Article

Monitoring the quality of Labour Force Survey results

Further detail about the work currently carried out to monitor the robustness of Labour Force Survey (LFS) results. The LFS is the source for estimates of employment, unemployment and economic inactivity.

Table of contents

1 . Author

Mark Chandler

2 . Executive summary

This article details the work currently carried out to monitor the robustness of Labour Force Survey (LFS) results. In summary:

- the methodology used to weight the LFS helps maintain robust headline estimates

- our analysis of interview patterns over the last 15 years indicates little change in the range and coverage of households surveyed

- sampling variability, indicated by a confidence interval, for both the employment rate and unemployment rate have remained stable over the past 10 years and were plus or minus 0.4 percentage points and plus or minus 0.3 percentage points respectively in 2016

3 . Introduction

This article highlights the activities undertaken to monitor the quality of Labour Force Survey (LFS) results and addresses important positive steps undertaken to address falling response rates to social surveys, including the LFS. This article looks at five areas:

- how the methodology is designed to help maintain robust headline estimates

- a review of the analysis work carried out as part of the regular production round to ensure results are robust

- a comparison of the LFS with other data sources

- identifying areas where LFS results are vulnerable to falling response rates

- an overview of the work planned and underway, with a discussion on how administrative data could be used to improve the robustness of LFS results

4 . Main approaches to address the robustness of the Labour Force Survey

4.1 Methodology

Each quarter's Labour Force Survey (LFS) sample of 50,000 households is made up of five sets of interviews or "waves", each of approximately 10,000 households. Each wave is selected for interview in five successive quarters. This means that in any one quarter, one wave will be receiving their first interview (wave 1), typically held face-to-face with the interviewer at the respondents home, one wave (wave 2) will be receiving their second interview and so on, with wave 5 receiving their fifth and final interview. Waves 2 to 5 are typically carried out over the telephone.

The LFS is a sample survey and requires population estimates in order to produce weights to ensure estimates from the survey are representative of the UK. The weighting of LFS results is updated every year, ensuring consistency with the most recent Office for National Statistics (ONS) population estimates and projections. The weighting is designed to account for non-response across several detailed dimensions (age, gender and area of residence). Seasonally adjusted LFS aggregates are constrained to ensure they add up to the population total. The size of the constraining adjustments has remained small (usually less than 0.04%), indicating a degree of consistency in the underlying data. Further information on the methodology of the LFS is available.

4.2 Regular analysis

Monthly waves and cohorts

Using the Labour Force Survey (LFS) microdata, Office for National Statistics (ONS) breaks the LFS aggregates down into building blocks (for example, based on the month they were interviewed and whether it was their first or one of their subsequent interviews), and examines them individually to ensure we understand their respective impacts on the overall aggregates. This exploits the wave structure of the survey, as discussed in the previous section, as an analytical tool for quality assurance. These estimates are published each month in Table X01.

The patterns in the disaggregated analysis have remained quite similar over the last 15 years, indicating little change in the range and coverage of households surveyed. For example, a cohort with an unusually high unemployment rate seems to be no more likely to enter the sample now than in the past and the employment rate generally increases between the first and second interview for any specific cohort, such as the change in collection mode from face-to-face to telephone.

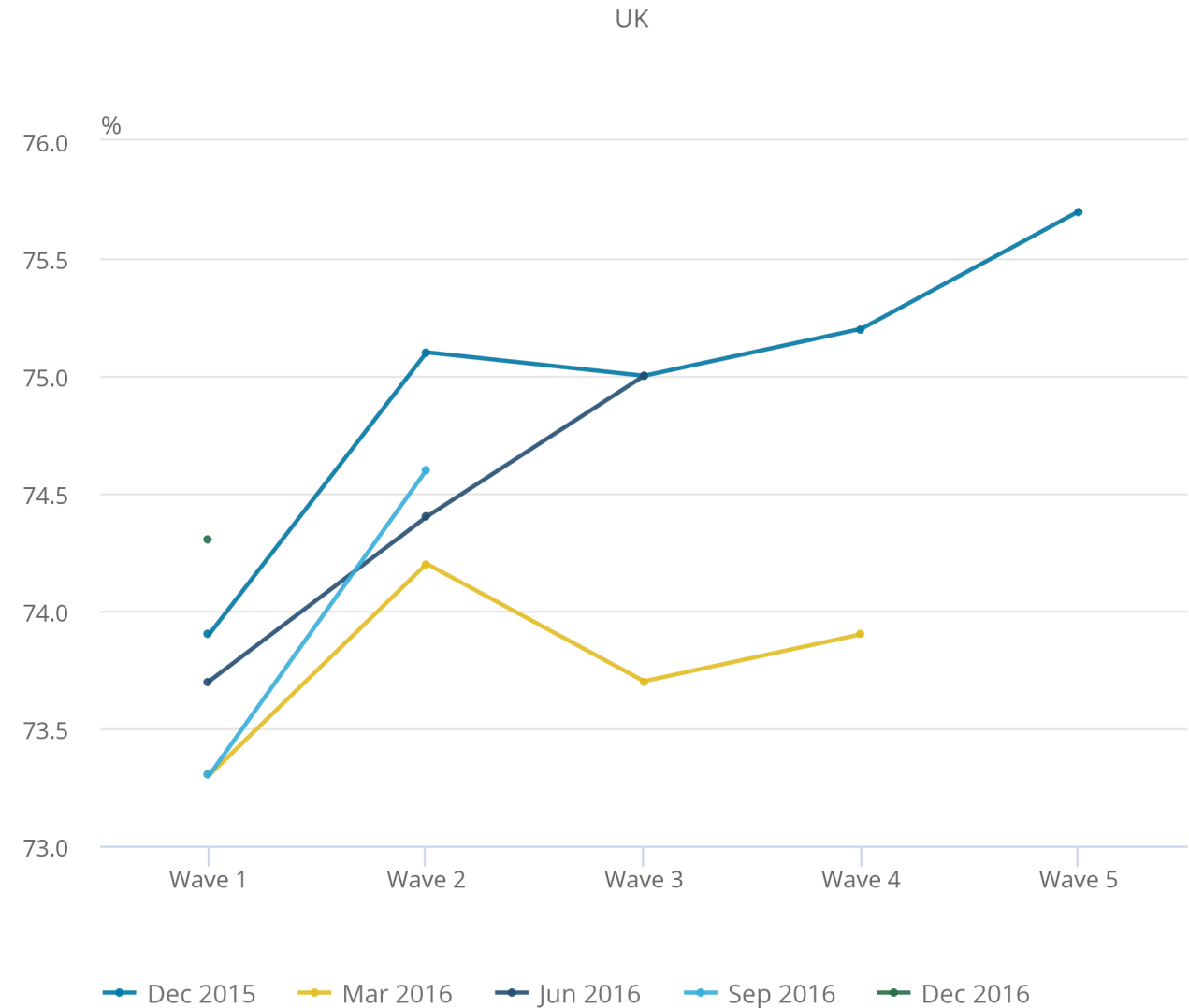### Wave and cohort analysis for employment from the LFS

