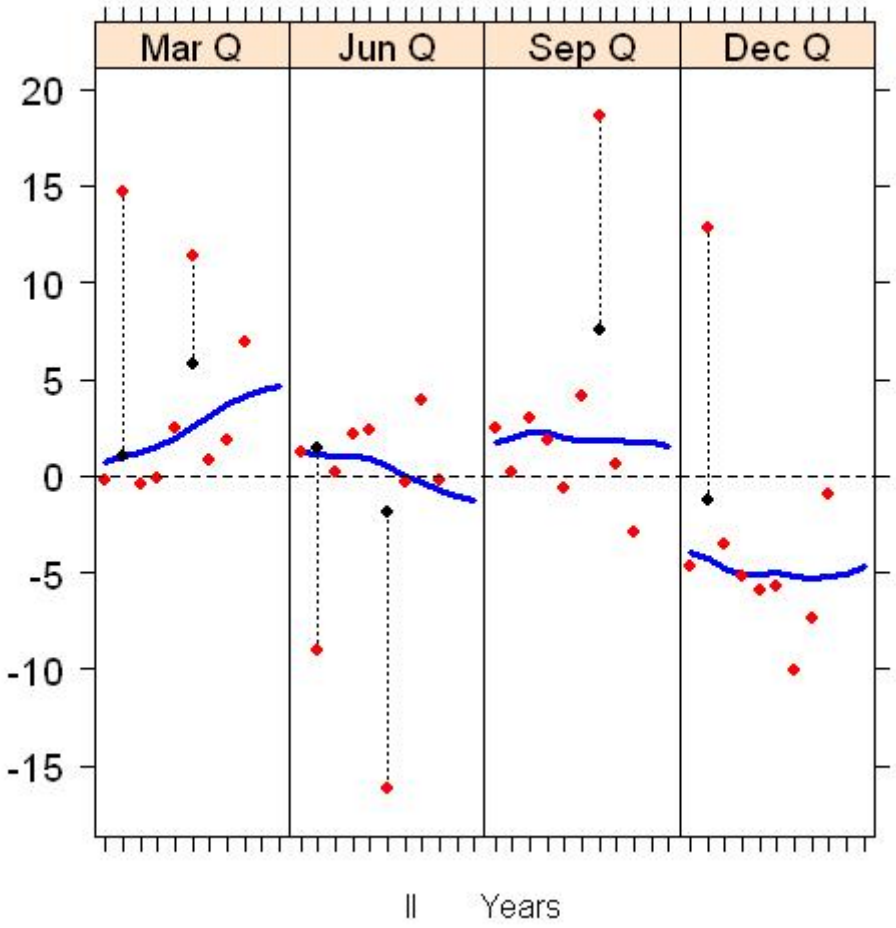
[6] The Seasonal Adjustment Review in ONS is iterative. Outliers and level shifts are only adjusted for if the business area can identify real-world causes – otherwise they are deemed random variation.

seasonal break, but no evidence was found in the SI chart (not shown). Neither were there any systematic or unusual amounts of weight reduction in table C17 for the irregular component – another indicator of a potential seasonal break. A plot of the seasonally adjusted series against the unadjusted series (not shown) also showed no anomalies. So no amendments to the automatic parameters were required.

Index of Infrastructure (II)

The default run found no evidence of seasonality in any test, and as the quality diagnostics (M-statistics) were reasonable and the SI chart (Figure 3) also showed very little evidence of seasonality, this series was not seasonally adjusted – it would not add any benefit to the interpretation of the series.

