

Monthly Early Estimates for working age client groups: Questions and Answers¹

Contents:	Page
Coverage	
Q 1: What makes the final National Statistics more accurate than the early estimates?	2
Q 2: Why are you using payment system scans and weighting them up to produce early estimates for these key working age client groups?	2
Q 3: Will you revise the early estimates when the National Statistics are published each quarter?	2
Q 4: Incapacity benefits scans are only available on a six weekly basis so how are you getting monthly estimates?	2
Q 5: ESA is a relatively new benefit replacing incapacity benefits. How have you combined the data for the two different benefits?	3
Q 6: Why are we not publishing JSA early estimates?	3
Q 7: When will you produce early estimates for the other client groups?	3
Q 8: When will you produce early estimates for the individual benefits?	3
Q 9: Why are the early estimates labelled as 'experimental'?	3
Timetable	
Q 10: When will subsequent figures be released?	4
Q 11: What changes were made to the methodology during Phase Two?	4
Q 12: Who is allowed early access to these estimate figures?	4
Technical	
Q 13: How do you adjust for the six weekly incapacity benefits scan dates?	5
Q 14: Why are we quoting an estimated range alongside the early estimates?	5
Q 15: Why was a 12 month rolling median used as the weighting factor opposed to any other method?	5
Q 16: Why are we using a rolling median but missing out the last four months?	5
Q 17: Why did you use uncarated instead of carated National Statistics? Note – carate is a blurring factor normally used on National Statistics to protect the identity of individuals.	6
Q 18: Why are the point estimates rounded to the nearest 5,000?	6
Q 19: How often will the error bound percentages be revised?	6

Coverage

Question 1: *What makes the final National Statistics more accurate than the early estimates?*

Answer 1: For the National Statistics there is a gap of around five and a half months between the publication and reference dates. See:
<http://research.dwp.gov.uk/asd/tabtool.asp>

The main reason for this lag is because we have to wait three months for a small but significant number of late claims to be processed. Waiting for these late claims makes the National Statistics figures more accurate than the early estimates where we simply predict the number of late claims based on historical levels.

Question 2: *Why are you using payment system scans and weighting them up to produce early estimates for these key working age client groups?*

Answer 2: The payment system scans are the basis of the published National Statistics.

However, timelier information is needed in the current economic climate and so we have developed monthly caseload estimates for the working age ESA and incapacity benefits and Income Support Lone Parent client groups.

Using the scans means that the early estimates are produced from the same data as the National Statistics but, instead of waiting for late claims to be processed, we estimate the number of late claims.

Question 3: *Will you revise the early estimates when the National Statistics are published each quarter?*

Answer 3: The quarterly National Statistics figures are the definitive figures for the quarter they cover. In that sense they are effectively revisions, at a quarterly level, for the early estimates for those months. The early estimates for non quarter months will not be revised.

Question 4: *Incapacity benefits scans are only available on a six weekly basis so how are you getting monthly estimates?*

Answer 4: The incapacity benefits scan nearest to the end of the month, regardless of which month it is actually in, is the starting point. This figure is then adjusted to reflect the daily change in the incapacity benefits figures and the number of working days by which the scan period and date of the end of the month differ. This is necessary because the introduction of ESA means that most new cases start on ESA rather than incapacity benefits and hence the incapacity benefits caseload is dropping by several thousand every week as people leave incapacity benefits.

For some months this correction for the scan dates is small, only around 7,000 cases, but for some months it is as large as 26,000 cases. But this will have less influence over time as ESA is rolled out.

Question 5: *ESA is a relatively new benefit replacing incapacity benefits. How have you combined the data for the two different benefits?*

Answer 5: We are estimating a combined ESA and incapacity benefits client group and we have weekly scans for ESA. Therefore we combine our estimated number of incapacity benefits cases at the end of the month with the number of cases on the ESA weekly scan closest to the end of the month.

Question 6: *Why are we not publishing JSA early estimates?*

Answer 6: JSA figures are currently produced on a monthly basis by ONS, with a time lag of approximately 5 weeks.

Question 7: *When will you produce early estimates for the other client groups?*

Answer 7: Phase Two of the early estimate work addressed the priorities identified amongst our customers for further work. Early estimates for other client groups were not identified as a high priority so at present early estimates for the other client groups are not being produced. Collectively these only account for approx 12 percent of the working age inactive group.

Question 8: *When will you produce early estimates for the individual benefits?*

Answer 8: We have no plans to do so at this stage.

