

ARTICLE

Jamie Jenkins and Mark Chandler
Office for National Statistics

Labour market gross flows data from the Labour Force Survey

SUMMARY

The article investigates the use of longitudinal datasets from the Labour Force Survey (LFS) as a tool for assessing the UK labour market. The main points are:

- LFS respondents are interviewed for five consecutive quarters. Therefore longitudinal datasets can be produced which link respondents over two and five quarters
- these datasets can be used to analyse the movement, or flow, between employment, unemployment and inactivity
- the gross flows between each labour market status every quarter tend to be much larger than the quarterly changes in the headline labour market aggregates
- longitudinal datasets are known to be subject to a number of biases

Introduction

The Labour Force Survey (LFS) is a quarterly sample survey of around 52,000 households living at private addresses in the UK, representing about 0.2 per cent of the population. The survey gathers information on a wide range of labour market characteristics and related topics. The survey has a rotating sample design, with the same households interviewed for five consecutive quarters and the last interview being a year after the first. The rotating design means there is replacement of one-fifth of households each quarter.

The main use of the survey is to provide cross-sectional data, but with households sampled consecutively, it is possible to link responses to provide longitudinal data, which is useful in identifying how individuals' economic circumstances change over time. The Office for National Statistics (ONS) produce two types of longitudinal datasets, linking respectively two and five quarters of data. These datasets include all respondents of working age, that is men aged 16 to 64, and women aged 16 to 59.

One of the key uses of these datasets is to analyse the gross flows of people moving between different labour force categories, most notably the three economic activity groupings of employment, unemployment and inactivity. Gross flows are the total number of people moving, for example from employment (E) to unemployment (U), or inactivity (N), and the total

number of people who move in the opposite direction. In total there are nine different flow categories for the three economic activity groupings, with another two – those individuals entering or leaving working age. Looking at the gross flows from one group to another, for example from employment to unemployment (EU), and from unemployment to employment (UE), shows more information than the net flow. This is the difference between the total number of people employed and unemployed at two different times. Each month, ONS publish net changes in the stocks of people in each of the three economic activity categories, and while these changes may be small, the underlying gross flows are much larger.

This article uses the two-quarter longitudinal data back to 1997, which is the first point at which consistently weighted time series are available. Firstly, it briefly describes the method used to produce the datasets, describes their limits, and then shows various analyses possible using the datasets.

Producing datasets

For both the two and five quarter longitudinal datasets, matching takes place using a unique identification variable for each individual for each quarter, with the age and date of birth of the matched cases checked to ensure correct matches. All unmatched individuals, as well as those individuals who do not respond but their data is carried forward in the cross-sectional dataset, or those individuals

where no information is available on their economic activity in any of the quarters are then removed from the dataset.

Two methodological issues arise when linking data and producing results. First, there are biases that arise from non-response, as different groups of people have different likelihoods of dropping out of the survey between interviews, which means they are not available for linking. Second, there are biases arising from response errors in the data, which produce false flows between economic states.

Compensating for non-response bias

There are many reasons for losing sample members after the first interview, either because they have moved house, it is not possible to contact them, or they refuse to continue to take part. As these groups of people are not typical of the general population, losing them will introduce bias.

Non-response is higher for young adults, single people, those living in London, those living in privately rented accommodation, the unemployed, and those in temporary employment. A multiple stage weighting method that initially constrains the weights by housing tenure to the cross-sectional datasets, then constrains the weights to population totals, and finally calibrates the weights using age, sex, region and economic status, is used to minimise this bias.

Response error bias

All surveys are susceptible to response errors, which is where the information from the respondent is not accurate. This could be for many reasons, such as misunderstanding the question, mis-recording by the interviewer, lack of knowledge, or through knowingly giving false information. Additionally, if a respondent is going through a transition at the time of the survey, they may give a false state by answering about the wrong reference week.

International research suggests that for the questions on economic activity status, the errors are not systematic and so when looking at the cross-sectional datasets, they cancel out. However, when linking individuals, such errors will lead to a change of economic activity status which is false, and therefore exaggerate and bias upwards the gross flows between each of the states. It is almost impossible to identify respondent errors directly, unless carrying out the interview again, but then that would raise doubt to the

correct answer. Investigations from other countries using re-interview surveys show that bias can vary considerably for each of the flows and to estimate the bias in the longitudinal datasets would need a UK re-interview survey.