Figure 3: SI chart of series II



Index of Other Work (IOW)

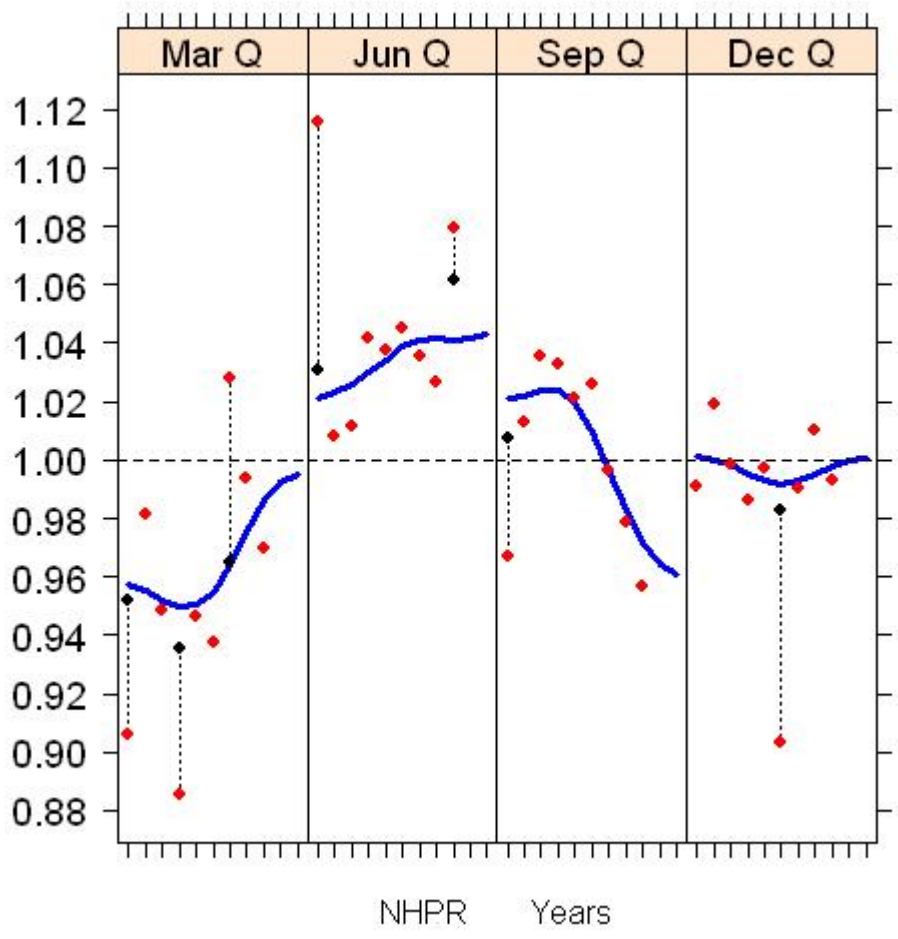
The default run indicated this series was not seasonal, however Easter was found to be significant (t = -2.05). This is a type I error – there is a 1 in 20 chance that a test statistic falling outside a 95% confidence interval could be that extreme by chance. The usual assumption is these errors are "unlikely to be happen at random" – but given the lack of seasonality in the series, the supplemental information invalidates the t-test. The quality statistics M7 - M11 fail, and there was no evidence of a

seasonal break from the SI chart (not shown) – so the series should not be seasonally adjusted.

New Housing: Private (NHPR)

