Do attrition and non-response adjustments really make a difference?

Debbie Cooper, Matt Greenaway, Andrea Lacey, Bethan Russ
Office for National Statistics

1.0 Introduction

Most UK household survey weights are the product of a design weight (calculated as the inverse of the selection probability) and a calibration weight. The calibration weight adjusts the design weights so that they match population totals for the variables included in the calibration model. This weighting process therefore takes sample design into account and also adjusts for imbalances in the selected sample. When non-responder characteristics are related to the calibration variables and to the survey outcomes, the calibration also accounts for non-response bias[1].

A number of UK household surveys include an additional step between calculation of design weights and calibration: sample-based non-response and attrition adjustments. Unlike calibration, sample-based adjustments are adjustments based on the characteristics of the sample as opposed to population characteristics. The aim of these adjustments is to decrease bias introduced by initial non-response and attrition. Initial non-response refers to non-response at Wave 1 of longitudinal surveys and non-response on cross-sectional surveys. Attrition refers to non-response from Wave 2 onwards of longitudinal surveys.

This paper examines the effectiveness of including the sample-based adjustments in the weighting process by:

(i) assessing the impact on estimates of removing sample-based attrition and non-response adjustments from household surveys that include them
(ii) investigating the impact on estimates of including sample-based attrition and non-response adjustments on household surveys that do not include them

This impact assessment was carried out for three UK household surveys: the Labour Force Survey (LFS), the Integrated Household Survey (IHS) and Personal Well-being (PWB). Each survey will be described in turn.

2.0 Labour Force Survey

The Labour Force Survey (LFS) is the largest UK social survey carried out on a regular basis[2]. It is a quarterly survey and uses a rotating panel design whereby selected households are retained in the sample for 5 consecutive quarters. Most households are interviewed face-to-face in their first Wave and then by telephone, if possible, from Wave 2 onwards.

2.1 Weighting methods

Weighting on the LFS involves calculation of a design weight (the inverse of the selection probability) followed by calibration. Calibration is carried out using the following calibration groups:

- Local Authority Districts (low level of geography, lower than NUTS3)

[1] Weeks, A., Fallows, A., Broad, P., Merad, S., Ashworth, K. 2013. Non-response Weights for the UK Labour Force Survey? Results from the Census Non-response Link Study. [online] Available at: http://www.ons.gov.uk/ons/guide-method/method-quality/specific/labour-market/articles-and-reports/non-response-weights-for-the-uk-labour-force-survey.pdf [Accessed 15th August2015]

[2] Office for National Statistics (2011), Labour Force Survey User Guide Volume 1: LFS Background and Methodology 2011. [Online] Available at: http://www.ons.gov.uk/ons/guide-method/method quality/specific/labour-market/labour-market-statistics/index.html. [Accessed 14th August 2015].

- Great Britain / Northern Ireland by sex by 'single year of age' for some age-groups
- Sex by age group by Government Office Region (GOR), equivalent to NUTS1

2.2 Impact on estimates of including attrition adjustments

Currently, the LFS does not include any initial non-response or attrition adjustments. However, one of the recommendations from the LFS National Statistics Quality Review[3] was to investigate the process of attrition on the LFS. Two main sets of analysis were carried out in order to do this. These included investigating the impact on key estimates when:

(i) A sample-based attrition adjustment was included in the weighting process
(ii) 'Tenure scaling' was applied