Each line (Figures 1 and 2) represents a single cohort and tracks its employment rate from the month of first interview (wave 1) to final interview (wave 5), a year later. For example, all respondents interviewed in March 2017 will consist of households in their first interview (March 2017), second interview (first interviewed in December 2016), third interview (first interviewed in September 2016) and so on. These are shown for 2 months only, December 2016 and March 2017, but it is possible to track these back to 1992. This is important in identifying any bias that could have been introduced to the survey, as the relationships between transitioning waves should fit a regular pattern.

The key shown in Figures 1 and 2 indicate the month in which each cohort received its first (wave 1) interview. The estimates are seasonally adjusted. Note that the cohorts are not identical at each wave as some people drop out of or join the survey as the cohort progresses through the waves. In general, though, the composition of each cohort remains much the same.

Figure 1: Employment rates, people aged 16 to 64, by wave (seasonally adjusted), December 2015 to December 2016

UK

Figure 1: Employment rates, people aged 16 to 64, by wave (seasonally adjusted), December 2015 to December 2016

UK

Source: Office for National Statistics

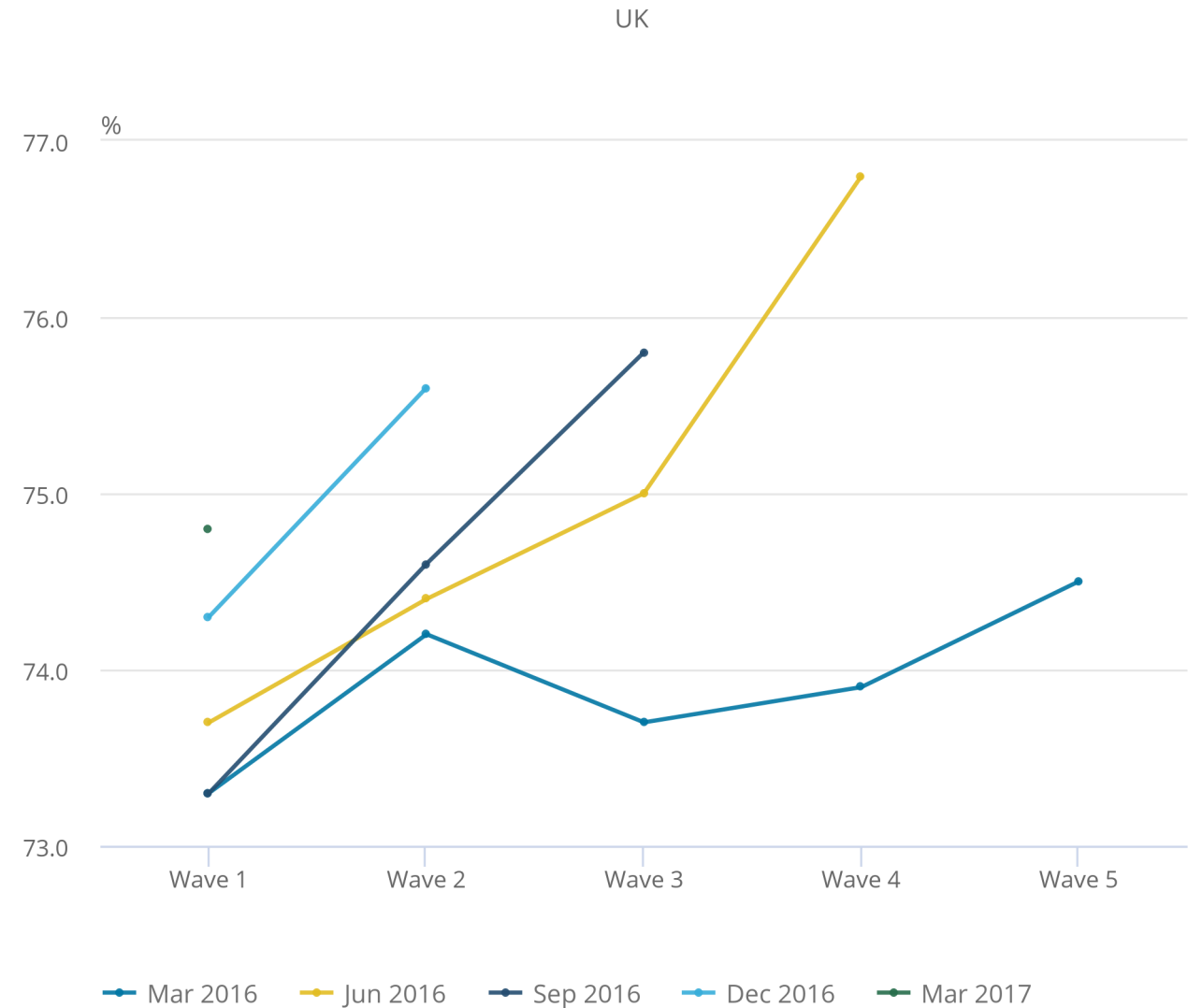**Source: Office for National Statistics**

Notes:

1. Each line represents a single cohort interviewed during December 2016 and tracks its employment rate from the month of first interview (wave 1) to final interview (wave 5), a year later. For example, all respondents interviewed will be in their first interview (December 2016), second interview (first interviewed in September 2016), third interview (first interviewed in June 2016) and so on. The final point of each line is the employment rate in the month being measured.