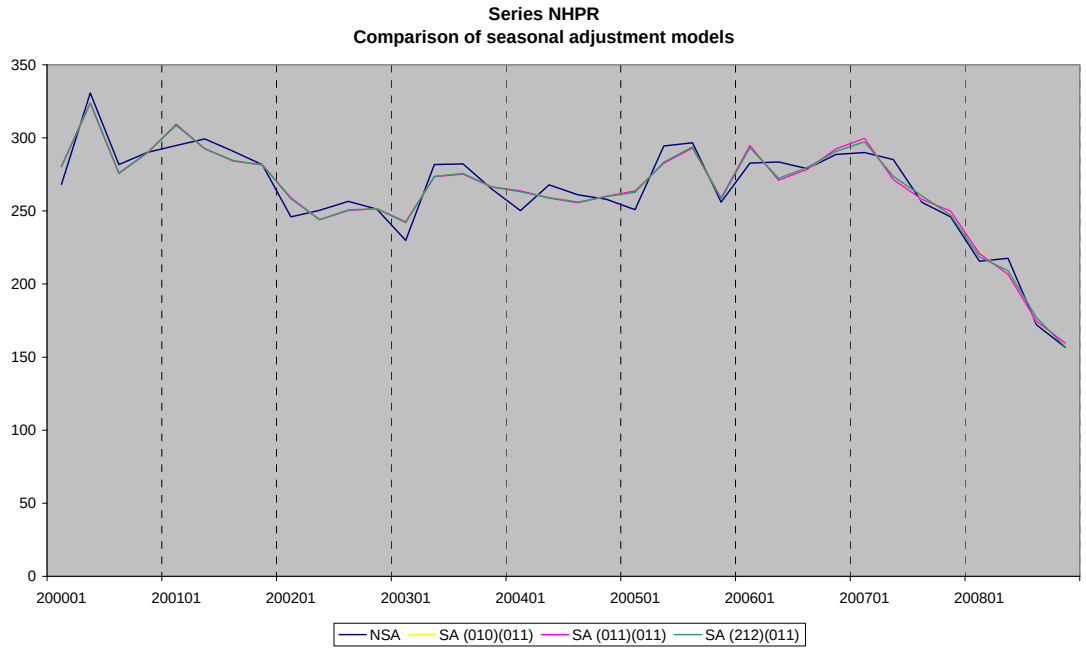
The default run indicated the series was seasonal, but fitted a complex (212)(011) ARIMA model automatically (with an acceptable fit, Q = 0.77[7]). However, five individual M-statistics failed, including all those relating to a seasonal break (M8 - M11). Examination of the SI chart (Figure 4) shows seasonality evolving in recent periods but no evidence of a break. The default run uses the "pickmdl" option to select an ARIMA model - as the result was unusual, the validity was verified by running "automdl"[8]. The result was a more parsimonious (010)(011) model, however the quality measures did not improve. Finally, an "airline" (011)(011) model was tested – it had the best fit in terms of quality statistics. Figure 5 shows very little difference in the seasonally adjusted series obtained from the three models. As the airline model is known to fit most economic series it is recommended for this series in the short term. The model should however be reviewed when more data are available.

Figure 4: SI chart of series NHPR



[7] The "Q" statistic summarises all the M-statistics (except M2) to assess overall goodness of fit.
[8] Pickmdl selects an ARIMA model from a short-list of five. Automdl selects from all possible models.

Figure 5: Three different ARIMA models for series NHPR



New Housing: Public (NHPU)

The default run indicated seasonality at the 1% level but none in the combined test. Five quality statistics failed (M7 – M11) and the default airline model was selected as none of the others passed the required tests[9]. No seasonal break was evident in the SI charts (not shown) that would resolve the M8-M11 statistic failures. The SI charts did show the possibility of seasonality. However, from using different spans of data and the automdl procedure, and still not finding any significant seasonal breaks, it is recommended this series is not adjusted for seasonality.

Other New work: Infrastructure (ONWI)

The default run indicated this series was not seasonal. The ARIMA model selected was (210)(011) - the last model to be tested[10]. However, the quality statistics showed this model was a poor fit - seven of the eleven measures failed and the model was rejected by the Q statistic. The SI chart (not shown) demonstrated random variation around zero giving more evidence to the series being not seasonal. We recommend this series is not adjusted for seasonality.

Other New Work: Public (ONWP)

