Approaches to the Seasonal Adjustment of Labour Market Flows.

Ricky Kanabar[1] and Daniel Ayoubkhani, Office for National Statistics, November 2012[2]

1. Introduction

Each quarter the Office for National Statistics (ONS) publishes estimates of the number of individuals who are in employment, unemployment and inactivity using data from the Labour Force Survey (LFS). The LFS is a national survey which follows households for five consecutive quarters, after which point they are rotated out. Therefore at any one quarter for one fifth of the sample it will be their first interview.[3] For more detailed information see the ONS LFS methodology website.[4]

Given the structure of the LFS the ONS is able to publish 2 quarter longitudinal datasets.[5] In effect these datasets follow an individual from one quarter to the next and therefore by definition constitute a panel.[6] Moreover in order to eligible to be part of this longitudinal study an individual should be present in each quarter, these are what are known as balanced panels.[7] It is possible that an individual can change economic status across the two quarters and therefore this investigation is related to flows between the three aforementioned economic states.

The ONS currently publishes estimates of labour market flows (REF to experimental flows data online) using data from the two quarter longitudinal datasets; however these are not currently seasonally adjusted. Seasonal adjustment aims to remove effects associated with the time of year or arrangement of the calendar, so that underlying movements within time series may be more easily interpreted. The indirect approach to seasonal adjustment involves adjusting subcomponent series to derive a seasonally adjusted aggregate, whilst the direct approach involves adjusting the aggregate itself. The aim of this investigation is to compare direct and indirect seasonally adjusted estimates of inflows and outflows

---

[1] Department of Economics and Related Studies, University of York, UK. Author was on a PhD internship sponsored by the ESRC whilst completing this work.

[2] The authors would like to acknowledge very helpful advice and comments from Nicholas Palmer and Gary Brown.

[3] This only holds true if there is no attrition otherwise as a proportion of the sample the refreshment will constitute more than one fifth of the sample.

[4] For more information: http://www.ons.gov.uk/ons/guide-method/method-quality/specific/labour-market/labour-market-statistics/index.html.

[5] The ONS also publishes five quarter longitudinal datasets however this is not the focus of this report.

[6] The questions in the LFS mean it is not possible to determine transitions that occur within these three month intervals.

[7] In some cases which are related to the age (75+) of the individual in their first interview, respondent's information is only collected in the first and last wave. In other cases information may be rolled forward if individuals are not contactable in two consecutive quarters. Cases such as these are not included in the longitudinal datasets.

associated with labour market status in the UK, and to subsequently recommend a preferred approach to the seasonal adjustment of these flows.

The aim of this report is to investigate the extent of seasonality in the flows estimates derived from the two quarter longitudinal datasets. We also want to determine which type of seasonally adjusted aggregate flows estimates if preferred, direct or indirect.

2. The Labour Market Flows and approaches to their Seasonal Adjustment

2.1 Description of the labour market

Conditional on being observed in two consecutive quarters on the two quarter longitudinal datasets, an individual can transition between any three states: (E)mployed, (U)nemployed or (I)nactive. This gives rise to nine possible transitions summarised in table 1.

Table 1: Transition matrix

Quarter	Q+1			
Q	Status	E	U	I
	E	X_{EE}	X_{EU}	X_{EI}
	U	X_{UE}	X_{UU}	X_{UI}
	I	X_{IE}	X_{IU}	X_{II}

Each of the elements (X_{ij}) in the matrix reflects the magnitude of the flows between two consecutive quarters. From this information, estimates of gross inflows and outflows for each of the statuses can be obtained, as summarised in table 2.

Table 2: Definitions of flows across two consecutive quarters

	Definition
Inflows to	
Employment	$X_{UE} + X_{IE}$
Unemployment	$X_{EU} + X_{IU}$
Inactivity	$X_{EI} + X_{II}$
Outflows from	
Employment	$X_{EU} + X_{EI}$
Unemployment	$X_{UE} + X_{UI}$
Inactivity	$X_{IE} + X_{IU}$

2.2 Seasonal adjustment in ONS

Seasonal adjustment aims to remove effects associated with the time of year or arrangement of the calendar, so that underlying movements within time series may be more easily interpreted. An economic time series may be thought of as an aggregation of a number of unobservable components, namely:

- **a trend-cycle:** the underlying long-term movement in the series, plus cyclical movements about it
- **a seasonal component:** within-year movements about the trend-cycle that remain roughly constant in magnitude and direction, and in the same month/quarter, from year to year
- **an irregular component:** residual variation after the trend-cycle and seasonal component (plus calendar-related effects) have been removed from the series, for example short-term shocks

The way in which the components are aggregated may be additive (whereby the components are summed) or multiplicative (whereby the components are multiplied together). Series that have a multiplicative pattern are typically log transformed to enforce additivity in order to estimate the components.

Seasonal adjustment aims to estimate the seasonal component and remove it (as well as calendar-related effects) from the observed time series to leave the trend-cycle and irregular component unchanged. Throughout ONS, the primary method for seasonal adjustment is the X-11 algorithm, as detailed in Ladiray and Quenneville (2001). The method is based on the iterative application of moving averages, which are used to decompose a time series into its unobserved components. However, due to the symmetrical nature of such moving averages, decomposed values for the most recent periods are obtained after the series is extended with forecasts. The forecasts are appended to “regARIMA” (regression with Autoregressive Integrated Moving Average errors) models, which are also used to derive adjustments for additive outliers and level shifts. These adjustments are known as “prior adjustments”, because they are made prior to the application of the X-11 algorithm (if prior adjustments were not made for such effects, seasonally adjusted estimates might be distorted). More information on ARIMA and regARIMA models can be found in Box et al (2008), Findley et al (1998) and US Census Bureau (2011).