There are many issues when trying to design a re-interview survey to measure the bias with a high-level of reliability. Linking the LFS to the 2001 Census suggests the development of such a survey would need a sample of over 10,000. Interviewing more individuals will increase respondent burden, and having a further interview may also increase attrition for the main survey.

Also the Census is only available at one particular time point (April), and given the seasonality of the flows between economic activity states, it is possible there may be seasonal variation in the misclassification. Overcoming seasonality is possible by using a four-quarter moving average, but these estimates will still be prone to response errors. However, assuming the response errors remain consistent over time, it is still possible to draw valid conclusions from the time series of changes between economic activity states. More information on estimating these response errors can be found in Brook and Barham (2006).

Analysis of the datasets

For analysis of the longitudinal datasets, the stock of the employed, unemployed and inactive at each quarter need to be obtained by summing the three corresponding flow categories. Therefore for the last quarter in the dataset, the stock of those who are:

- **employed** – is given by summing those employed at both quarters

(EE), those who move from unemployment to employment (UE) and those moving from inactivity to employment (NE)

- **unemployed** – is given by summing those moving from employment to unemployment (EU), those remaining unemployed (UU) and those moving from inactivity to unemployment (NU), and
- **inactive** – is given by summing those moving from employment to inactivity (EN), those moving from unemployment to inactivity (UN) and those remaining inactive (NN)

The stocks derived from the longitudinal datasets differ slightly to those obtained from a quarterly cross-sectional dataset due to attrition and also because those who are entering or leaving working age are excluded.

The unemployment rate derived from the flow variable, which is expressed as a percentage of working age, will also be lower and is not consistent with the unemployment rate given in ONS's Labour Market statistical bulletin. This is defined, according to International Labour Organisation definitions, in terms of all adults aged 16 and over as a percentage of the economically active (employed and unemployed) population.

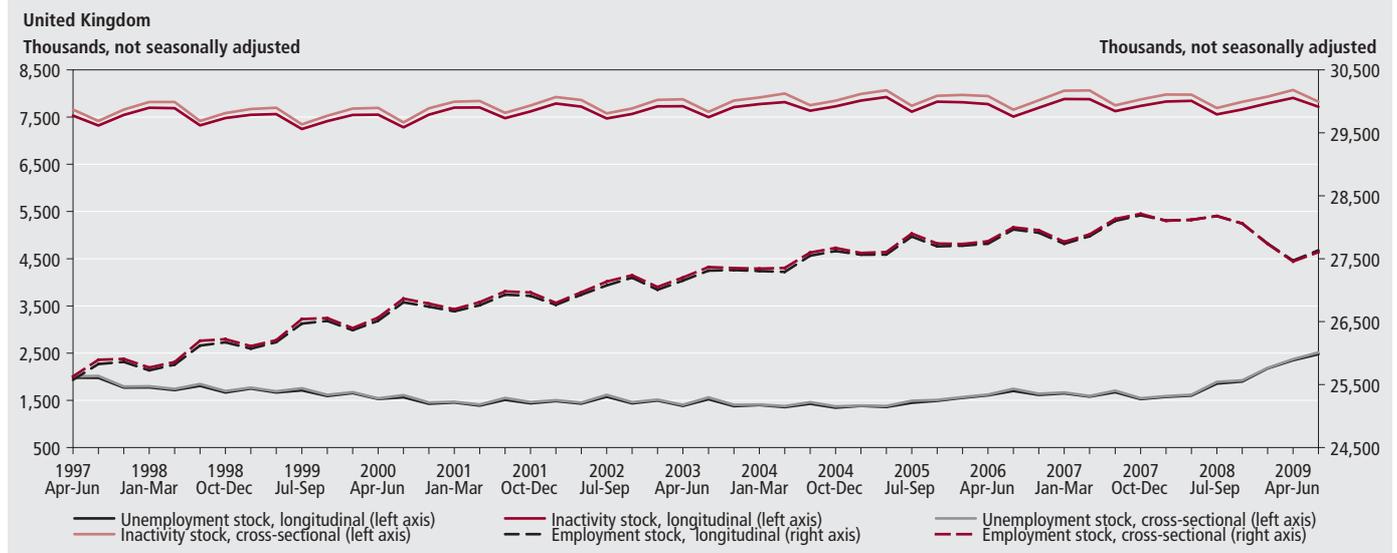
Table 1 shows the weighted gross flow for each category and the associated sample size. Looking at the proportions of people who have remained in the same economic activity status between the two LFS interview periods, it is clear that the vast majority of people do not change their status between waves, that is they remain in employment, unemployment or inactivity over both quarters. In 2005 the proportion was 92 per cent and in

