Measuring and reporting reliability of Labour Force Survey and Annual Population Survey estimates

Identification of existing threshold rules in Labour Force and Annual Population Surveys, with recommendations to improve consistency and the communication of uncertainty.

Table of contents

1 . Introduction

A review of our existing Office for National Statistics (ONS) publishing practices helped identity different quality threshold criteria for publishing Labour Force Survey (LFS) and Annual Population Survey (APS) estimates, but offered no clear reason as to why they are not consistent. The purpose of this paper is to:

- identify the existing quality threshold rules that are in place for these datasets (for example publications, ad-hoc requests, Parliamentary Questions (PQs), Freedom of Information (FOI) requests) across ONS directorates

- determine whether quality thresholds are layered on top of disclosure thresholds and review the alternative uncertainty measures that are published to communicate quality

- share these findings and make recommendations on how we can make our approach consistent, and more clearly communicate the level of uncertainty in our estimates

2 . Office for National Statistics thresholds

As the Labour Force Survey (LFS) and Annual Population Survey (APS) are sample surveys, these estimates are subject to an associated sampling error that decreases as the sample size increases. It is the nature of sampling variability that the smaller the group whose size is being estimated, the (proportionately) less precise that estimate is. In order to avoid misuse of estimates which are not sufficiently robust, reliability thresholds are applied which can be based on different methods, for example, the actual number of respondents, the grossed number of respondents or the coefficient of variation of an estimate. In all cases, a minimum number of responses is required to protect against risk of disclosure because of the small sample underpinning the estimates. The purpose of this guidance note is to determine the best approach to publish.

Currently, the Labour Market and Households Division (LMHD), Social Survey Division (SSD) and Public Policy Analysis Division (PPAD) routinely produce estimates from two LFS datasets – the quarterly dataset and the annual dataset – and use a threshold approach to determine whether estimates are suitable for publication. The main variations on these approaches are:

- For the LMHD, for UK and regional tables, a general release rule is employed where any output with a weighted population of 4,500 persons or more should be released.

- For outputs covering more detailed geographic areas, a threshold release rule is employed where any output with a person count of three or more should be published.

In addition to headline tables, the following rules are applied:

- For Parliamentary Questions, we calculate the coefficient of variation (CV) and use descriptions as our key to statistical robustness, suppressing estimates where the value has a CV>20%.

For SSD a general disclosure threshold release rule is employed where any output with a person count of three or more should be published.

Within the PPAD, rules also vary between branches:

Team	Suppress below (all unweighted)	Publish with shading	Publish no shading
Policy Evidence and Analysis Team	3	3-29 (uw)	30 (uw)
Cities Analysis	3	<6,000 (w)	6,000 (w)
Regional Economics Team	10	-	10 (uw)
Population Statistics	30	-	-
Migration Statistics	3	-	-

Source: Office for National Statistics

All branches have a threshold for suppression below which estimates will not be published. In addition, two branches use shading measures to denote quality precision if the output requires lower count measures to be published. Where requested by customers, other measures of quality such as coefficients of variation and/or confidence intervals have also been produced.

3 . Methodological approach

Working with the Methodology Division, the method proposed provides a minimum value estimate – for LFS and APS levels – that enables us to give an indication of the precision of the estimates.

The method used is based on an approximation of the sampling error of small groups; the minimum value n_min is given by:

$N\_min = \text{Design effect}/(CV)^2$

For the LFS, we set the design effect to 1 and the highest allowed CV to 20% which yield n_min=25

In the APS, the small groups occur in estimates by local authority, and the design effect is approximately similar to that in the LFS. So n_min=25

In the longitudinal LFS, the design effect is slightly higher because there is usually more variation in the weights by a factor of about 1.1. So n_min=27

As the LFS and APS are sample surveys, response rates will change, although the unweighted counts recommended will not change with rising or falling response rates. However, as response rates fall, for example, the weights will increase. Therefore, flagging the same unweighted count means that higher weighted counts get flagged.

Measures of uncertainty for LFS and APS rate and change estimates will be determined by the outcome of the methodological approach agreed for level estimates.

The LMHD will conduct an annual review with the Methodology Division to validate the continued use of the n_min unweighted values and/or agree updated values.

4 . Recommendations

Because of the number of estimates published, it is not possible to create a coefficient of variation (CV) or provide a measure of sampling variability for each estimate. Instead it is recommended to continue to provide measures of sampling variability for headline indicators in labour market reports and use the proposed methodological approach for the remainder.

In summary, adopting these methods leads to following specific recommendations:

- Any LFS or APS output with a person count of less than three should not be published and will be suppressed under disclosure threshold rules. Secondary disclosure rules will also continue to apply.

- On the quarterly datasets, any output with a person count of three or more should be published and a shading measure to denote quality precision (on counts of more than three and less than or equal to 25).

- On the annual datasets (published via LMHD and SSD), any output with a person count of three or more should be published and a shading measure to denote quality precision (on counts of more than three and less than or equal to 25).

- On the longitudinal datasets, any output with a person count of three or more should be published and a shading measure to denote quality precision (on counts of more than three and less than or equal to 27).

- Measures of uncertainty for LFS and APS rate and change estimates will be determined by the outcome of the methodological approach agreed for level estimates.

The above recommendations will apply to regular and ad-hoc LFS and APS outputs, Parliamentary Questions and Freedom of Information requests. Examples of shading and guidance text are provided in Annex A.

5 . Next steps

- This guidance note will be published on the ONS Labour Market methodology web page and links directing users to this note will be included in relevant outputs

- We will begin introducing the shading measure and accompanying footer guidance to LFS and APS outputs

6 . Annex A

Quality measures – colour coding and descriptions

The proposal is to use colour coding to highlight relative levels of quality (working on the basis of a 25% shading option). This recommended shading option, with reference to survey sample sizes, as the key to statistical robustness is presented in this annex.

To accompany the launch of this improved uncertainty communication, we plan to insert an additional footnote to our Labour Force Survey and Annual Population Survey publications to note the following.

Shaded estimates

Shaded estimates are based on a small sample size. This may result in less precise estimates, which should be used with caution.

Unshaded estimates

Unshaded estimates are based on a larger sample size. This is likely to result in estimates of higher precision, although they will still be subject to some sampling variability.