Figure 2: Employment rates, people aged 16 to 64, by wave (seasonally adjusted), March 2016 to March 2017

UK

Figure 2: Employment rates, people aged 16 to 64, by wave (seasonally adjusted), March 2016 to March 2017

UK

Source: Office for National Statistics

Source: Office for National Statistics

Notes:

1. Each line represents a single cohort interviewed during March 2017 and tracks its employment rate from the month of first interview (wave 1) to final interview (wave 5), a year later. For example, all respondents interviewed will be in their first interview (March 2017), second interview (first interviewed in December 2016), third interview (first interviewed in September 2016) and so on. The final point of each line is the employment rate in the month being measured.

Figures 1 and 2 show an example of the cohort analysis routinely carried out. In Figure 2, the new cohort that entered the LFS at wave 1 in March 2017 had an employment rate of 74.8%. The wave 5 value for the cohort it replaced received their wave 5 interview in December 2016 and had an employment rate of 75.7% (Figure 1). Investigations of this wave from the LFS microdata show that this difference mainly represented a random sampling effect (that is, it is within the sampling error) and with the anticipated mode change effect, a small increase in subsequent waves is expected. The difference between these two points is typical and similar to relationships between previous waves 1 and 5.

Figure 2 also shows the transition between wave 1 and wave 2 for the cohort that joined in December 2016 and the transitions between waves 2 to 3, and so on. The movement between waves 3 to 4 stands out; the increase between waves is strong relative to other wave 3 to 4 movements (Figure 1, for example), with the June 2016 cohort employment rate at 76.8% at wave 4. This is supported by the September 2016 cohort also showing a strong increase between waves 2 to 3 and with the March 2016 cohort, containing respondents with the lowest employment rate, also experiencing an increase since their last (wave 5) interview.

The conclusion to be drawn is that the growth in March 2017 was mainly due to an increase in the reporting of employment by respondents in follow-up interviews in their third, fourth and fifth waves during this period, rather than a transition effect between the departing wave 5 and the new wave 1. A sustained period of new cohorts entering the survey that greatly differ from those exiting the survey could indicate potential sample changes that may impact quality. Highly erratic movements between waves could signify attrition-related issues.

Estimates presented here and published in Table X01 are Experimental Statistics and are not designated as National Statistics. However, the production and evaluation of the estimates is an important part of our quality assurance of the 3-monthly averages published in the UK labour market statistical bulletin.

Seasonal adjustment

Results are also subject to "sense checks" to ensure that factors that we would expect to impact on results do so. As an example, analysis of the data collected on actual hours worked by individual weeks continues to identify regular effects, such as school and bank holidays, to the same extent as the data has done in the past. This is reassuring and helps to ensure optimal seasonal adjustment. Additionally, the hours worked data behaves as expected with irregular bank holidays (for example, the Royal Wedding and the Diamond Jubilee), with a small reduction when those occur.

Seasonal adjustment of the LFS is subject to a regular annual review to ensure whether any unusual occurrences, such as questionnaire changes to the LFS, have had an effect on the time series and to determine if the existing seasonal adjustment is optimal. Seasonality in the data has continued to be well-defined generally, with no significant discrepancies or sudden changes in patterns. We decompose the seasonal adjustment every month into its component parts and investigate any irregularities. This is vital to understand whether data movements are due to outliers in the survey data or because labour market conditions have changed. All estimates of employment and unemployment are published on a seasonally adjusted and non-seasonally adjusted basis for users.
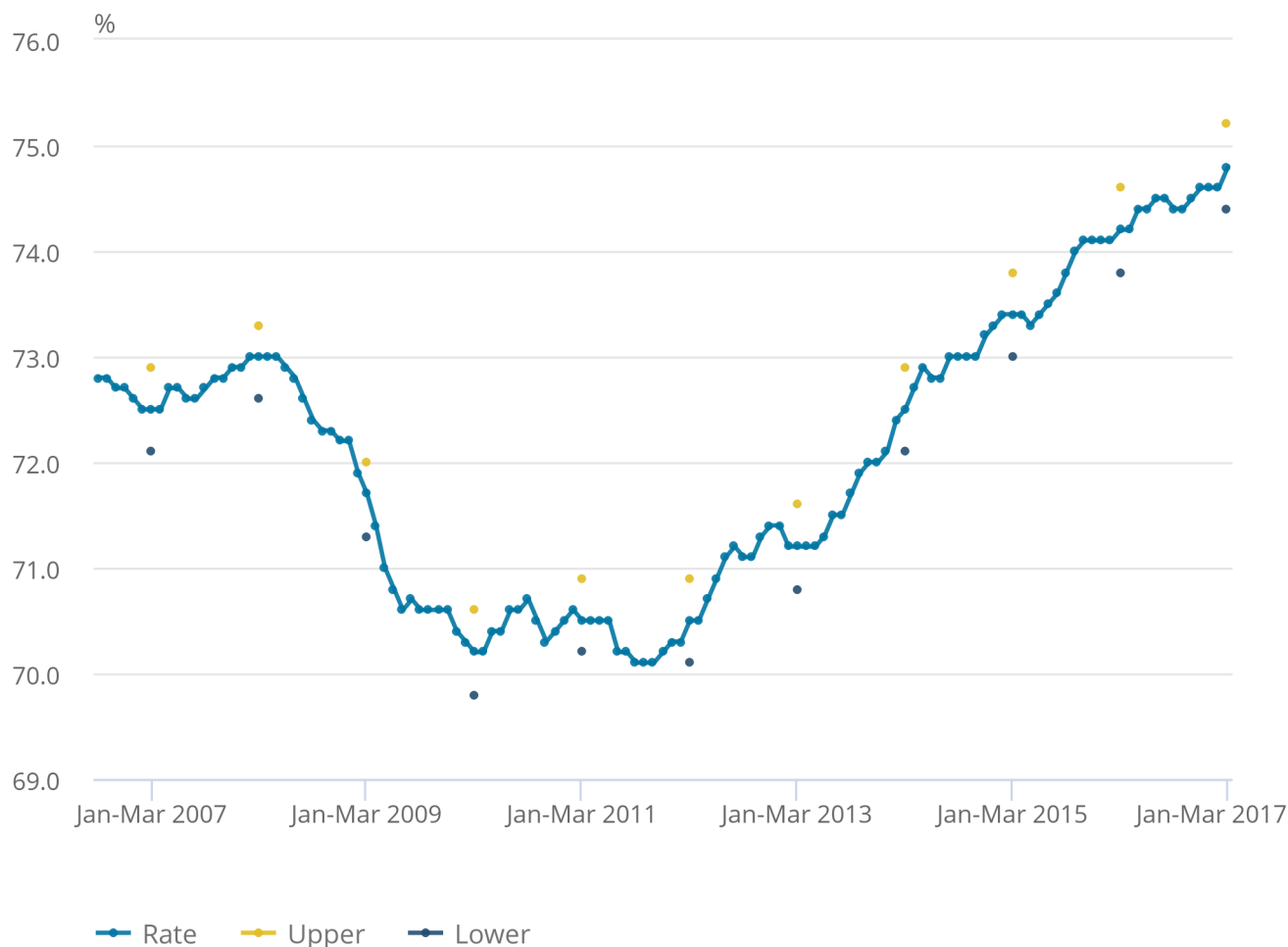
Sampling variability