2.3 Approaches to seasonally adjusting the labour market flows

Seasonal adjustment of each of the six aggregate flows series as listed in table 2 can be done in two ways:

1. **direct estimation:** seasonally adjust each of the six aggregate flows series
2. **indirect estimation:** seasonally adjust each of the six off diagonal elements in table 1 and sum according to table 2 in order to derive seasonally adjusted aggregate flows

Of course, seasonal adjustment of net inflows to a particular employment status also leads to a choice between direct and indirect estimation; should one seasonally adjust aggregate inflows and outflows to derive net inflows, or seasonally adjust net inflows itself? However, the analysis reported on here focused solely on seasonal adjustment of the six aggregate flows series.

Note that these are not equivalent and will not give exactly the same estimates. The direct approach is conceptually appealing, as seasonal adjustment is performed directly on the series of interest. However, the indirect approach guarantees that the sum of the seasonally adjusted subcomponents is equal to the seasonally adjusted aggregate, which is not the case for direct estimation. More information and discussion on the two approaches can be found in Eurostat (2009), and Ladiray and Mazzi (2003).

In order to obtain directly and indirectly seasonally adjusted estimates, each of the six transition series and six aggregate flows series was seasonally adjusted using the X-12-ARIMA software package (US Census Bureau 2011). Many of X-12-ARIMA's automatic modelling and selection capabilities were utilised during this process, including:

- a test for the need for a log transformation
- selection of the ARIMA model orders
- additive outlier and level shift detection and correction
- selection of the trend and seasonal moving average lengths

Where necessary, X-12-ARIMA's automatic selections were amended via manual intervention following, for example, visual inspection of the plotted original and seasonally adjusted data.

Figures 1 to 6 illustrate directly and indirectly seasonally adjusted estimates for each of the six aggregate flows series. The sample used for estimation is all individuals aged 16-64.[8] All of the series exhibit obvious seasonal behaviour. Some series are more volatile than others,

[8]Figures 1-6 are quoted in 000's.

for example outflows to employments, and some contain abrupt changes in level, for example inflows to and outflows from unemployment.

Figure1: Inflows to employment

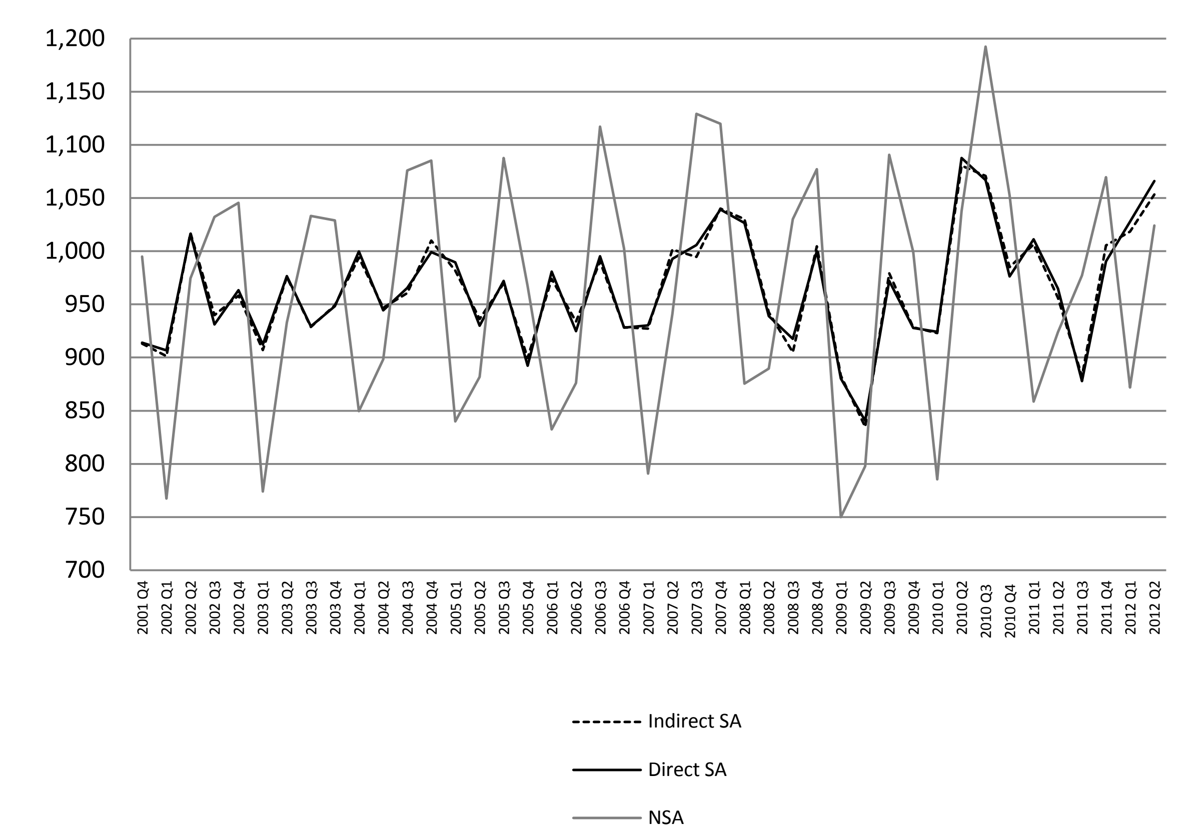
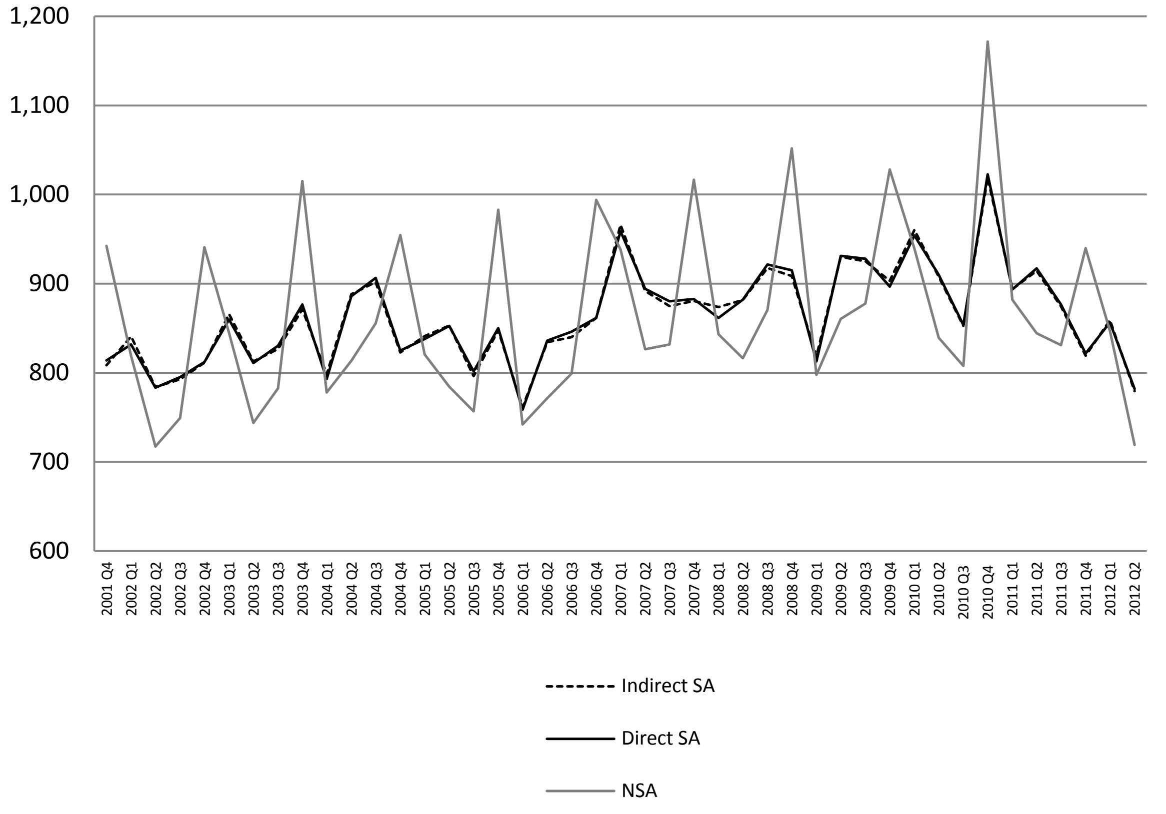


