

Measuring Adult Educational Attainment Using the Labour Force Survey – plans for improvements in response to statistical research

1. Introduction

- 1.1. In July 2009, BIS and the ONS jointly commissioned a statistical research project to understand the strengths and weaknesses of the existing method of measuring adult educational attainment using the Labour Force Survey (LFS) and to identify the most optimal measure available for use at the present time. This project is part of wider programme of work, the LFS Education and Training Review (ETR). The other projects under the ETR are: data matching work to support the statistical research; a “Wave 6” LFS follow-up survey to investigate the quality of responses; and a detailed review of the education and training section of the LFS questionnaire.
- 1.2. An [interim report](#) on the statistical research conducted by RM Data Solutions was published on 22 October 2009, setting out the issues that the first phase of research had identified and what further research should be undertaken to fully understand these issues. BIS and ONS have presented these findings to LFS user groups during November and December. The research is now complete and RM Data Solutions have successfully identified firm proposals for adjusting the various attainment estimates to compensate for the biases they have identified. The research team has produced a [final report](#) which includes recommendations for the next steps and further work that BIS and ONS should consider as a result of their findings.
- 1.3. The research has been subject to external quality assurance from Denis Allnutt, who has reported on his views on the quality and robustness of the research and its findings (see Annex A of this paper). Members of the Government Statistical Service (GSS) from BIS and ONS have worked with the researchers and the Quality Assurer throughout the project to provide peer review for the research as it has progressed. The recommendations of RM Data Solutions have been commented on by the Quality Assurer and reviewed and accepted by the GSS professionals in BIS and ONS. This paper sets out the practical steps necessary to respond to those recommendations, including some changes to reporting of PSA Targets and further research and development work in both the medium and long term.

2. Background

- 2.1. Data from the Labour Force Survey are used across Government and beyond for a wide range of activities. Within BIS they are used to examine the relationship between qualifications, employment and earnings as well as to calculate the rates of return to qualifications and to measure educational attainment levels. The attainment data are used to measure the Skills PSA targets, which require increases in levels of attainment within the working age population.

2.2. Qualifications attained through provision funded by the Learning and Skills Council are identified in the Individualised Learner Record (ILR), which is managed by the Further Education Data Service. In recent years, there has been a high growth rate in educational attainment observed from the ILR and modelling work suggested this was not being reflected in the growth rate observed in the LFS data. It is crucial that BIS are able to monitor attainment accurately if we are to ensure we are able to understand, develop and monitor Government policies and therefore BIS and ONS have been working together to understand the strengths and weaknesses of the current method and the underpinning sources.

2.3. A BIS/ONS [discussion paper](#) on the initial investigations was published in July 2009, setting out the apparent discrepancies between the different sources of data. This work suggested that further investigation was required and, as a result, a substantial statistical review of the robustness of the existing method of measuring attainment using the LFS was instigated. RM Data Solutions have now completed their research and BIS have also undertaken further work to develop their modelling of historical growth based on the administrative data.

3. Outcomes of the statistical research

3.1. The research by RM Data Solutions has identified that there are a number of potential sources of bias affecting the current method for estimating attainment. These biases are not independent of each other and RM Data Solutions have therefore proposed an adjustment solution that addresses as much of the overall bias as is possible, within the time and resource constraints of this project, and takes account of the interactions. Further details of the areas of potential bias investigated by the research are provided in sections 2 and 3 of the main research report.

3.2. The proposed solution broadly comprises of:

- Maximising the use of information collected during the 5 waves of LFS (to “boost” the sample in the reference quarter)
- Replacing LFS data by NPD information for 19-21 year olds
- Adjusting responses where ILR data from the previous 4 years definitively indicate a higher NQF level than reported
- Using personal response data to replace information given by proxy, and preferring the final response given by a subject where more than one response was given in person
- Imputing proxy responses, and responses where ‘Other’ is the highest stated qualification

3.3. Making these adjustments will increase the percentage of the population reported as holding qualifications. This is the case for each of the levels for which BIS reports on PSA targets, i.e. Level 2 and above (for both the working

age and economically active populations), Level 3 and above, and Level 4 and above. The adjustments also result in increased growth over recent years in the series for Level 2 and above and Level 3 and above. Further details of the impact are set out below within section 4 of this document.