Standard errors are calculated every month for the headline series for levels, rates and changes. Office for National Statistics (ONS) produces estimates of the 95% confidence interval for all headline LFS estimates of UK employment, unemployment, economic inactivity and hours worked. These are published as part of the monthly Labour Market statistical bulletin and provide quality indicators not only for the headline level and rate, but also the respective quarterly and annual changes. Confidence intervals are available for the past 10 years in Table A11.

Figures 3 to 6 show the employment rate for people aged 16 to 64 and the unemployment rate for people aged 16 and over, along with the respective quarterly changes and the upper and lower boundaries of the 95% confidence interval for selected points in time.

UK

Figure 3: Employment rate people aged 16 to 64, with 95% confidence interval, Quarter 3 (July to Sept) 2006 to Quarter 1 (Jan to Mar) 2017
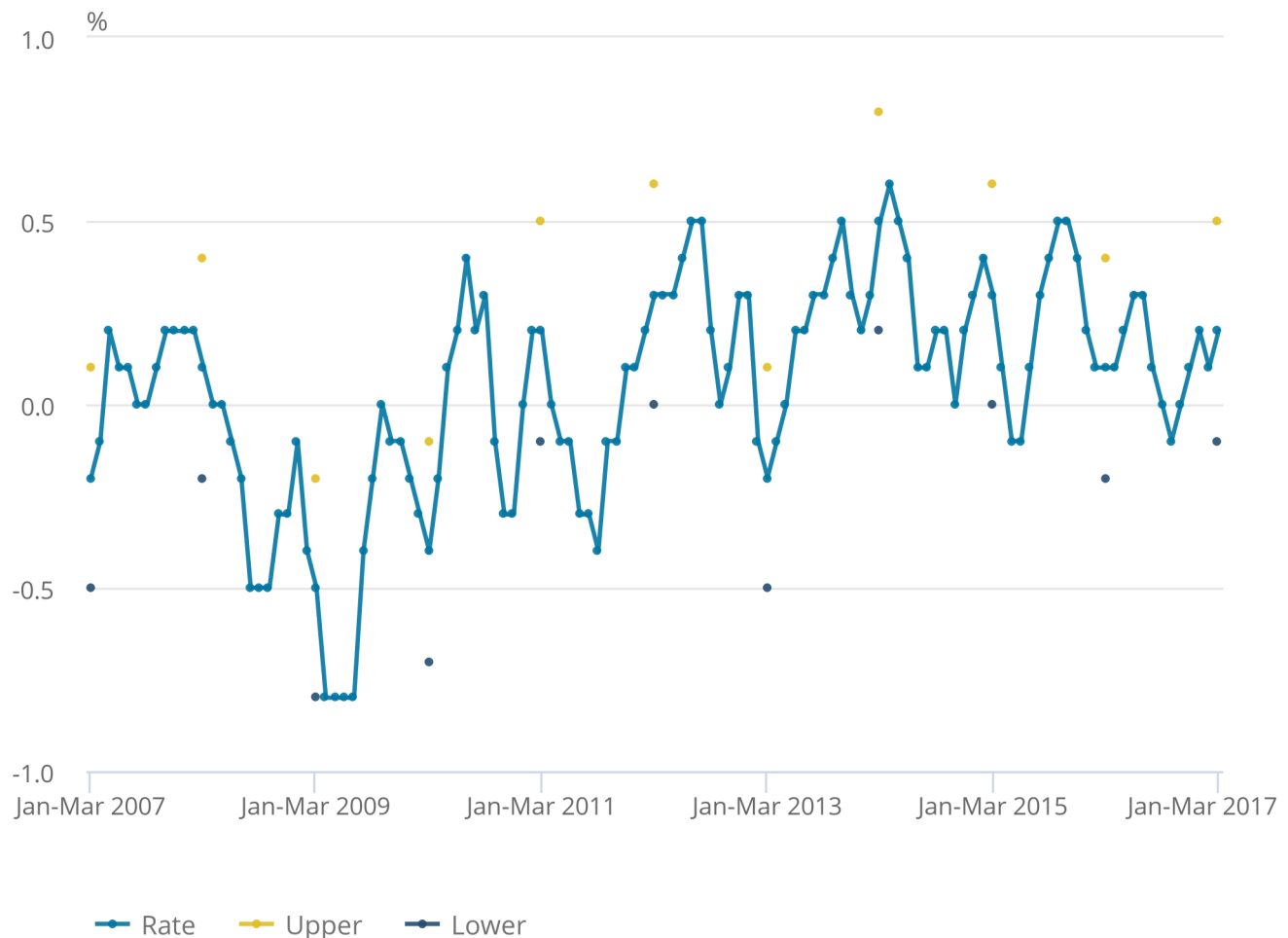
UK