Figure 2: Inflows to unemployment

Figure 3: Inflows to inactivity



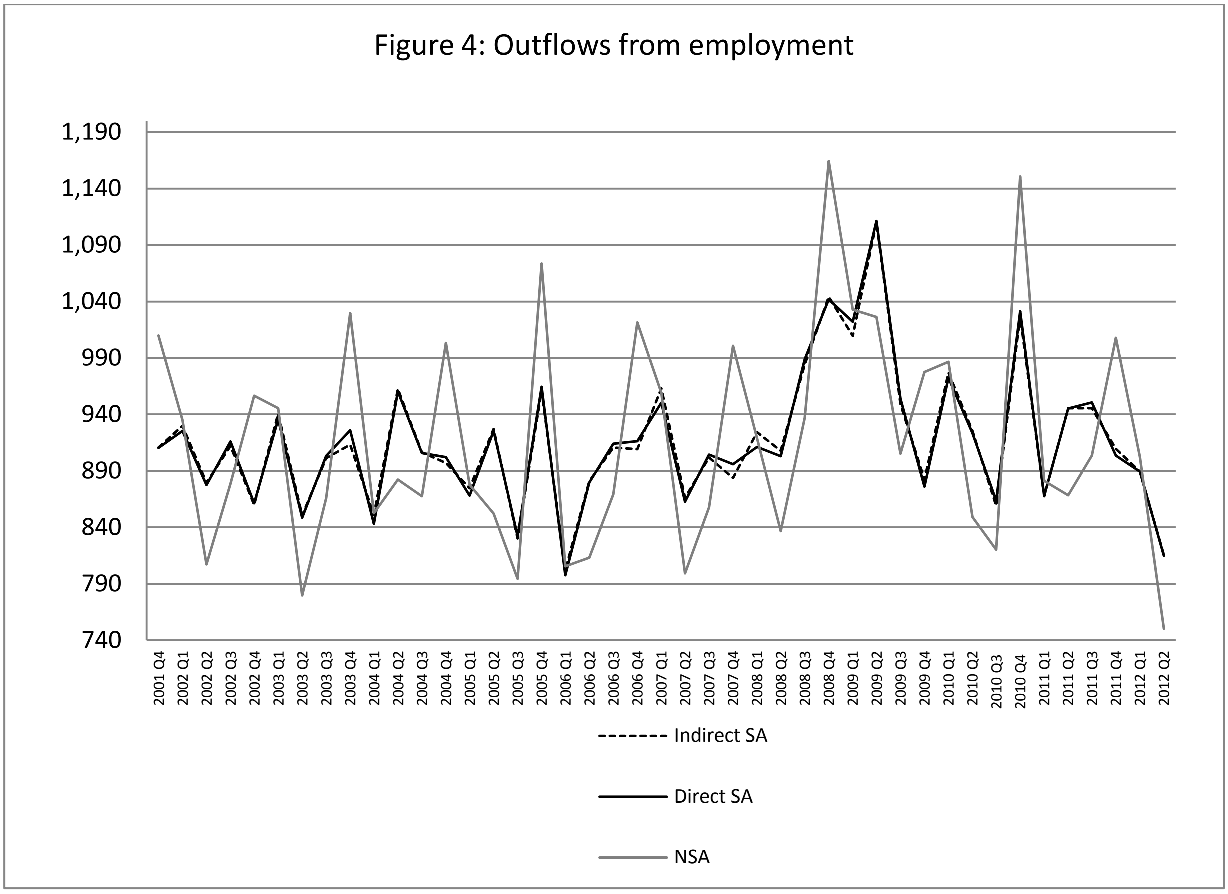


Figure 4: Outflows from employment
1,190
1,140
1,090
1,040
990
940
890
840
790
740
2001 Q4
2002 Q1
2002 Q2
2002 Q3
2002 Q4
2003 Q1
2003 Q2
2003 Q3
2003 Q4
2004 Q1
2004 Q2
2004 Q3
2004 Q4
2005 Q1
2005 Q2
2005 Q3
2005 Q4
2006 Q1
2006 Q2
2006 Q3
2006 Q4
2007 Q1
2007 Q2
2007 Q3
2007 Q4
2008 Q1
2008 Q2
2008 Q3
2008 Q4
2009 Q1
2009 Q2
2009 Q3
2009 Q4
2010 Q1
2010 Q2
2010 Q3
2010 Q4
2011 Q1
2011 Q2
2011 Q3
2011 Q4
2012 Q1