Question 9: *Why are the early estimates labelled as 'experimental'?*

Answer 9: Early estimates have been labelled 'experimental' as we are still assessing the quality of the methodology. Phase One resulted in us publishing early estimates in mid June using a simple approach. Phase Two has improved the accuracy of early estimates and the first estimates using this improved methodology were published on 20 January 2010. We will continue to monitor the accuracy of the early estimates before making a decision on whether the 'experimental' label can be dropped.

Timetable

Question 10: *When will subsequent figures be released?*

Answer 10: Each month we will release, at the same time as the Labour Market Statistics, figures for the calendar month which ended approximately six weeks earlier.

Question 11: *What changes were made to the methodology during Phase Two?*

Answer 11: Phase Two started after the release of information on 17 June 2009 and concluded at the end of December 2009. The work in this phase was prioritised following discussion with stakeholders.

The main aim of Phase Two was to look at improving the accuracy of the early estimates. The first stage in Phase Two was to produce monthly datasets equivalent to the quarterly National Statistics figures published. These monthly datasets then enabled us to investigate various methodologies for producing early estimates. These methodologies were, computed using a rating factor based on the ratio between National Statistics and payment system scan caseloads, simple regression analysis, and regression analysis including various other variables that it was felt might help explain the differences between the National Statistics and payment system scan caseloads.

Following comparisons of estimates produced from each method there was little difference in the accuracy between the methods. In part, because there are fewer data points for the regression approach. As there was little difference in accuracy the method that is easier to implement each month was the chosen methodology, namely computing a monthly rating factor based on the ratio between the National Statistics and payment system scan caseloads.

Due to an increased level of confidence in the accuracy of the estimates, resulting from the improvements made in Phase Two and by comparisons with the final published National Statistics, we have changed the way we calculate and display the error bounds. They are now determined by using an average error rate method taking into account the percentage difference between early estimates and National Statistics over a 12 month period.

As well as the changes to the methodology described above, there are differences between the Phase One and Phase Two estimates in the rounding policy. The Phase Two point estimates have been rounded to 5,000 rather than 10,000 as the new methodology produces more accurate estimates and therefore we have been able to reduce the rounding figure to reflect this improved accuracy.

Question 12: *Who is allowed early access to these estimate figures?*

Answer 12: The early estimates are designated Official Statistics and we apply the same pre-release access rules as for our National Statistics. These arrangements are set down in the Pre Release Access Order. In summary, unless individuals are involved with quality assurance work the first access will be the pre release access group 24 hours before release of the figures. The members of the pre release group are restricted and are published on the web.

Technical

Question 13: *How do you adjust for the six weekly incapacity benefits scan dates?*

Answer 13: The *average change* in the number of cases per working day is calculated between two adjacent scans. The later of the two scans may then be adjusted, by this change figure, and used to estimate the number of cases at a month end. The adjustment takes the *average change* figure and multiplies it by the number of working days by which the scan differs from the month end date. This typically gives an adjustment of plus or minus between 100 and 26,000 cases depending on the month. The adjustment is then applied to the scan caseload figure.

Question 14: *Why are we quoting an estimated range alongside the early estimates?*

Answer 14: We cannot accurately produce a point estimate given the changes both for Lone Parent Obligations and with the introduction of ESA. The range is a useful reminder of the uncertainty that surrounds the estimates and that the real figure is likely to lie within a range of values.

Question 15: *Why was a 12 month rolling median used as the weighting factor opposed to any other method?*

Answer 15: Several potential methods of producing early estimates were considered. These included using a mean, median, rolling mean and rolling median. Each option has different strengths and weaknesses. The 12 month rolling median minimised the greatest error between the estimated quarterly figures and actual National Statistics.

Additionally, medians are less affected by extreme values than means and using a rolling average takes into account the most recent trends.

Question 16: *Why are we using a rolling median but missing out the last four months?*

Answer 16: We have used the information from four months behind due to timing issues over when figures are available and when estimates are needed. This method means that we can obtain rating factors for four months after the most recent published figures. For example, we can obtain a rating factor for November 2009 based on data up to July 2009.

Question 17: *Why did you use uncarated instead of carated National Statistics? Note – carate is a blurring factor normally used on National Statistics to protect the identity of individuals.*

Answer 17: The numbers are already well protected by the rounding process.

Question 18: *Why are the point estimates rounded to the nearest 5,000?*

Answer 18: Changing the rounding of the point estimates from 10,000 to 5,000 reflects the improvements in accuracy in the Phase Two estimates. These figures are estimates and rounding them reflects this and avoids spurious levels of accuracy.

Question 19: *How often will the error bound percentages be revised?*

Answer 19: The error bound percentages will be regularly monitored and will be revised every six months or earlier if the figures warrant it.

¹ Version two of question and answer document