2.2.1 Including a sample-based attrition adjustment in the weighting process

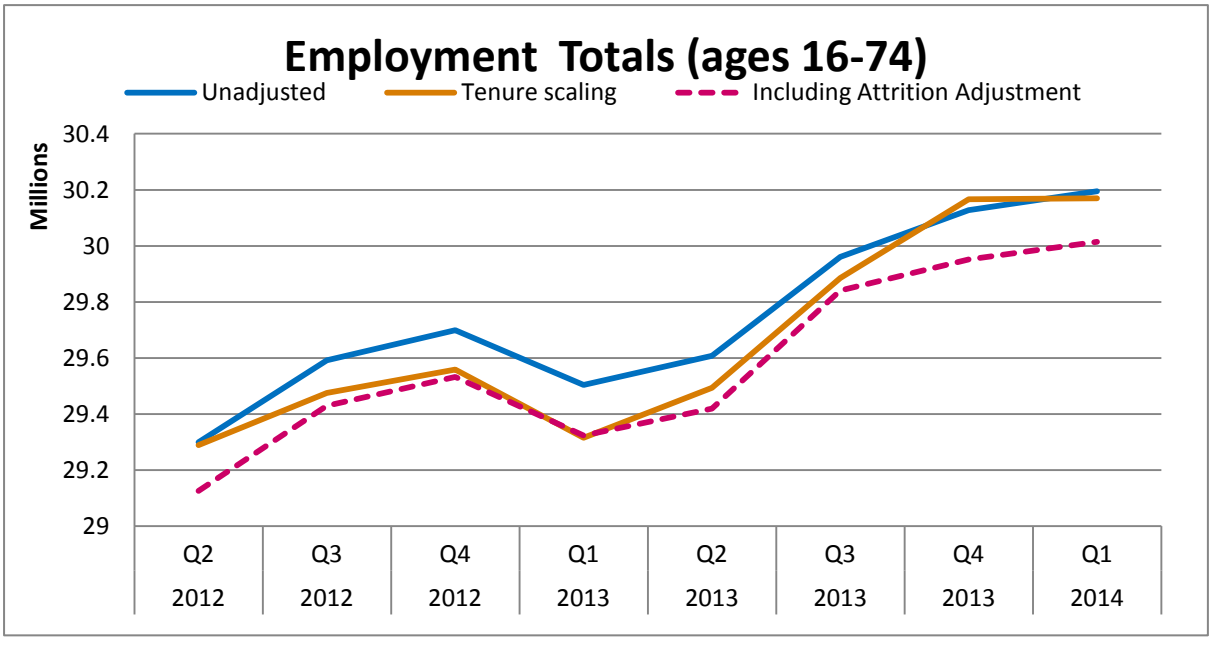
Following some exploratory analysis a logistic regression model was created in order to calculate the probability of an individual staying in the survey. With attrition adjustments, a wealth of information is available regarding individuals who drop out of the survey because these would have responded in an earlier wave. Consequently, a large number of possible independent variables could be included in the regression model. The variables chosen for inclusion in the logistic regression model were: age, tenure, GOR, disability, ethnicity, household type and whether an individual responded by proxy at the previous wave. The sample-based attrition adjustment was calculated by multiplying the design weights by the reciprocal of the probability of staying in the survey. The results are shown in Figure 1.

2.2.2 Applying 'Tenure scaling'

Calibration is currently carried out using totals by age, sex and region. Since the attrition modeling indicated that tenure appears to be linked to attrition, we investigated a simplified sample-based attrition adjustment, whereby the distribution of the design weights by tenure at Waves 2-5 was scaled to replicate the distribution of the design weights at wave 1.

The results of both sets of analysis described above for the employment estimate are provided in Figure 1.

Figure 1: A comparison of employment estimates (ages 16-74) with and without adjustments



[3] Office for National Statistics (2013). National Statistics Quality Review (NSQR) Series (2) Report Number 1: Review of Labour Force Survey. [Online] Available at: http://www.ons.gov.uk/ons/guide-method/method-quality/quality/quality-reviews/list-of-current-national-statistics-quality-reviews/nsqr-series--2--report-no--1/index.html [Accessed 14th August 2015].

Figure 1 indicates that including an attrition adjustment consistently lowers the employment estimate. The tenure scaling also seems to generally reduce the employment estimate although the picture is not as clear cut as it is for the attrition adjustment.

3.0 Integrated Household Survey

The IHS is a composite survey which is made up of a number of ONS social surveys. Each of the component surveys includes a set of IHS 'Core' questions. These include questions on: basic information (e.g. age, sex etc), components of identity, education and employment, perceived general health, and accommodation. Additional questions, such as smoking prevalence, which are sponsored by other government departments, are also asked on the IHS.

The aim of having a composite survey is to have a sample size large enough to produce higher precision estimates for lower levels of geography. The component surveys tend to have a rotating panel design. Some component surveys are cross-sectional and others are longitudinal.