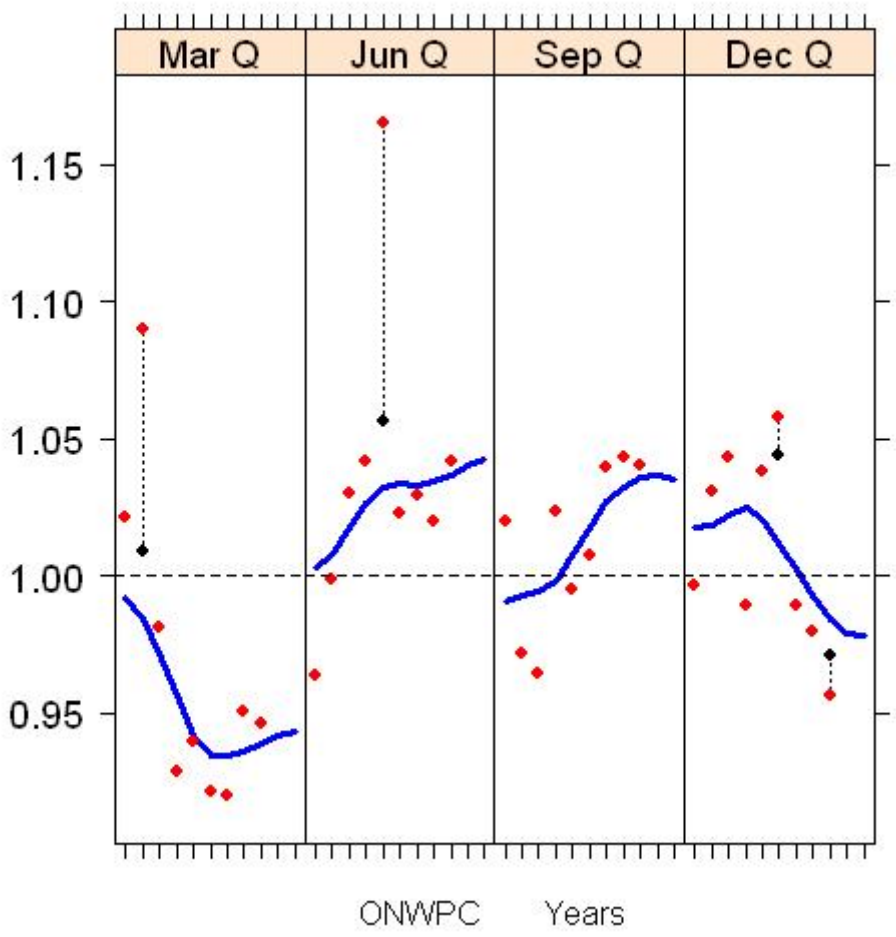
The default run indicated this series was not seasonal. The airline model automatically selected by X-12-ARIMA didn't fit well, with six of the eleven quality measures failing and the model being rejected. High amounts of irregularity were present. We recommend this series is not adjusted for seasonality.

[9] If none of the five models satisfies the quality criteria, Pickmdl selects the airline model by default.
[10] Pickmdl options are "select best model" or "select first model that passes criteria" (the ONS choice).

Other New Work: Private Commercial (ONWPC)

The default run indicated this series was not seasonal, however Easter was found to be significant (just). Further analysis revealed that the M7 statistic only just failed, and examination of the SI chart (Figure 6) shows a sharp drop in quarter 1 2003 - a highly significant seasonal break (p=0.000). On adjusting for this break, Easter became more significant and seasonality tests became positive. We recommend this series is seasonally adjusted with an adjustment for a seasonal break at quarter 1 2003.