Table 1
Categories of flows, Q2 2009 to Q3 2009

United Kingdom				
Status in first quarter	Status in second quarter	Sample Size	Weighted gross flow (thousands)	Gross flow (per cent)
Aged 15	Working age	206	202	0.5
Employed	Employed	24,342	26,561	69.6
Employed	Unemployed	332	413	1.1
Employed	Inactive	384	453	1.2
Unemployed	Employed	438	555	1.5
Unemployed	Unemployed	1,164	1,432	3.8
Unemployed	Inactive	312	386	1.0
Inactive	Employed	430	518	1.4
Inactive	Unemployed	511	632	1.7
Inactive	Inactive	6,416	6,878	18.0
Working age	Above working age	189	152	0.4
Total		34,724	38,182	100.0

Source: Labour Force Survey

Figure 1
Comparison of labour market stocks from cross-sectional and longitudinal datasets



Source: Labour Force Survey

2009 it was 91 per cent. Despite this, one of the strengths of the longitudinal datasets is their ability to look at some of the transitions between different states, as these can give some insights into the flows which contribute to the overall numbers in each state.

Time-series comparison

Figure 1 shows a comparison over time of the stock of employed, unemployed and inactive people of working age estimated from the LFS cross-sectional datasets and from the two-quarter longitudinal datasets. It can be seen that the estimates for employment and unemployment from both datasets are broadly in agreement. The wider difference between the measures of inactivity is due to the absence of those entering working age in the latter quarter, as this cohort is more likely to be inactive than those leaving working age. This is noticeably wider during the seasonal peaks.

The seasonal effects apparent in the LFS cross-sectional datasets equally apply to the longitudinal datasets. Typically, there is a significant shift of people from economic inactivity to activity during the third quarter of each year, partly due to those leaving full-time education, with the reverse happening in the first quarter of each year.

These seasonal effects can make measuring the individual gross flows from one quarter to the next difficult. Therefore, it is recommended that users present time series data using a four-quarter moving average. This has been adopted for the charts present in this article.

Gross flows

Figure 2 shows the estimated gross flows, that is the total inflow or outflow for working age employment, unemployment and inactivity from one calendar quarter to the next. The period covers April–June (Q2) 2009 and July–September (Q3) 2009. The stocks for each status represent the latter period and are the non-seasonally adjusted aggregates for the working age population, as published in the Labour Market statistical bulletin.

Comparing these gross flows to the published quarterly changes in the headline LFS aggregates reveal how

substantial the underlying movements hidden behind these values are.

Hazard rate

The hazard rate is a measure of the probability that an individual will change status over the quarter. This is calculated by taking the gross flow from the second quarter as a percentage of the total stock from the previous status in the first quarter. Table 2 shows that between 2000 and 2007, on average, 30 per cent of those who were in unemployment in the first quarter had moved into employment by the next quarter. By 2009 it was 23.4 per cent.

Figure 2
Quarterly working age¹ population flows, Q2 2009 to Q3 2009²



Notes:

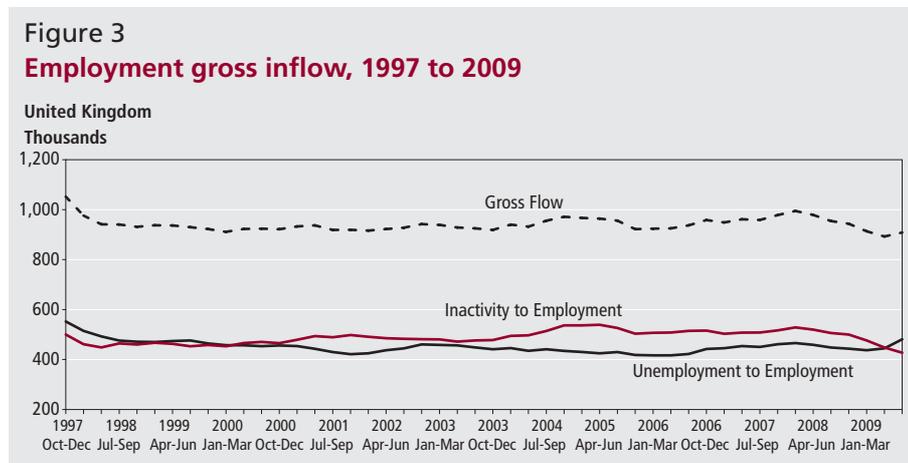
- 1 Men aged 16–64 and women aged 16–59.
- 2 The stocks included in the diagram are the published non-seasonally adjusted estimates: www.statistics.gov.uk/statbase/product.asp?vlnk=8292. These are the preferred source and differ from the stock estimates derived from the longitudinal dataset.