- 3.4. Given that some of the data series used in the adjustment process do not extend back to 2001, further work is required to establish how estimates for economically active adults prior to 2006 can be adjusted. In March, BIS will publish an interim estimate of the growth in the number of economically active adults qualified to Level 2 or above; this will be based on the sum of:
 - the growth in the unadjusted series from 2001 to 2006; and
 - the growth in the adjusted series from 2006 to 2009.
- 3.5. BIS will also commission, jointly with ONS, further research on whether or not an adjusted measure of growth can be produced for 2001 to 2006. This research would seek to produce a “next best” adjustment for 2001 (which would also be applied to 2006), and the main research tasks will be:
 - verifying the applicability of the existing “sample boosting” approach;
 - producing a modified imputation model, making use of a reduced set of variables; and
 - analysing more closely the suitability of the alternative administrative data sources for young people (in the absence of NPD data being available for 2001).
- 3.6. BIS do not propose to take forward further matching work using FE data sets predating the ILR, because it is highly unlikely that those datasets can be configured to provide sufficient comparability. Furthermore, it is our judgement that any attempt to achieve comparability will involve disproportionate cost.
- 3.7. This research will be completed and reported on in summer 2010. If the research does not identify an improved measure for 2001 to 2006, BIS will adopt the interim measure (set out in 3.4 above) for reporting on growth since 2001 in the number of economically active adults qualified to Level 2 or above.
- 3.8. As it was not possible to match HESA data to the LFS within the life of the current project, the additional research now planned will also establish whether analysis of such data suggests further refinements are required to the adjustment method for the Level 4 and above series. Additional work will also be undertaken to assess the impact of matching ILR and HESA to the expanded LFS sample in the reference quarter rather than only to those respondents in the standard Q4 data set.
- 3.9. Because the scope of the research was limited to investigation of the measures used to report on the percentages of the population holding particular levels of qualifications, the adjustments that are being recommended are only designed to be used in producing England level estimates of attainment. The research

did not consider the possibility that the biases identified are relevant to other LFS measures or other uses of the education data, nor have adjustments been produced at this stage for statistics below England level. The research has identified some areas for further investigation in respect of the original scope; further details of these recommendations are set out under Section 5 of this paper

4. Plans for implementing the research outcomes

- 4.1. BIS and ONS have accepted the recommendations that RM Data Solutions have made and BIS will implement the proposed adjustments when reporting in March on progress against PSA targets as at Q4 2009. BIS will produce a series on the previous basis in March and this will be made available as part of the methodology documentation accompanying the Statistical First Release. More specifically, for the PSA target relating to the economically active population BIS will continue to publish the unadjusted series for 2001 to 2006, until a full assessment can be made of how adjustments might be made in earlier years.
- 4.2. In practical terms, the use of the adjustments that have been developed will result in a provisional set of NQF estimates which will be finalised once data for the following Q4 has been received. Thus the statistics for Q4 2009 will be finalised in the Statistical First Release published in March 2011, although in practice any changes are likely to be small. A broad outline of how the publication of provisional and final results will work can be found in section 6 of the final research report.
- 4.3. The statistics presented here are all provisional, as the estimates will need to be revised in light of reweighted base datasets recently produced by ONS. Furthermore, as the HESA data have not yet been matched to LFS data, the Level 4 and above statistics for all years will remain provisional in March 2010. Similarly, due to matched ILR data not being available at the present time for 2007, estimates for the Level 2 and above and Level 3 and above series for 2007 will also remain provisional in March 2010.
- 4.4. The current intention is that all of the provisional estimates produced in March 2010 will be finalised in March 2011. Incorporation of the additional matched data will be investigated as part of the follow-up research set out in section 3 above. If, against expectations, this results in large changes to any of the provisional estimates we will produce a revision in advance of March 2011.
- 4.5. The timetable for publication of results is therefore as follows:

Available within this report	Provisional adjusted estimates for Q4 2006 Provisional adjusted estimates for Q4 2007 Provisional adjusted estimates for Q4 2008
------------------------------	--

25 March 2010	Final adjusted estimates for Q4 2006 Final adjusted estimates for Q4 2008 Provisional adjusted estimates for Q4 2009
March 2011	Final adjusted estimates for Q4 2007 Final adjusted estimates for Q4 2009 Provisional adjusted estimates for Q4 2010
March 2012	Final adjusted estimates for Q4 2010 Provisional adjusted estimates for Q4 2011
March 2013	Final adjusted estimates for Q4 2011 Provisional adjusted estimates for Q4 2012