Figure 6: SI chart of series ONWPC

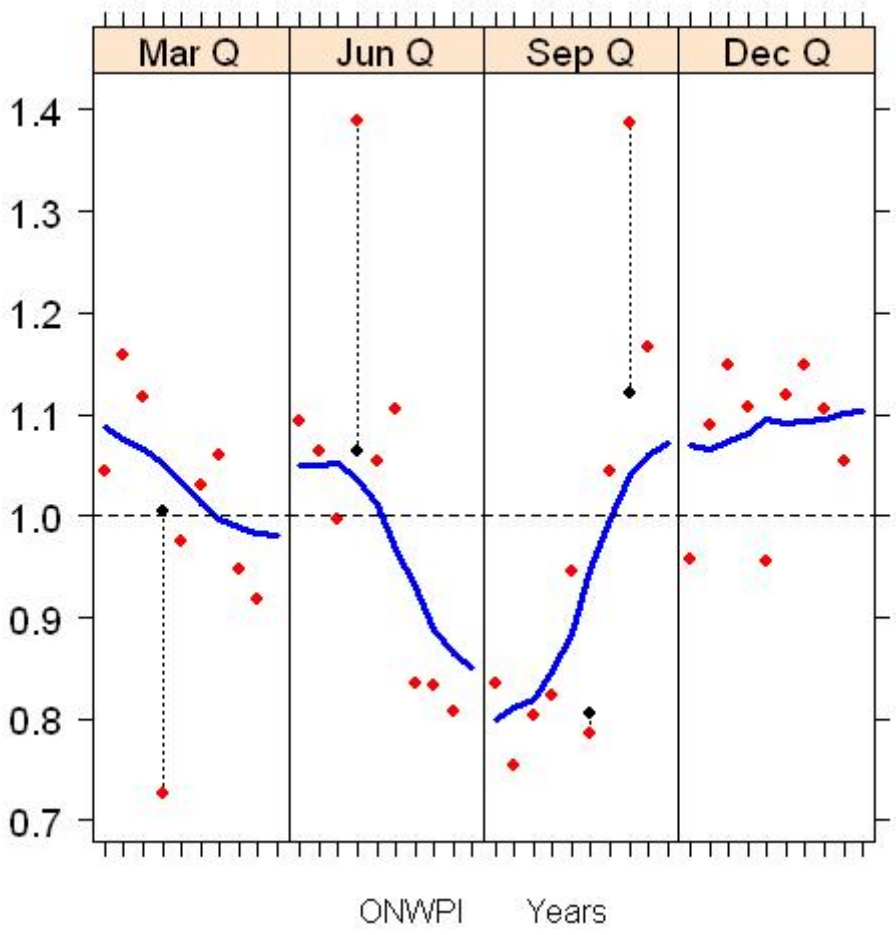


Other New Work: Private Industrial (ONWPI)

The default run indicated this series was not seasonal, and returned very poor quality M-statistics. However, on examination of the SI chart (Figure 7), a distinct change can be seen in quarter 2 of 2006. The seasonal break was highly significant (p=0.000). Normal practice requires three full years of data after a break to ensure that the effect is modelled correctly. However, the break is so significant here that two and a half years is acceptable. After including this break, the quality statistics for the irregular component still failed. Figure 7 shows four values have been amended - three quite considerably (indicated by the dotted line). When all of these points were included in the regression specification as outliers, only quarter 1 2003 had a significant t-value. Adjusting for this outlier alone improved the irregular component and hence the quality of the seasonal adjustment overall. We recommend this series is adjusted for

seasonality adjusting for a seasonal break in quarter 2 2006 and an additive outlier in quarter 1 2003.