3.1 Weighting methods

The IHS uses a multi-stage weighting procedure which accounts for probability of selection and adjusts for non-response. The weight produced for each responding or imputed case (person) will be a product of the following:

(i) The initial address-level design weight
(ii) A sample-based non-response adjustment for cross-sectional surveys and Wave 1 of longitudinal surveys
(iii) A sample-based attrition adjustment for longitudinal surveys
(iv) A scale factor for pooling surveys
(v) A calibration adjustment.

Stages (i) to (iii) above are calculated separately for the component IHS surveys. Following this, a scaling factor is calculated in order to pool the surveys and finally a calibration adjustment is carried out. Stages (ii), (iiii) and (v) are the stages of interest in this paper.

3.1.1 Non-response adjustment

In order to calculate the sample-based non-response adjustment a design-weighted binary indicator of response/non-response is regressed onto Output Area Classification (OAC) group, GOR and some system variables (to account for the complex sample design).

OAC is based on Census data and contains a number of groups (OAC groups) that describe the sociodemographic makeup of small geographic areas. OAC is believed to be related to a number of IHS core variables and is therefore used to help reduce the bias of the estimator. The reciprocal of the response propensity from the non-response model is then used as the non-response adjustment.

3.1.2 Attrition Adjustment

In order to reduce bias resulting from differential dropout rates after the first wave of the longitudinal surveys an attrition adjustment is used. This is also based on a logistic regression model.

With initial non-response not much auxiliary information about non-respondents is available. On the other hand, with attrition adjustments, much more detailed information about non-respondents is available. Consequently, more variables are available to include in the logistic regression model for

the attrition adjustments. A number of different variables are included in the regression model, depending on the survey, however, the main regressors included are:

- GOR
- OAC group
- household reference person age, sex and ethnicity
- length of time at address
- age of youngest dependent child in household aged under 19

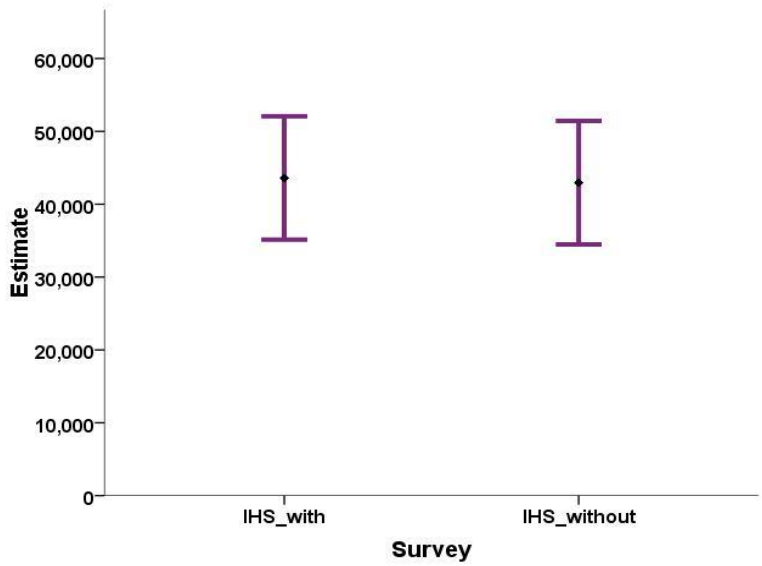
3.1.3 Calibration Adjustment

The design weight is multiplied by the non-response adjustment and, where relevant, the attrition adjustment. The product of this is then calibrated to known population totals by age, sex and region. Integrative calibration is used on the IHS so that all individuals within a household are given the same weight.

3.2 Impact on estimates of removing non-response and attrition adjustments

Estimates with and without sample-based non-response and attrition adjustments were compared for the 2013 IHS dataset.

Figure 2: Total estimates for Non-Smokers who have never worked or are long term unemployed in Essex.



Key: IHS_with: estimate calculated from the dataset containing attrition and non-response adjusted weights; IHS_without: estimate calculated from the dataset containing weights which have not been adjusted for attrition and non-response.

Figure 2 indicates that removing the attrition and non-response adjustments from the IHS weighting process has a negligible impact on the estimate. Moreover, as the confidence intervals demonstrate a considerable amount of overlap, sample variability alone may be responsible for a similar change in the estimates year-on-year.