2012 Q2
Indirect SA
Direct SA
NSA

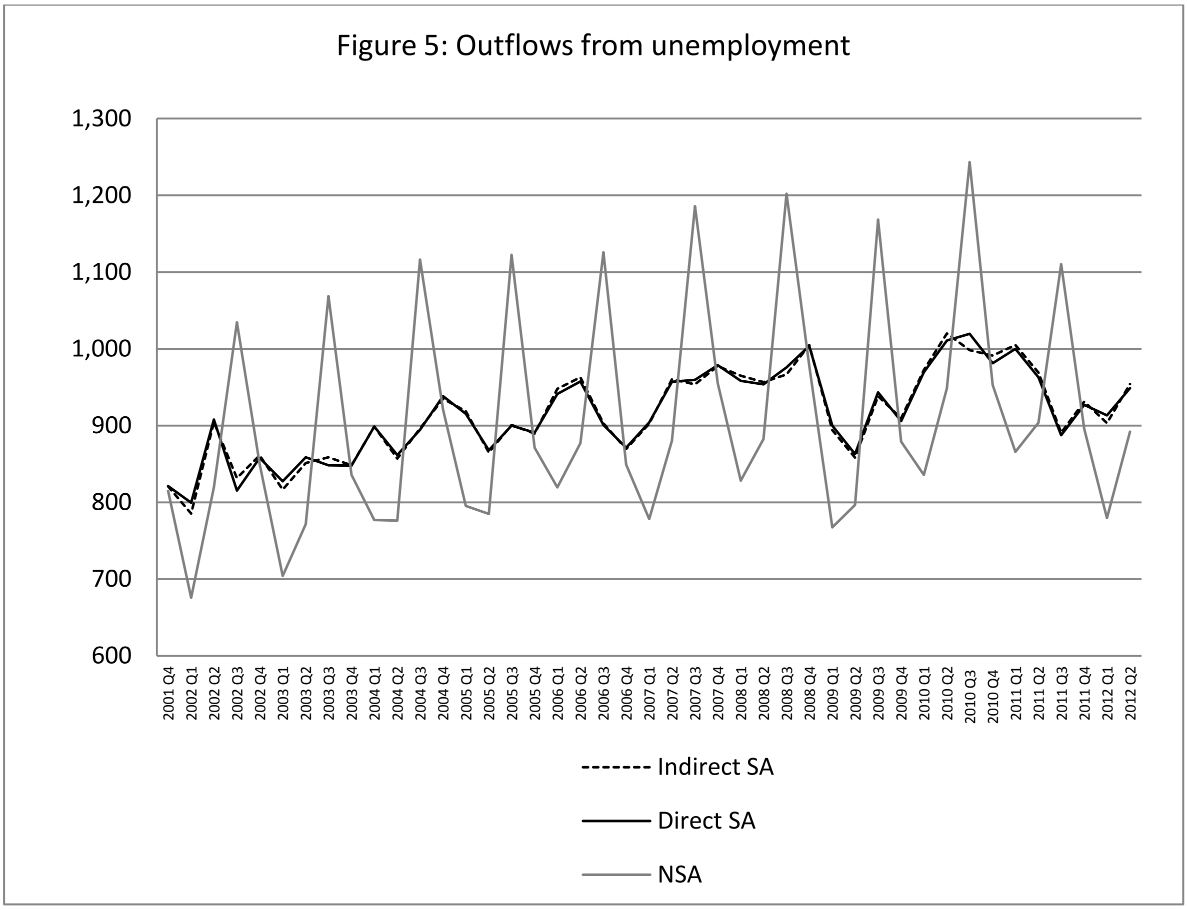


Figure 5: Outflows from unemployment
1,300
1,200
1,100
1,000
900
800
700
600
2001 Q4
2002 Q1
2002 Q2
2002 Q3
2002 Q4
2003 Q1
2003 Q2
2003 Q3
2003 Q4
2004 Q1
2004 Q2
2004 Q3
2004 Q4
2005 Q1
2005 Q2
2005 Q3
2005 Q4
2006 Q1
2006 Q2
2006 Q3
2006 Q4
2007 Q1
2007 Q2
2007 Q3
2007 Q4
2008 Q1
2008 Q2
2008 Q3
2008 Q4
2009 Q1
2009 Q2
2009 Q3
2009 Q4
2010 Q1
2010 Q2
2010 Q3
2010 Q4
2011 Q1
2011 Q2
2011 Q3
2011 Q4
2012 Q1
2012 Q2
Indirect SA
Direct SA
NSA