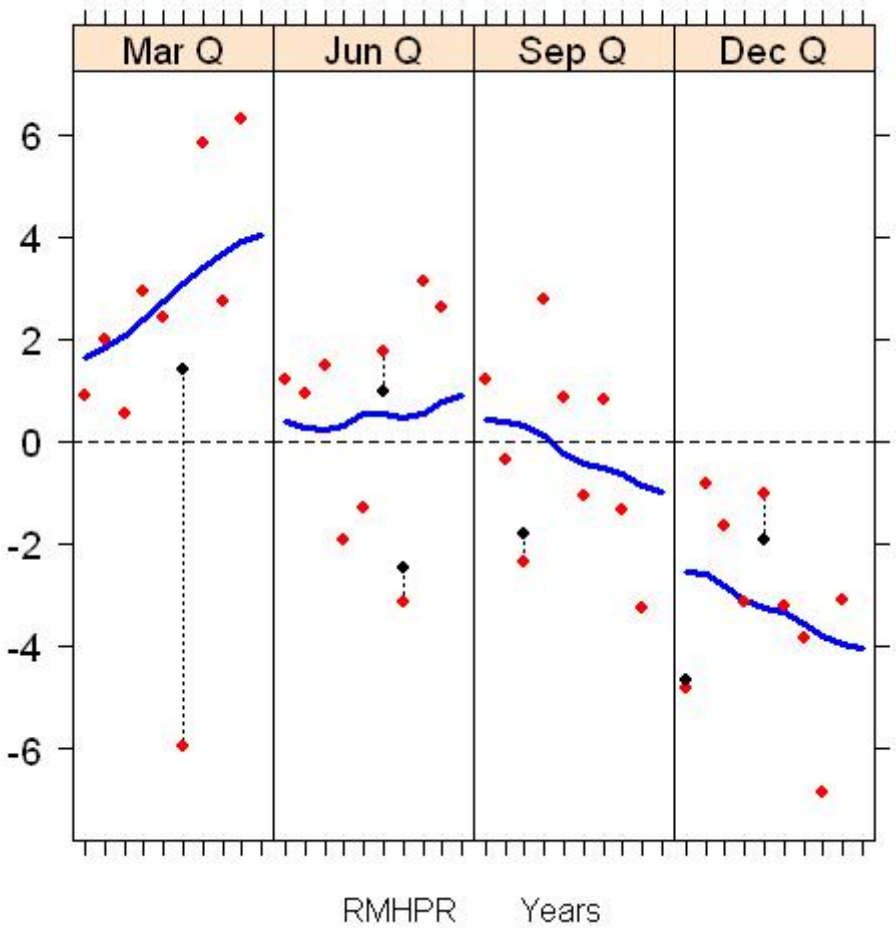
Figure 7: SI chart of series ONWPC



Repair and Maintenance – Housing: Private (RMHPR)

The default run indicated this series was seasonal at the 1% level, but the combined test concluded there was no seasonality. However, the SI chart (Figure 8) shows clear seasonality in quarter 1 (above) and quarter 4 (below). The value for quarter 1 2005 has been considerably adjusted – including this as an outlier in the regression specification resulted in positive seasonality tests, quarter 1 2005 being a significant regressor, the decomposition of the model changing to multiplicative and Easter becoming significant. We recommend this series is seasonally adjusted using the parameters as stated.

Figure 8: SI chart of series RMHPR



Repair and Maintenance – Other Work: Private (RMOWPR)

The default run indicated this series was not seasonal. Pickmdl selected the airline model as a default, but this did not fit well, and the SI chart was random (not shown). Hence we recommend not adjusting this series for seasonality.

Repair and Maintenance – Other Work: Public (RMOWPU)

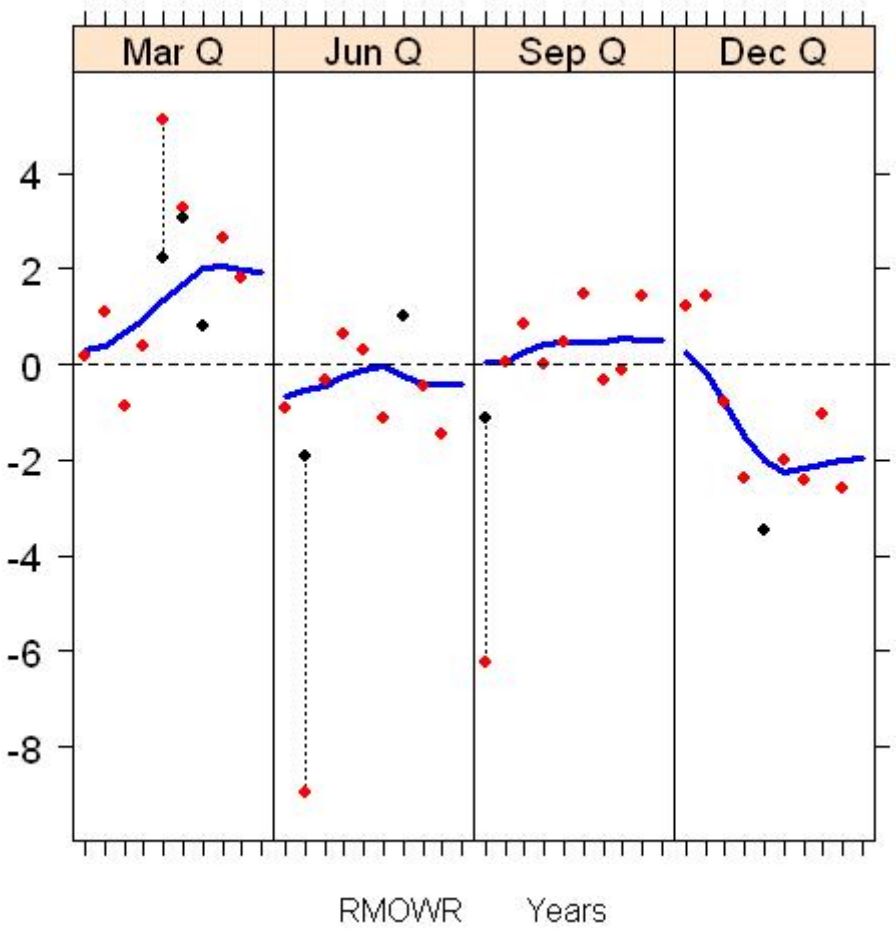
The default run indicated this series was seasonal. The airline model was chosen as a default, and proved a fair fit - only one quality measure failed and the Q statistic was 0.68. However, when the automdl procedure was run, an even more parsimonious (001)(011) was obtained and Easter was found to be significant. This combination gave a vast improvement in the quality of the adjustment, no quality measures failed and the overall Q statistic was 0.32. We recommend this series is seasonally adjusted using a (001)(011) ARIMA model with an adjustment for Easter.