Source: Labour Force Survey

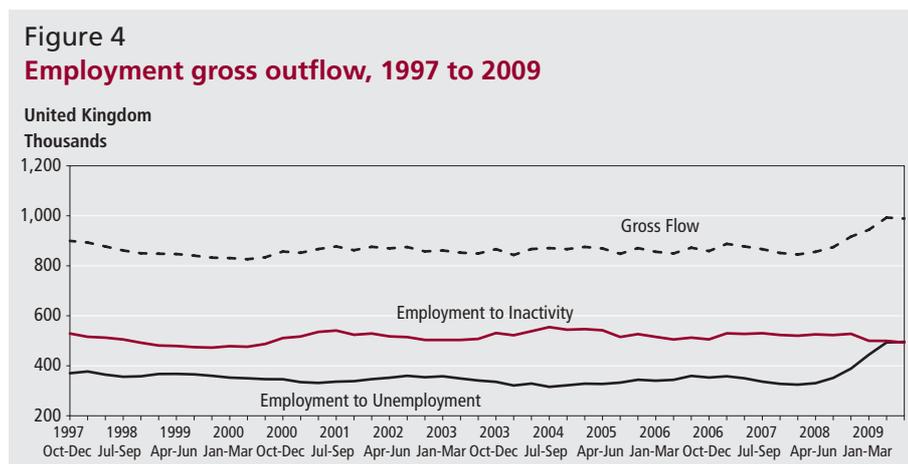
Table 2
Annual hazard rates, 1997 to 2009

United Kingdom									
	Per cent, Q3 each year								
	Remain in employment	Employment to Unemployment	Employment to Inactivity	Remain in unemployment	Unemployment to Employment	Unemployment to Inactivity	Remain in Inactivity	Inactivity to Employment	Inactivity to Unemployment
1997	96.5	1.5	1.8	57.0	26.3	16.5	85.2	7.5	6.0
1998	96.8	1.3	1.7	56.9	26.5	16.5	84.4	8.2	6.1
1999	96.9	1.3	1.6	55.7	27.9	16.3	85.1	7.8	5.7
2000	96.9	1.2	1.7	53.4	29.6	16.8	85.3	8.0	5.5
2001	96.7	1.3	1.8	53.3	28.6	18.0	86.0	7.6	5.4
2002	96.6	1.4	1.8	53.3	29.9	16.6	85.9	7.5	5.5
2003	96.7	1.3	1.8	53.1	28.6	18.2	85.7	7.7	5.5
2004	96.7	1.1	2.0	48.8	30.9	20.2	85.0	8.5	5.8
2005	97.0	1.1	1.6	47.9	32.3	19.7	85.4	7.8	5.9
2006	96.6	1.3	1.7	52.8	28.8	18.3	84.7	8.2	5.9
2007	96.8	1.1	1.8	52.9	28.5	18.4	84.9	8.1	6.3
2008	96.6	1.4	1.7	54.0	25.3	20.8	84.5	7.5	7.3
2009	96.7	1.5	1.7	60.3	23.4	16.3	85.3	6.4	7.8

Source: Labour Force Survey



Source: Labour Force Survey



Source: Labour Force Survey

The transitions need to be taken in context relative to the stocks for each series. For example, 1.5 per cent moved from employment to unemployment in Q3 2009 which corresponds to approximately 400,000 people. This number is similar in size to the 16.3 per cent of people moving from unemployment to inactivity.

Employment

The gross inflow to employment (Figure

3) has been falling since the peak in Q1 2008. Gross inflows to employment from inactivity have been driving this primarily. The inflow from unemployment picked up sharply in the latest quarter, surpassing the inflow from inactivity for the first time since 1999.

The gross outflows (Figure 4) to unemployment show a sharp increase from Q3 2008 that has levelled off in recent quarters. The series has reached a

similar level to the outflow to inactivity which has been broadly flat throughout the past 3 years.