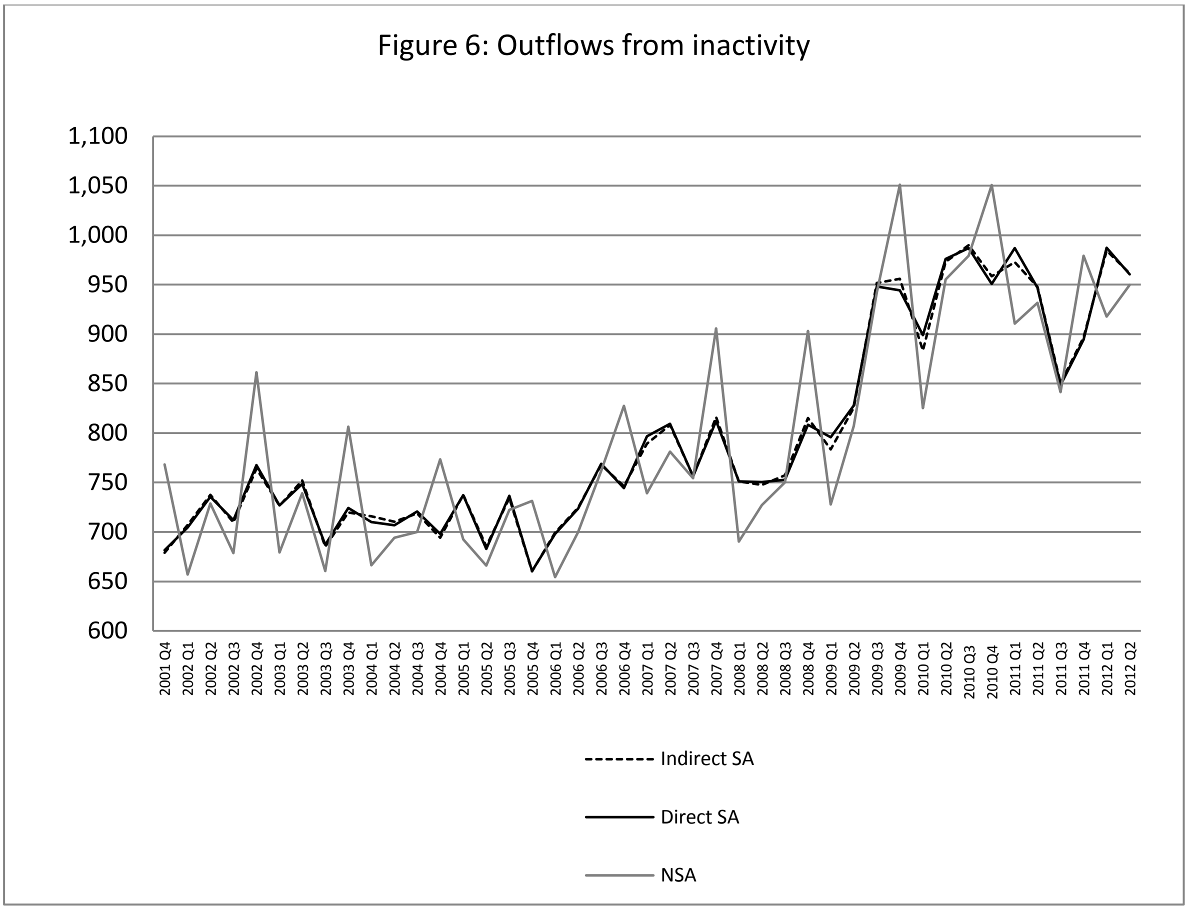


Figure 6: Outflows from inactivity
1,100
1,050
1,000
950
900
850
800
750
700
650
600
Indirect SA
Direct SA
NSA

Table 3 summarises mean and maximum absolute percentage differences (MeAPD and MxAPD respectively) between directly and indirectly seasonally adjusted estimates for each series, calculated over the span 2001Q4 to 2012Q2 (43 observations) as:

$$MeAPD = \frac{1}{43}\sum_{t=1}^{43}\left|\left(\frac{DSA_t - ISA_t}{ISA_t}\right) \times 100\right|$$

$$MxAPD = Max\left|\left(\frac{DSA_t - ISA_t}{ISA_t}\right) \times 100\right|$$

$t = 1, 2, \ldots, 43$

Where:
DSA_t = directly seasonally adjusted estimate at time t
ISA_t = indirectly seasonally adjusted estimate at time t

Table 3: Magnitudes of differences between directly and indirectly seasonally adjusted estimates

United Kingdom

	MeAPD	MxAPD
Inflows to		
Employment	0.77	1.89
Unemployment	0.91	5.04
Inactivity	0.76	3.36
Outflows from		
Employment	0.43	1.22
Unemployment	0.62	3.27
Inactivity	0.75	3.16

Source: Office for National Statistics

There are clearly differences between the direct and indirect estimates, and these differences are typically larger for some series, and for some time periods, than others. Table 3 indicates that the largest MeAPD for inflows is into unemployment, whilst for outflows it is from inactivity. In terms of MxAPD, the pattern is similar for both inflows and outflows; the difference is relatively large for unemployment and – to a lesser extent – inactivity, and relatively small for employment. Had the direct and indirect estimates yielded only marginal differences for all six series, there would be little motivation to formally assess which of the two approaches is preferred; one might opt to use the indirect approach as it has the property of its constituent components being additive. However, as some of the differences are quite substantial, a formal comparison involving quantitative measures is required.