Repair and Maintenance – Other Work: Roads (RMOWR)

The default run indicated this series was seasonal at the 1% level but not at the combined level. The SI chart (Figure 9) also indicated seasonality, and a seasonal break at quarter 1 2002 was significant. This break was not clear in Figure 9, but was evident in the plot of the unadjusted series (not shown). We also can see from Figure

9 that the values for quarter 2 2001 and quarter 3 2000 have been amended considerably. These were added to the regression specification as outliers and found to be significant. The resulting seasonal adjustment still has some quality failures but the series is quite volatile and the overall Q statistic of 0.65 is acceptable. We recommend seasonally adjusting this series with a seasonal break in quarter 1 2002 and adjusting for additive outliers in quarter 2 2001 and quarter 3 2000.

Figure 9: SI chart of series RMOWR

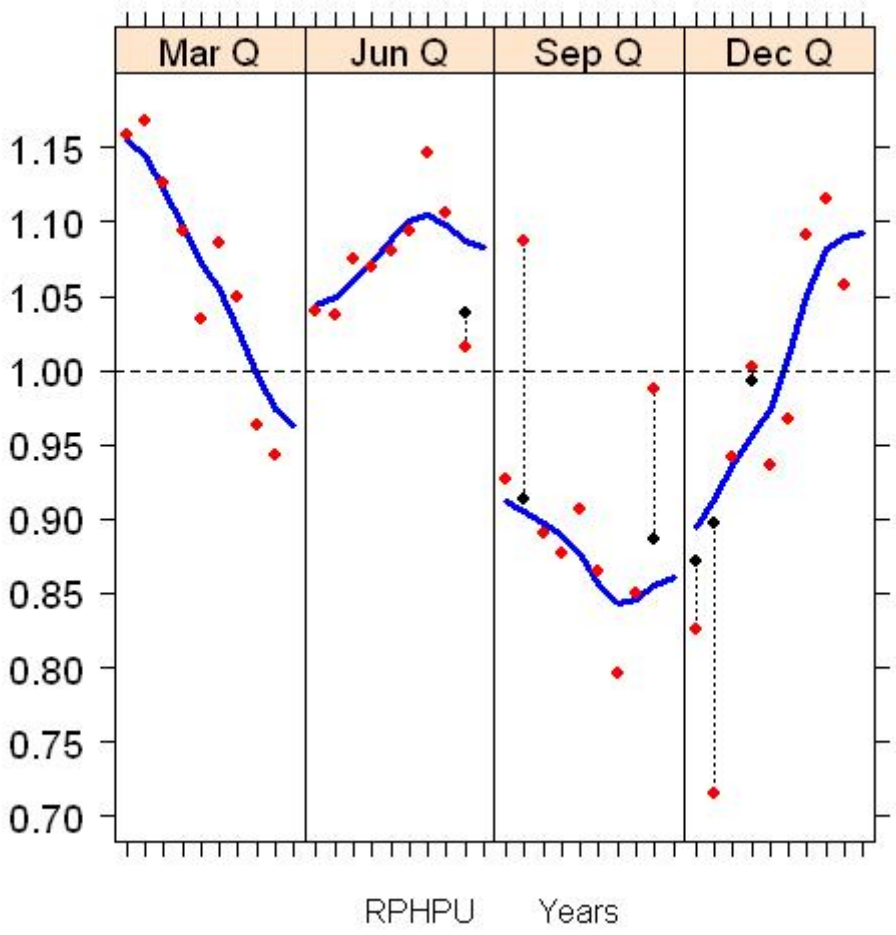


Repair and Maintenance – Housing: Public (RPHPU)