Between 2002 and 2005 patterns of declining flows from unemployment to employment, and increasingly flows from inactivity to employment were developing. Although the likelihood of moving from unemployment to employment is higher than moving from inactivity to employment (see *hazard rates*), the stock of people who are inactive is much greater than the stock of those who are unemployed. Since 2007 the two lines have been converging, meaning that although the rates of flow are different, in reality there is only a relatively small difference in the numbers of people entering employment from both unemployment and inactivity.

Unemployment

Nearly one million people moved into unemployment between Q2 2009 and Q3 2009 (Figure 5). Between 2003 and 2008, inactivity was the main driver, growing steadily since 2004. This is an effect of an increased transition from a larger pool of inactive people.

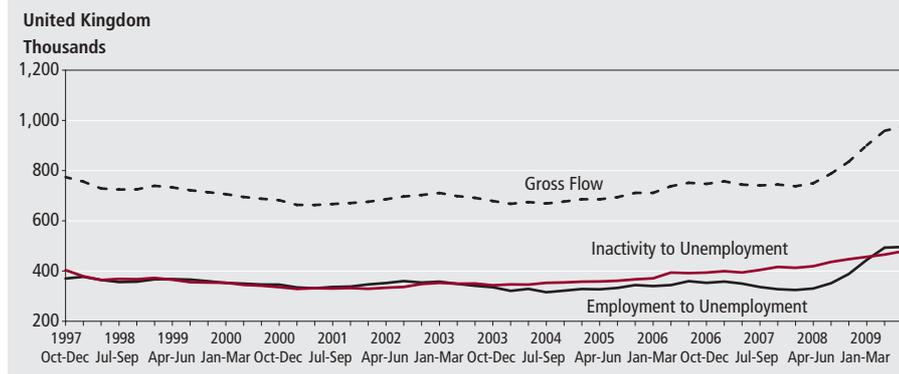
The flow from employment between 1997 and 2008 was consistently below 0.4 million and reached a low prior to the sharp increase in 2008.

Figure 6 shows an outflow from unemployment that has exceeded 0.8 million. This outflow is significantly higher than in the same period a year ago, and is mainly driven by the flow to employment. The flow to inactivity has been on an upward trend since 2006.

Inactivity

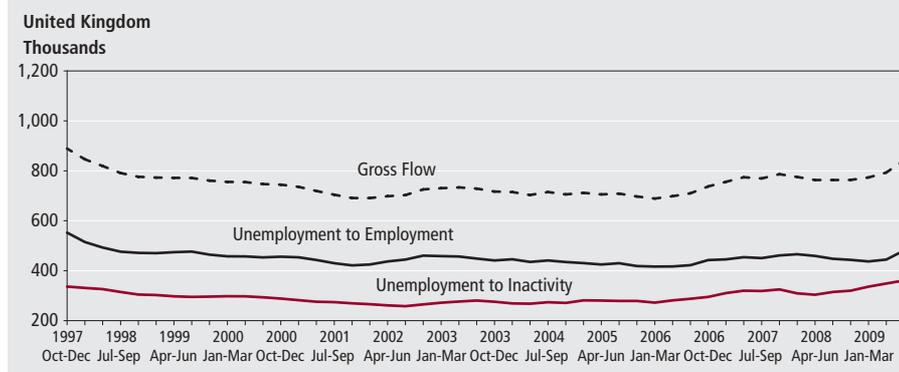
The gross inflow to inactivity was has remained over 0.8 million (Figure 7). This

Figure 5
Unemployment gross inflow, 1997 to 2009



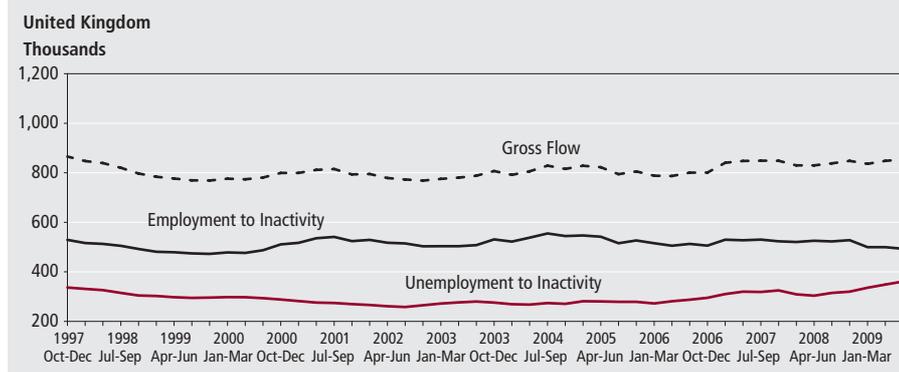
Source: Labour Force Survey

Figure 6
Unemployment gross outflow, 1997 to 2009



Source: Labour Force Survey

Figure 7
Inactivity gross inflow, 1997 to 2009



Source: Labour Force Survey

is broadly similar to the inflow for the same period a year ago; however the main driver has switched from employment to unemployment in the past year.