4.0 Personal Well-being

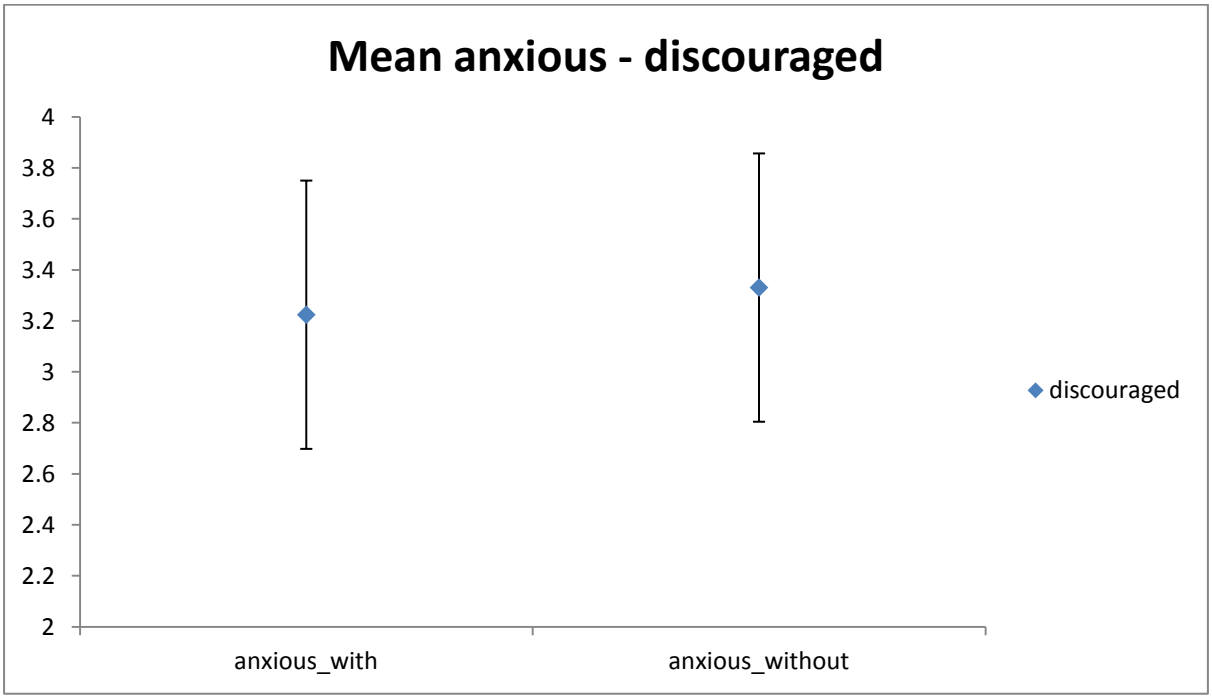
Personal well-being (PWB) is measured using 4 questions that are asked on the Annual Population Survey (which uses a rotating panel design). The weighting methods for PWB are very similar to

those of the IHS described above, although the wellbeing questions are only asked to over-16 respondents, they are not asked by proxy, and 'integrative' weighting is not used.

4.1 Impact on estimates of removing non-response and attrition adjustments

A similar impact analysis to that carried out on the IHS was performed for the key PWB estimates. Below is a graph showing one of the largest differences in the key estimates.

Figure 3: Mean estimate for anxious for 'discouraged' category



Key: anxious_with: estimate calculated from the dataset containing attrition and non-response adjusted weights; anxious_without: estimate calculated from the dataset containing weights which have not been adjusted for attrition and non-response.

Figure 3 indicates that removing the attrition and non-response adjustments has a negligible impact on the estimate. Moreover, as the confidence intervals demonstrate a considerable amount of overlap, sample variability alone may be responsible for a similar change in the estimates year-on-year. The largest differences between estimates with and without attrition adjustments for PWB all had quite large confidence intervals because of relatively small sample sizes in these particular categories.

Points for discussion:

- Is calibration sufficient in weighting or should sample-based attrition and non-response adjustments be included where possible? Rather than including non-response and attrition adjustments might it be better to focus on improving calibration?
- Most weights are produced within an automated production system – is it worth calculating attrition and non-response adjustments in such a case? There is higher chance of errors occurring. Moreover, it is generally too complex and costly to put regression models into an automated system so adjustment factors are generally used. However, these may not be reviewed regularly which results in them being outdated.