Source: Office for National Statistics

Source: Office for National Statistics

For the employment rate (Figure 3), a 95% confidence interval of plus or minus 0.3 percentage points in 2006 is now plus or minus 0.4 percentage points in 2016. For the quarterly change (Figure 4), this has hardly moved from plus or minus 0.3 percentage points over the time period shown. The employment figures are based upon a very large sample, so even with falling response and the subsequent drop in the achieved sample size, this has had little effect on the sampling variability.

Figure 4: Employment rate, people aged 16 to 64, change on quarter with 95% confidence interval, Quarter 1 (Jan to Mar) 2007 to Quarter 1 (Jan to Mar) 2017

UK

Figure 4: Employment rate, people aged 16 to 64, change on quarter with 95% confidence interval, Quarter 1 (Jan to Mar) 2007 to Quarter 1 (Jan to Mar) 2017
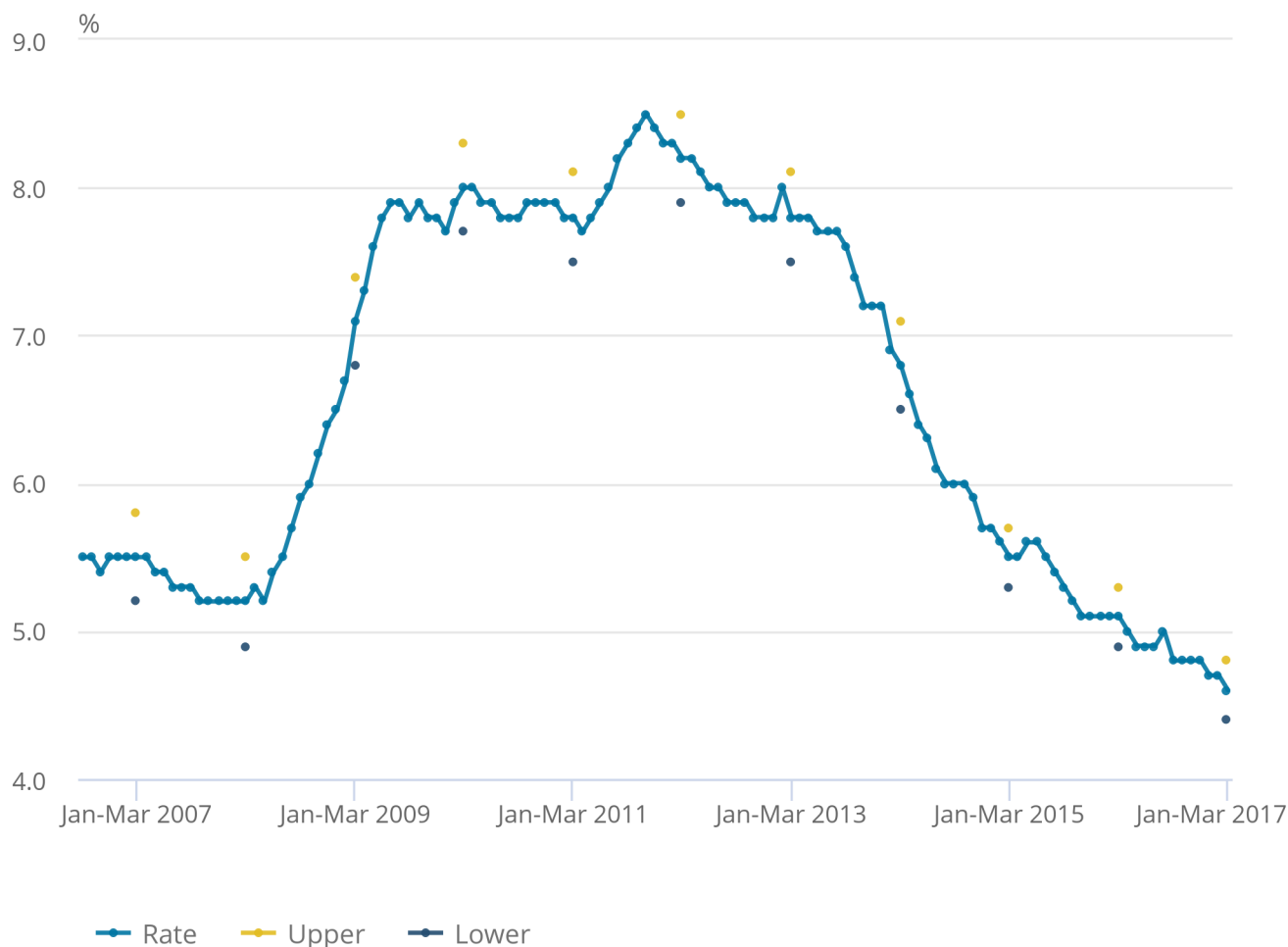
UK

Source: Office for National Statistics

The unemployment rate has relatively higher sampling variability (95% confidence intervals) of plus or minus 0.3 percentage points in 2006 and remains the same in 2016.