3. Quality Comparison of the Direct and Indirect Approaches to Seasonal Adjustment

For each of the six aggregate flows series, three dimensions of seasonal adjustment quality were used to compare the direct and indirect approaches. These dimensions are:

1. residual seasonality
2. revisions
3. roughness

The measures associated with each of these three dimensions were produced using X-12-ARIMA, and are described in detail in Findley et al (1998), ONS (2007) and US Census Bureau (2011).

3.1 Residual seasonality

Residual seasonality can be defined as seasonality present in a time series after seasonal adjustment has been performed. This definition makes it an obvious indicator of seasonality adjustment quality, because a completely seasonally adjusted series should be free of seasonality.

Residual seasonality was tested for in two ways. The first was an F-test for stable seasonality, performed on the seasonally adjusted data (after taking differences between successive values, in order to remove the trend in the series). This test is obtained from the classical one-way analysis of variance (ANOVA) approach, whereby total variance is broken down into a variance of the four quarterly mean levels, due to seasonality, and a residual variance (Ladiray and Quenneville 2001).

For each of the six aggregate flows series, two null hypotheses were tested:

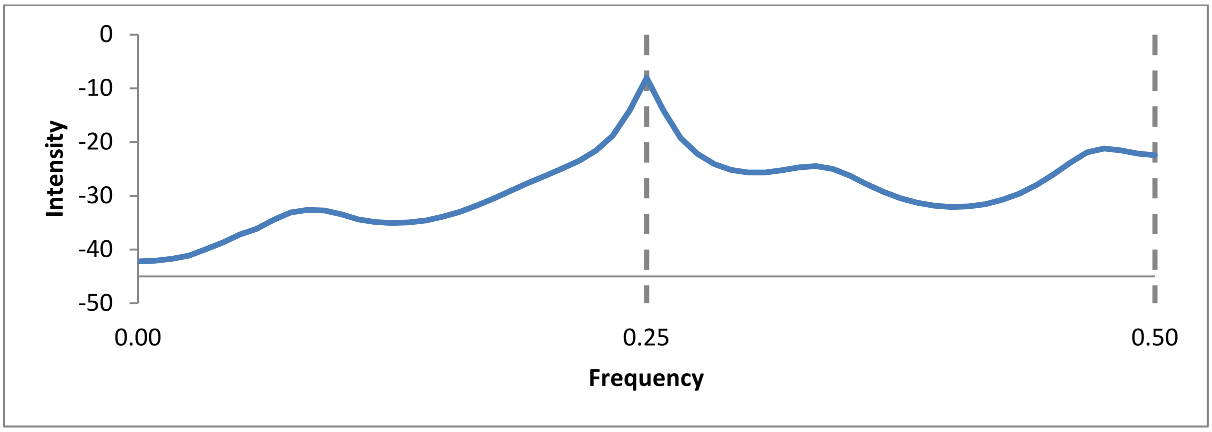
- H_{01}: no residual seasonality in the entire series
- H_{02}: no residual seasonality in the last three years of the series (reflecting the relative importance generally placed on more recent observations)

Neither of these null hypotheses was rejected for any of the six series at the one per cent significance level.

The second test for residual seasonality was visual inspection of the spectral plot for the seasonally adjusted data. If a quarterly time series is seasonal, one would expect the observation in the current time period to be related to that four periods ago, the observation in the previous time period to be related to that five periods ago, and so on. Such a series is said to exhibit cyclical variation with a “period” of four quarters, or a “frequency” of 0.25 cycles per quarter (so that the cycle repeats itself every four quarters). The estimated spectrum is a plot of intensity by frequency. The total area under the curve represents the variance of the series, so a peak at a certain frequency marks an important contribution to the variance by a cyclical component of that frequency. Hence one would expect to see a peak in the estimated spectrum of a seasonal quarterly series at the frequency 0.25, but no such peak in the estimated spectrum of the corresponding seasonally adjusted data if the series has been completely seasonally adjusted. More details on spectral analysis of time series can be found in, for example, Nerlove (1964).

Figure 7 illustrates the estimated spectrum for non-seasonally adjusted inflows to employment. As expected, there is an obvious peak at the frequency 0.25 cycles per quarter, indicating the presence of seasonality. Figures 8 and 9 illustrate the estimated spectra for directly and indirectly seasonally adjusted inflows to employment respectively. Neither of the estimated spectra have a peak at the seasonal frequency, corroborating the result obtained from the F-tests: that there is no evidence of residual seasonality in either of the seasonally adjusted estimates. Spectra for the inflows to employment series have been arbitrarily selected for illustration, as the results are representative of seasonal adjustments for the other five aggregate flows series (whose spectra are not shown here).

Figure 7: Estimated spectrum for non-seasonally adjusted inflows to employment



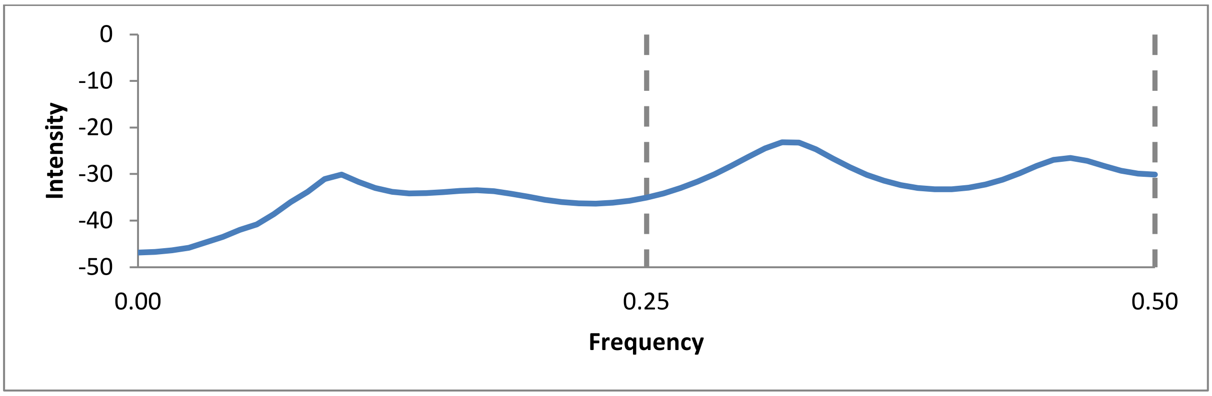
Source: Office for National Statistics

Notes:

1. The spectrum has been estimated from the first differenced, log transformed, prior adjusted data

2. Intensity is presented on the decibel ($10\times\log_{10}$) scale

Figure 8: Estimated spectrum for directly seasonally adjusted inflows to employment



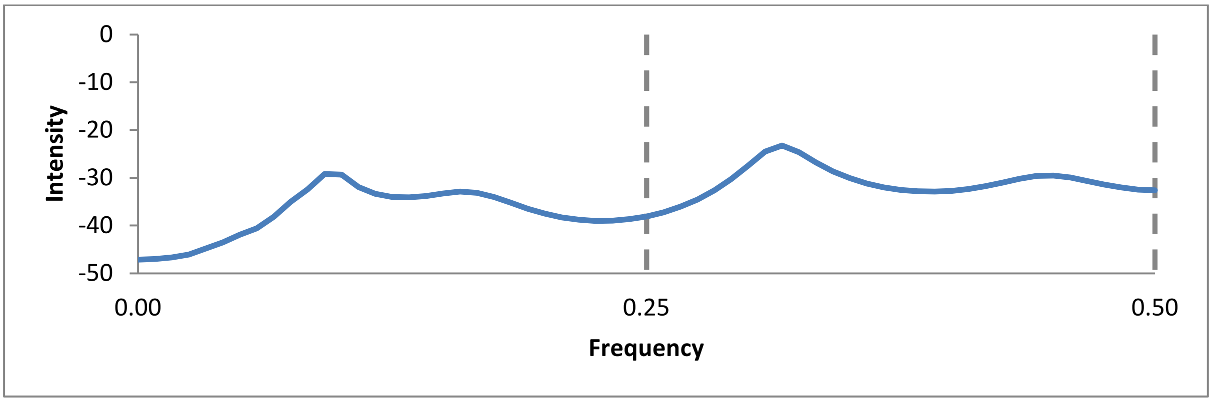
Source: Office for National Statistics

Notes:

1. The spectrum has been estimated from the first differenced, log transformed data

2. Intensity is presented on the decibel ($10\times\log_{10}$) scale

Figure 9: Estimated spectrum for indirectly seasonally adjusted inflows to employment



Source: Office for National Statistics

Notes:

1. The spectrum has been estimated from the first differenced, log transformed data

2. Intensity is presented on the decibel ($10\times\log_{10}$) scale

3.2 Revisions

A desirable characteristic of a seasonally adjusted series is stability. Whenever a new observation is added to the current end of a time series, and the series is then seasonally adjusted, there will be revisions to seasonally adjusted estimates, particularly in the most recent time periods. Seasonally adjusted estimates that are subject to relatively small revisions over time are said to be stable. On the other hand, seasonally adjusted estimates that are subject to relatively large revisions over time are said to be unstable. In the latter case, seasonally adjusted estimates may not be a reliable source of information for users, particularly if revisions result in changes in direction as well as magnitude.

An analysis of revisions history was performed for each of the six aggregate flows series in order to assess the stability of the directly and indirectly seasonally adjusted estimates. The process for each series is summarised below, and described visually in figure 10.

1. 2006Q4 was selected as the start date for the span to be analysed
2. the series was seasonally adjusted up to and including the start date
3. the start date was moved forward one time period and the series was seasonally adjusted again; this process continued until the end of the series was reached
4. the percentage revision between first and final seasonally adjusted estimates was calculated for each time period included in the analysed span (excluding the most recent period, for which only one seasonally adjusted estimate is available)
5. the mean and maximum absolute percentage revisions (MeAPR and MxAPR respectively) across all 22 time periods included in the analysed span were calculated as:

$$MeAPR = \frac{1}{22}\sum_{t=1}^{21}\left|\left(\frac{Final_t - First_t}{First_t}\right) \times 100\right|$$

$$MxAPR = Max\left|\left(\frac{Final_t - First_t}{First_t}\right) \times 100\right| \qquad t = 1, 2, \ldots, 22$$

Where:
$First_t$ = first seasonally adjusted estimate at time t
$Final_t$ = final seasonally adjusted estimate at time t

Figure 10: Overview of the process used for the revisions history analysis

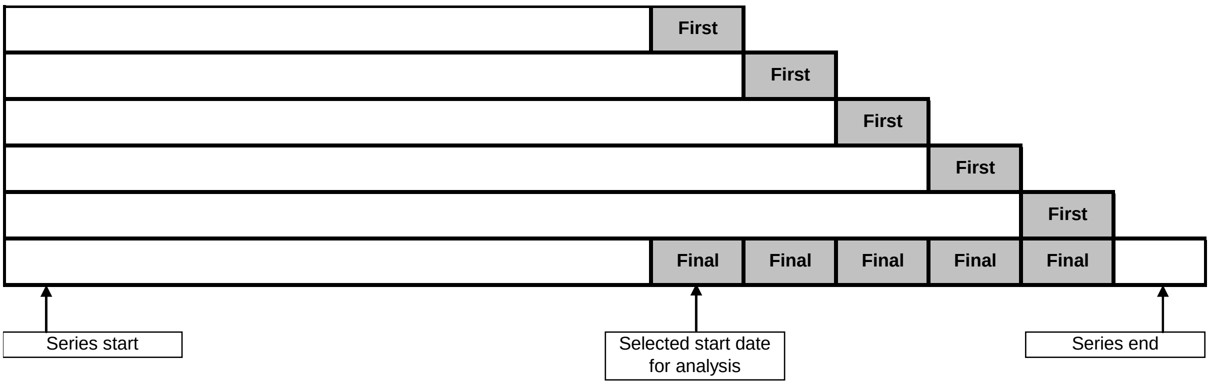


Table 4 summarises MeAPR and MxAPR values for the direct and indirect approaches for each of the six aggregate flows series. The results are mixed, and neither of the approaches can be said to be preferable. Whether considering MeAPR or MxAPR, each approach is preferred for exactly half of the series.