4.6. The impact of introducing the adjustments for the historical series back to 2006 is as follows:

	L2+		L3+		L4+	
	Published	Adjusted	Published	Adjusted	Published	Adjusted
2006	69.9%	71.7%	49.4%	52.3%	30.0%	32.3%
2007	70.7%	73.1%	50.6%	53.9%	30.9%	33.2%
2008	71.2%	73.9%	50.8%	54.5%	31.2%	33.7%
95% CI	+/-0.4 pp	+/-0.3pp	+0.5 pp	+/-0.4pp	+0.5 pp	+/-0.4pp
Growth 2006 to 2008	1.3pp	2.2pp	1.4pp	2.2pp	1.2pp	1.4pp
95% CI	+/- 0.6pp	+/-0.5pp	+0.7 pp	+/-0.5pp	+0.6 pp	+/-0.6pp

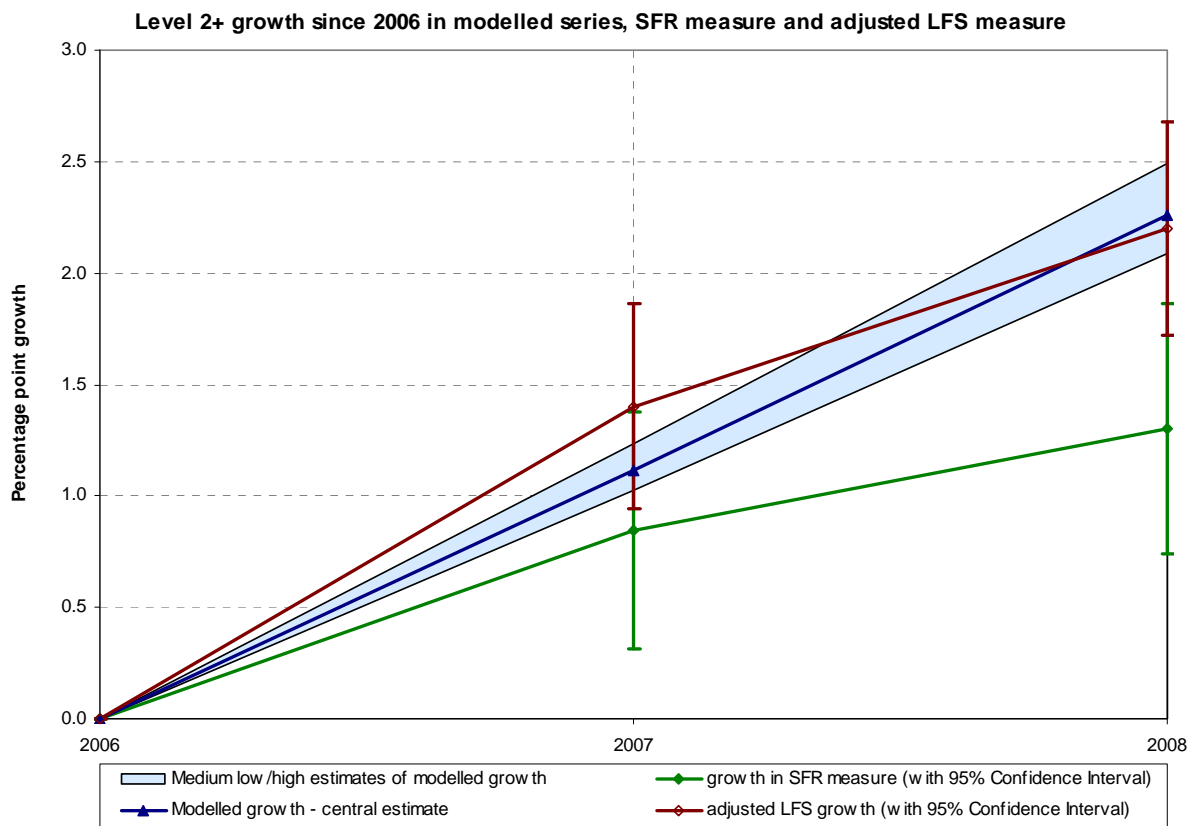
4.7. Applying the adjustment method to the economically active adult population increased growth between 2006 and 2008 from 539 thousand to 820 thousand more economically active adults qualified to Level 2 or above. As described in section 3 above, further work would be required before the adjustment method could be used prior to 2006, therefore no adjustment has been made to growth between 2001 and 2006 at the present time.

4.8. As described in section 2 above, BIS have also been undertaking further work on modelling historical data from other sources such as the ILR, as a means to understand the quality of the series with which the LFS statistics were being compared in the July 2009 discussion paper. As a result of this work, detailed in a separate [background paper](#), BIS now better understand the quality of the modelled historical data, and have been able to improve these estimates, in particular those for the Level 2 and above series.