UK

Figure 5: Unemployment rate, people aged 16 to 64, with 95% confidence interval, Quarter 3 (July to Sept) 2006 to Quarter 1 (Jan to Mar) 2017
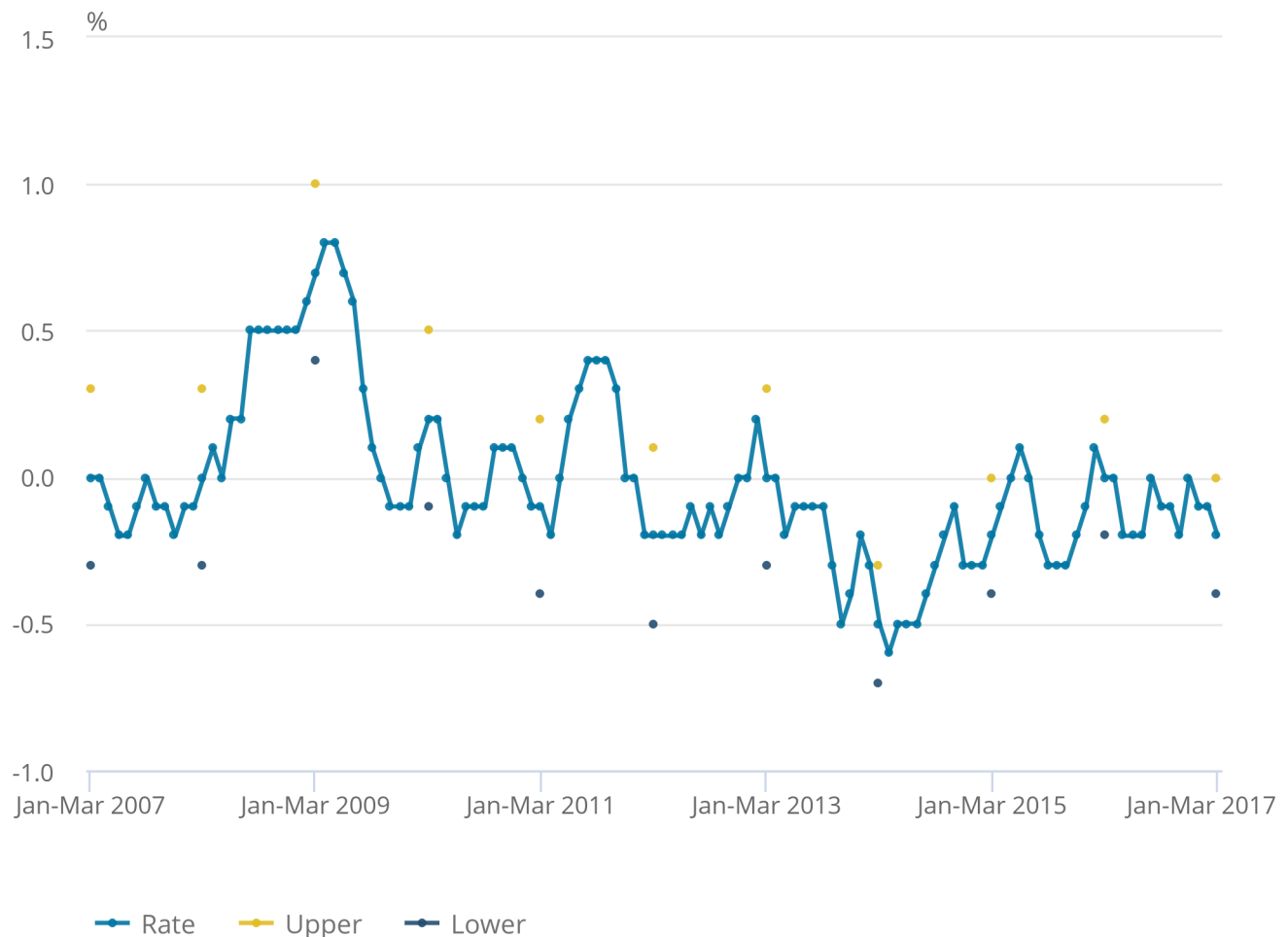
UK

Source: Office for National Statistics

Source: Office for National Statistics

Figure 6: Unemployment rate, people aged 16 to 64, change on quarter with 95% confidence interval, Quarter 1 (Jan to Mar) 2007 to Quarter 1 (Jan to Mar) 2017

UK

Figure 6: Unemployment rate, people aged 16 to 64, change on quarter with 95% confidence interval, Quarter 1 (Jan to Mar) 2007 to Quarter 1 (Jan to Mar) 2017

UK

Source: Office for National Statistics

Source: Office for National Statistics

There would need to be a significantly large drop in the sample size to increase the confidence interval to the point that published UK estimates are perceived to be unusable due to a lack of precision.