Table 4: Mean and maximum absolute percentage revisions between first and final seasonally adjusted estimates

	MeAPR		MxAPR	
	Direct	Indirect	Direct	Indirect
Inflows to				
Employment	1.21	1.29	3.57	3.08
Unemployment	0.43	0.83	1.26	2.18
Inactivity	1.13	0.90	3.71	2.73
Outflows from				
Employment	1.00	0.72	2.25	1.74
Unemployment	1.34	1.14	2.75	3.88
Inactivity	0.69	0.82	1.82	2.12

Source: Office for National Statistics

3.3 Roughness

In general, there is no requirement for a seasonally adjusted time series to be smooth. If the non-seasonally adjusted data are relatively volatile, then one should expect the seasonally adjusted data to also be volatile, as seasonal adjustment aims to remove the seasonal component – but not the irregular component – from the time series. However, smoothness (or roughness) can be a useful criterion for comparing competing seasonal adjustments.

For each of the six aggregate flows series, roughness is measured as the root mean square difference (RMSD) between the seasonally adjusted data and a smooth trend estimate, where the latter is obtained via a differencing operation. This measure is reported for the entire series and for the latest three years. A relatively large RMSD value indicates a relatively rough seasonally adjusted series, whilst a relatively small RMSD value indicates a relatively smooth seasonally adjusted series.

Table 5 summarises RMSD values for the direct and indirect approaches for each of the six aggregate flows series. The results are again mixed, and neither of the approaches can be said to be preferable.

Table 5: Root mean square differences between seasonally adjusted and trend estimates

	RMSD (entire series)		RMSD (latest three years)		Percentage change (%)	
	Direct	Indirect	Direct	Indirect	Entire Series	Last 3 years
Inflows to						
Employment	66537.635	66222.343	74144.975	73686.199	0.474	0.619
Unemployment	37116.597	40069.773	29107.154	32302.116	-7.956	-10.977
Inactivity	61800.624	61027.854	77935.740	77343.429	1.250	0.760
Outflows from						
Employment	78377.381	77990.028	90559.676	89069.141	0.494	1.646
Unemployment	47798.861	49004.201	55131.262	56619.386	-2.522	-2.699
Inactivity	45088.870	46845.652	41322.213	42824.227	-3.896	-3.635

Note: positive percentage changes indicate that the indirect seasonally adjusted composite is smoother than the direct seasonally adjusted composite.
Source: Office for National Statistics

4. Conclusion

It has been shown that time series within the employment flows dataset exhibit seasonality. Moreover, two approaches to seasonal adjustment have been investigated – direct and indirect estimation – and have been shown to produce different estimates. The two approaches have been compared in terms of three dimensions of seasonal adjustment quality: residual seasonality, revisions and roughness. On balance, neither of the approaches is obviously superior; there is no evidence of residual seasonality in any of the six aggregate flows series for either adjustment, and the results for the revisions and roughness dimensions are mixed. It is therefore recommended that the six aggregate flows series be seasonally adjusted via indirect estimation. This approach does not appear to produce seasonally adjusted estimates that are inferior to those of the direct approach, and guarantees that the sum of seasonally adjusted subcomponents is equal to the corresponding seasonally adjusted aggregate. This additivity property is potentially valuable from a user viewpoint, as it allows, for example, contributions to unusually large quarter-on-quarter changes in an aggregate series to easily be attributed to changes in its subcomponents.

References

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (2008), *Time Series Analysis: Forecasting and Control*, Fourth Edition, Wiley: New Jersey.

European Central Bank (2003), *Seasonal Adjustment*, Edited by Manna, M. and Peronaci, R., European Central Bank: Frankfurt am Main, available at: www.ecb.int/pub/pdf/other/statseasonaladjustmenten.pdf

Eurostat (2009), *ESS Guidelines on Seasonal Adjustment*, Office for Official Publications of the European Communities: Luxembourg, available at: http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-09-006/EN/KS-RA-09-006-EN.PDF

Findley, D.F., Monsell, B.C., Bell, W.R., Otto, M.C. and Chen, B.-C. (1998), "New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program", *Journal of Business and Economic Statistics*, No. 16, available at: http://www.census.gov/ts/papers/jbes98.pdf

Ladiray, D. and Quenneville, B. (2001), *Seasonal Adjustment with the X-11 Method*, Lecture Notes in Statistics, No. 158, Springer-Verlag: New York.

Nerlove, M. (1964), "Spectral Analysis of Seasonal Adjustment Procedures", *Econometrica*, Vol. 32, No. 3 (July 1964), pp. 241-286.

ONS (2007), *Guide to Seasonal Adjustment with X-12-ARIMA*, Draft Version, available at: http://www.ons.gov.uk/ons/guide-method/method-quality/general-methodology/time-series-analysis/guide-to-seasonal-adjustment.pdf

ONS (2012), Labour Force Survey, available at: http://www.ons.gov.uk/ons/guide-method/method-quality/specific/labour-market/labour-market-statistics/index.html

US Census Bureau (2011), *X-12-ARIMA Reference Manual*, Version 0.3, US Census Bureau: Washington DC, available at: http://www.census.gov/ts/x12a/v03/x12adocV03.pdf