4.9. The conclusion of this further work is that the modelling work is useful for providing an approximate view of the trend of attainment in the population, although it cannot provide a measure of the level without reference to the LFS. While the types of assumption being made do limit the accuracy of the growth estimates that can be produced, where the modelling suggests a change in the rate of growth we believe this can provide an insight to how one might expect the trend in the LFS should be changing.

4.10. The modelling work has also yielded a measure of uncertainty around the Level 2 and above series. The Level 3 and above and Level 4 and above series are considered to be more approximate, due to more assumptions being required in the modelling. Taking these limitations into account, the revised modelled data appear to suggest that the outcomes of the statistical research have largely resolved the discrepancies that were highlighted in the July 2009 paper.

4.11. The chart below compares the modelled growth for Level 2 and above, and the medium uncertainty range around this (see section 3 of BIS [report on modelling historical growth](#)), with the growth estimates that arise from both the SFR and adjusted LFS measures. 95% confidence intervals for the SFR and adjusted LFS growth estimates are illustrated by error bars for the 2007 and 2008 points.



4.12. The table below compares the growth in the modelled series with the growth in the adjusted LFS series for each of the three attainment levels.

	Modelled growth 2006 to 2008	Growth in adjusted LFS 2006 to 2008
Level 2 and above	2.3pp	2.2pp
Level 3 and above	1.9pp	2.2pp
Level 4 and above	1.3pp	1.4pp

5. Other actions being taken forward by BIS and ONS

5.1. In addition to the specific short term follow up research set out in section 3, RM Data Solutions have recommended that BIS and ONS conduct longer term follow-up work relating to this research. Where any of this new research suggests the adjustment approach requires refinement, an announcement would be made in advance of new data being published. The main areas in which it is recommended further research takes place are:

- Considering whether plans to improve the Education and Training section of the LFS questionnaire in 2011 necessitate any refinement of the adjustment approach
- In taking forward proposed general work on the following areas, ONS should seek to analyse impacts on qualifications data:
 - Further research, through the Census link study and related work, on possible bias arising from households that never respond to the LFS
 - Further research on uncertainty in population totals in terms of considering whether grossing methods should take account of household composition and recent migrants
- BIS and ONS should consider conducting a follow up survey to directly test accuracy of proxy responses and imputed proxy responses
- BIS should consider whether to expand the NPD database with a view to covering qualifications achieved by all adults

5.2. These recommendations will be fed into a review, by BIS and ONS, of the Education and Training Review programme in spring 2010. This process will establish a work plan for medium and longer term projects following on from the current range of ETR work described in section 1.

5.3. BIS anticipates an annual process of reviewing the adjustment approach, as new information comes to light. Refinements to the method would be made only if they can be shown to add statistical value to the data, in terms of reducing any bias still present. Where refinements are proposed, notice would be given in advance of the publication of the next Statistical First Release reporting on these measures.

5.4. The approximate timetable for review in the medium term is therefore:

Summer 2010	Any refinements on basis of matched HESA and LFS data
February 2012	Any refinements required to account for 2011 questionnaire changes
February 2013	Any refinements on basis of census linking work in respect of non-response, measurement error and migration

Annex A

REPORT FROM EXTERNAL QUALITY ASSURER

I confirm that, in my professional judgement, the revised methods of calculating qualification rates which are to be introduced following this project represent appropriate use of the available data – both from the Labour Force Survey and from administrative sources – to produce more accurate estimates of both the level of, and temporal changes in, qualification rates than were produced by the method used hitherto for the Statistical First Release.

I further confirm that the intended approach for handling the revisions to the data required by the new method are appropriate and have due regard to the availability of data and the costs of data matching. These necessitate the production of provisional estimates, based on the LFS data available in the spring after the reference quarter and incorporating an adjustment (based on previous years' differences between provisional and final estimates) for the expected impact of further LFS data and of the use of ILR data. Final estimates are then produced one year later, using the full LFS data and the matching of LFS and ILR data for individuals.

Throughout this project, both in the work to develop revised estimation methods and in the review of the BIS modelling of trends in L2+ qualification rates using data from a variety of sources, those undertaking the work have considered, and responded appropriately to, all suggestions for further work or revision to their analyses.

The further work planned on the estimation of the proportion qualified at L4+ (incorporating data from the Higher Education Statistics Agency into the analysis), and on attempting to develop – using the more restricted range of available data – an improved measure of the change in qualification rates between 2001 and 2006 is both necessary and appropriate to the development of improved measures.

The specification of any future work on weighting methods, or assessment of non response bias or the accuracy of proxy responses in the Labour Force Survey should address the impact of any methodological changes on the measurement of qualifications, and lead to an assessment of whether these estimates could be further improved.

Denis Allnutt