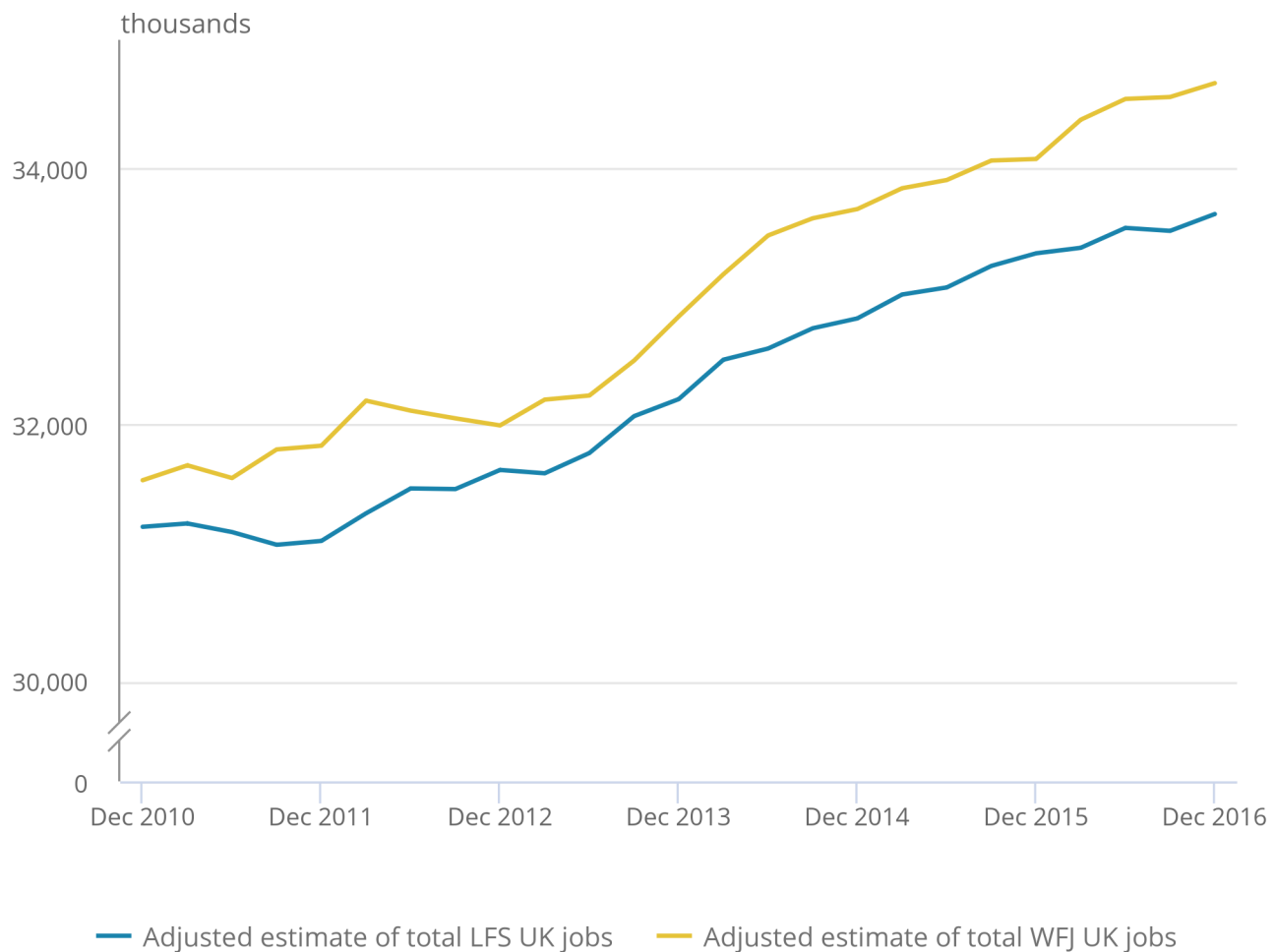## 4.3 Comparison with other data sources

Workforce jobs

Office for National Statistics (ONS) publishes a [quarterly comparison of Labour Force Survey (LFS) employment](#) with Workforce Jobs (WFJ), a measure of jobs produced from surveys of businesses. The comparison shows the time series of each respective series, as published, and after adjustments for known and measurable differences, determined by the [Employment and Jobs review in 2006](#). There are no significant differences in the employment trends between the two sources, although there is a "gap" between the estimates of the level, likely to be caused by non-quantifiable differences between the series.

Figure 7 shows the LFS and WFJ series adjusted for known differences. The final data point for December 2016 (WFJ) and November 2016 to January 2017 (LFS) shows the two series slightly diverging, with the WFJ series showing stronger growth in recent quarters. However, this is usual and is similar to movements seen when employment started to increase in 2011. When the trend in employment starts to change, the two series can diverge in the short term, partly due to WFJ being a point-in-time estimate and the LFS a 3-month average figure.

Figure 7: Labour Force Survey and Workforce Jobs estimates of jobs adjusted for measurable differences, seasonally adjusted, December 2010 to December 2016

UK

Figure 7: Labour Force Survey and Workforce Jobs estimates of jobs adjusted for measurable differences, seasonally adjusted, December 2010 to December 2016