The default run indicated seasonality at the 1% level but not using the combined measure. The SI chart (Figure 10) showed seasonality in quarters 2 and 3 with evolving seasonality in the other two quarters. The default airline ARIMA model was chosen by default, so the automdl procedure was run – this chose (200)(011). Other parameters obtained from the automdl procedure included the constant term in the regression (t<-10). A test for a possible seasonal break evident from Figure 10 in quarter 4 2006 was positive (p=0.000). After adjusting for these parameters, the irregularity of the series was still affecting the quality of the adjustment. Two final regressors were identified: one via examination of Figure 10 – a large amendment to the quarter 3 2001 point; and one via examination of the unadjusted plot (not shown) – a level shift in quarter 1 2002. When these two regressors were added the quality measures and the overall fit improved sufficiently for us to recommend this series is

seasonally adjusted with a seasonal break in quarter 4 2006, and adjustments for a level shift in quarter 1 2002 and an additive outlier in quarter 3 2001.

Figure 10: SI chart of series RPHPU



5. Results

Table 1 summaries the results discussed in section 4.

Table 1: Parameters for seasonal adjustment

Name	Decomposition	Seasonal?	TMA	SMA	Model	Regressors	t-values
ANW	Multiplicative	Some	5	3x5	(0 1 1)(0 1 1)		
ARM	Multiplicative	Yes	5	3x3	(0 1 1)(0 1 1)	Easter[15]	-3.02
AW	Multiplicative	Yes	5	3x9	(0 1 1)(0 1 1)	AO2001.1	4.40
IH	Multiplicative	Yes	5	3x5	(0 1 1)(0 1 1)		
II		No					
IOW		No					
NHPR	Multiplicative	Yes	5	3x5	(0 1 1)(0 1 1)		
NHPU		No					
ONWI		No					
ONWP		No					
ONWPC	Multiplicative	Yes	5	3x5	(0 1 1)(0 1 1)	SB2003.1 Easter[15]	p=0.000 -2.63
ONWPI	Multiplicative	Yes	5	3x5	(0 1 1)(0 1 1)	SB2006.2 AO2003.1	p=0.000 3.43
RMHPR	Multiplicative	Yes	5	3x5	(2 1 2)(0 1 1)	Easter[8] AO2005.1	-2.05 -3.75
RMOWPR		No					
RMOWPU	Multiplicative	Yes	5	3x5	(0 0 1)(0 1 1)	Easter[15]	-4.04
RMOWR	Multiplicative	Yes	5	3x9	(0 1 1)(0 1 1)	SB2002.1 AO2000.3 AO2001.2	p=0.00 -5.26 -6.59
RPHPU	Additive	Yes	5	3x5	(2 0 0)(0 1 1)	Constant AO2001.3 LS2002.1 SB2006.4	-8.92 4.35 9.97 p=0.00

6. Recommendations

a) Seasonal series should be adjusted according to the parameters specified in Table 1. The X-12-ARIMA specification files provided with this report should be used for this purpose.

b) Non-seasonal series should not be seasonally adjusted – it would distort rather than clarify the picture.

c) The ".rmx" files specific to each series contain regressors to adjust for seasonal breaks until quarter 4 2010 (extending these files for future years is simply a case of extending the sequence of 0's or 1's as appropriate). These ".rmx" files need to be saved in the same folder as the ".spc" specification files for the seasonal adjustment to work.

d) Seasonal adjustment models and parameters should be reviewed annually – for instance, the best model for series NHPR is currently slightly ambiguous and another year's data might clarify it. The starting point for subsequent reviews should be the current model.