Gross outflows (Figure 8) from inactivity to employment have been lower than the same time a year ago. Outflows to unemployment are now driving the overall outflow, rather than the flow to employment, as in the past. Total outflows are lower than a year ago.

Students form the largest group moving out of inactivity and account for over half

of all those moving into employment. Students also contribute largely to the seasonal movement of the flow to unemployment. It is clear from comparing the two outflows that people transitioning into the labour market in the latest quarter are now more likely to be moving into unemployment.

Conclusion

Understanding the quarterly changes in the levels of employment, unemployment and economic inactivity is aided by an

analysis of changes in the labour market status of respondents to the LFS. For this reason, the two-quarter longitudinal datasets have been used. The sample is also more robust and less subject to sampling variation than the five-quarter counterparts. These datasets are weighted using the same population estimates as those used in the main quarterly LFS datasets, although the weighting methodology differs due to response bias. Consequently the estimates are broadly but not entirely consistent with the published aggregates. Also, the datasets are limited to working age people.

The aim of this article has been to give some background to how the longitudinal datasets are produced and to show some of their uses, looking at gross flows over time between the states of employment, unemployment and inactivity. The patterns shown here generally reflect the changes which have been seen in the recent cross-sectional figures produced from the LFS, in terms of an increase in unemployment and the decline in employment. In addition they give some insight into the size of the movements between categories, especially the continued increase in the transitions the economically inactive make into the labour market.

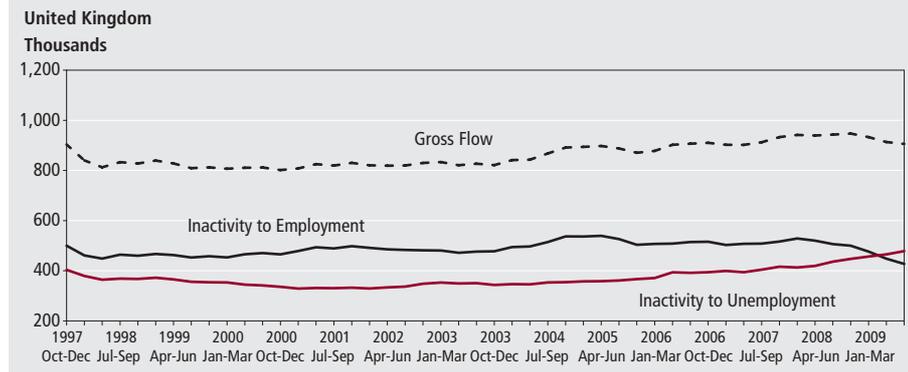
In response to user interest the quarterly results from the two-quarter longitudinal datasets are now included in Annex 4 of the *Labour Market Overview* (available at www.statistics.gov.uk/downloads/theme_labour/LMS_QandA.pdf). These are classed as experimental statistics, due to some methodological issues and as such, are not an accredited National Statistic. Further work by ONS will consider options to improve the method of producing the datasets, along with extending the coverage in response to the change in female state pension age (see Clegg, Leaker and Kent 2010). A further analytical article, which will utilise the dataset and look in more detail at the recession, will be produced later in the year.

Users considering independent analysis of the datasets need to consider the sample size and seasonality of the flow, especially when conducting micro-level analysis. It should be noted, however, that estimated changes over time in the gross flows should be largely free from any response bias.

CONTACT

✉ elmr@ons.gov.uk

Figure 8
Inactivity gross outflow, 1997 to 2009



Source: Labour Force Survey

REFERENCES

Brook K and Barham C (2006) 'Labour market gross flows data from the Labour Force Survey', *Labour Market Trends* July 2006. Available at www.statistics.gov.uk/articles/labour_market_trends/gross_flows.pdf

Clegg R, Leaker D and Kent K (2010) 'Implications of the change in female state pension age for labour market statistics', *Economic and Labour Market Review* January 2010. Available at www.statistics.gov.uk/elmr/01_10/downloads/ELMR_Jan10_Kent.pdf