UK

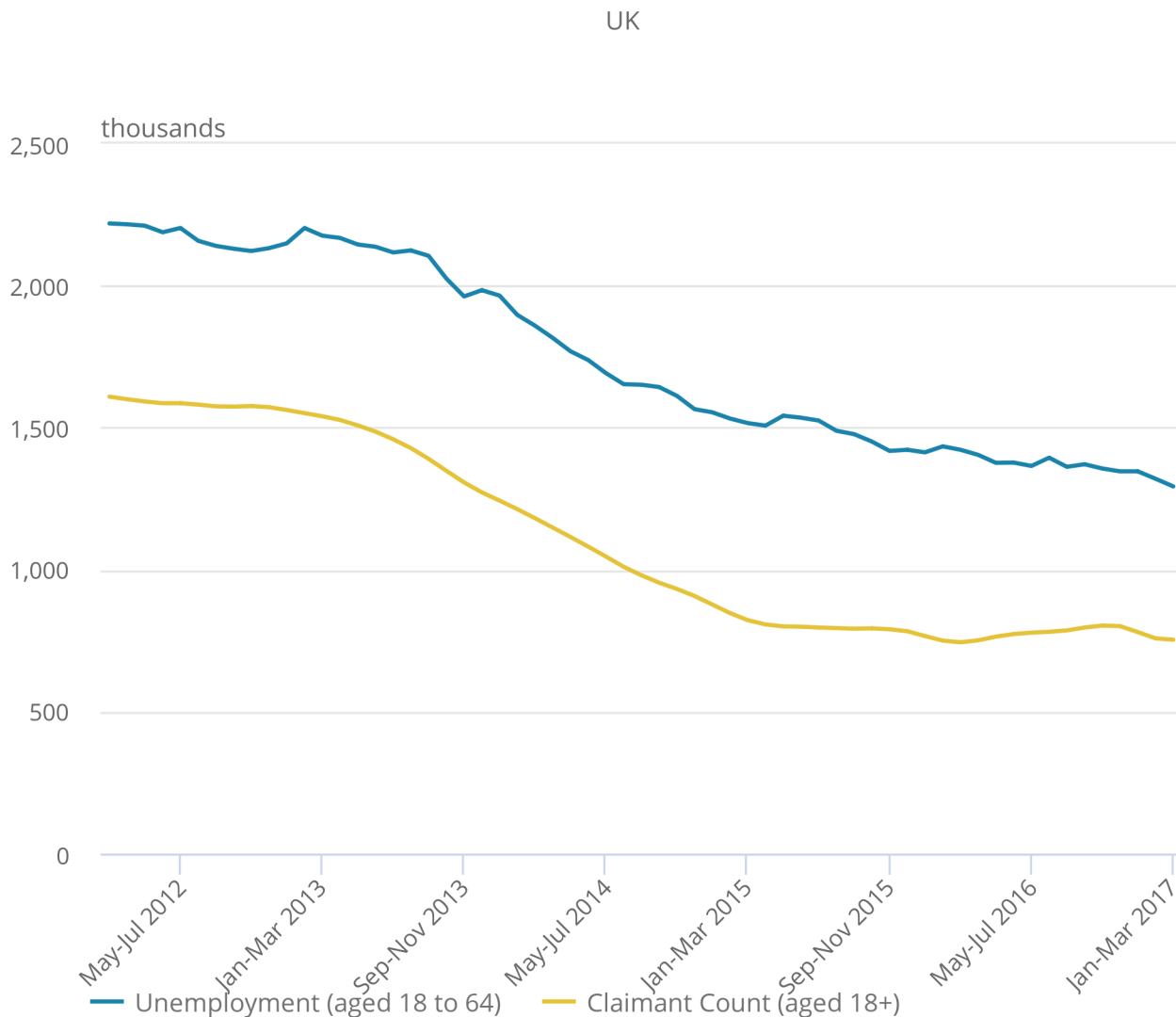Source: Office for National Statistics

Claimant count

The Claimant Count measures the number of people claiming unemployment-related benefits. The Claimant Count estimates are currently designated as Experimental Statistics because the Universal Credit estimates are still being developed by the Department for Work and Pensions (DWP). The roll out of the full service Universal Credit over the coming months means that the Claimant Count may be volatile from month to month, potentially affecting the seasonal adjustment of the data. However, the Claimant Count estimates do provide the best available estimates of the number of people claiming unemployment-related benefits in the UK. The Claimant Count includes people who claim unemployment-related benefits but who do not receive payment. For example, some claimants will have had their benefits stopped for a limited period of time by Jobcentre Plus. Some people claim Jobseeker's Allowance in order to receive National Insurance Credits.

In the UK labour market statistical bulletin we compare quarterly movements in unemployment, from the LFS, with quarterly movements in the Claimant Count in Table X05. Some claimants will not be classified as unemployed. For example, people in employment working fewer than 16 hours per week can be eligible to claim Jobseeker's Allowance depending on their income. The unemployment estimates used in this comparison exclude unemployed people aged 16 to 17 and aged 65 and over, as well as unemployed people aged 18 to 24 in full-time education. This provides a more meaningful comparison with the Claimant Count than total unemployment because people in these population groups are not usually eligible to claim unemployment-related benefits. Figure 7 shows a comparison of the levels for LFS unemployment and the Claimant Count.

Figure 8: Unemployment and Claimant Count comparison, Quarter 1 (Jan to Mar) 2012 to Quarter 1 (Jan to Mar) 2017

UK

Figure 8: Unemployment and Claimant Count comparison, Quarter 1 (Jan to Mar) 2012 to Quarter 1 (Jan to Mar) 2017

UK

thousands

Source: Office for National Statistics

Source: Office for National Statistics

Notes:

1. The unemployment estimates shown in this comparison exclude unemployed people in the 16 to 17 and 65 and over age groups as well as unemployed people aged from 18 to 24 in full-time education. This provides a more meaningful comparison with the Claimant Count than total unemployment because people in these population groups are not usually eligible to claim unemployment related benefits.

Census

To date, the main mechanism for accurately assessing potential bias in the LFS has been the Census Link Studies (CLS). After every census we examine the characteristics of survey non-respondents to assess non-response bias. The results from 2011 suggested slightly lower bias overall than in 2001, even though the LFS response rate had fallen from 80% to 62% in this time. This is as a result of the 2011 survey seemingly being more representative of the population overall than in 2001. The most recent CLS did not recommend the introduction of bias adjustments for the LFS.

4.4 Areas of vulnerability

Whilst the 2011 Census Link Studies (CLS) provided reassurance overall, it highlighted concerns within some socio-demographic groups and some survey variables – Black, Asian, and Minority Ethnic economic activity for example. This is the inherent nature of survey non-response bias; it is statistic specific, not survey based. So, for example, bias may be present within one survey variable and not another.

A currently high-profile potential vulnerability is related to country of birth and nationality of respondents. The Labour Force Survey (LFS) weighting does not account for any overrepresentation or underrepresentation in the sample of foreign nationals. Consequently, the estimates may show some effects of this from time to time, although the sample does not appear to be biased either way. Also, revisions to Office for National Statistics (ONS) population estimates, arising from updated net migration estimates, do not necessarily feed through directly to increased estimates of foreign nationals.

4.5 Opportunities to utilise administrative data

One way of assessing and improving the robustness of the Labour Force Survey (LFS) results is to use available administrative data on tax and benefits. Initial work is already underway in ONS's Social Survey Division, which expects to:

- complete linking of LFS non-responders with benefits and income from administrative data

- replicate linking methodologies, an approach already proven for the Census Link Studies

- publish an interim analysis of findings this summer

- continue to explore what is possible from other administrative and "Big Data" sources, turning this into an ongoing programme of work

5 . Plans for future updates

This article details the work currently carried out to monitor the robustness of Labour Force Survey (LFS) results as published in the monthly [UK labour market statistical bulletin](). Future updates will focus on the data source and discuss the short-term, medium-term and long-term actions to address declining response and to investigate non-response bias in the LFS. This will include commentary on the long-term strategic solution of survey and non-survey data source integration and statistical redesign.