

## **New publication rules for the APS and the 2003/04 Annual LFS data**

Until the present release of the 2003/04 annual data, all published LFS data have had thresholds applied, where estimates below these threshold have been suppressed because they were considered 'statistically unreliable'. The thresholds were calculated so that they were approximately equivalent to suppressing if the standard error of an estimate was greater than 20 per cent of the estimate itself. For annual LFS data, different areas had different thresholds as some areas had larger samples than others.

However, ONS has recently decided that estimates from surveys should not be suppressed on the grounds of reliability. Some estimates will continue to be suppressed because they are considered to be disclosive, that is individuals could be identified from them. Instead of suppressing because estimates are not reliable, the estimates will be presented with an indication of their reliability. For the 2003/04 annual LFS data, estimates and approximate 95 per cent confidence intervals will, in most cases, be published.

The exceptions to this rule are:

- i) where an estimate is based on fewer than three observations: this estimate will be suppressed as disclosive,
- ii) where an estimate is based on three to nine observations: in this case the estimate will be published, but the confidence interval will not be published, as the confidence interval is considered to be not reliable, and could give users a misleading impression of the reliability of the estimate (the estimate itself will not be reliable, and a 'health warning' will be published with it), and
- iii) where the estimate of a rate is 100 per cent: in this case it is not possible to have a confidence interval around the estimate (often this rate will be based on a small number of observations, and one of i) or ii) above will apply).

What does the 95% confidence interval mean? If, for example, we have an LFS estimate and confidence interval of 63% +/- 0.27%, this means that 19 times out of 20 we would expect the true rate to lie between 62.73% and 63.27%. Only in exceptional circumstances (1 in 20 times) would we expect the true rate to be outside the confidence interval around the LFS estimate. Thus the narrower the confidence interval, the more confident we are that the true value is close to the estimate, and hence the narrower the confidence interval, the more reliable is the estimate.

The calculation of confidence intervals is based on formulae given in Volumes 1 and 6 of the LFS User Guides. (Volume 1 is 'LFS Background and Methodology' and Volume 6 is 'Local Area Data'). All LFS User Guides can be found at

<http://www.statistics.gov.uk/statbase/product.asp?vlnk=1537.>)

The formulae used are:

- i) for levels, the approximate standard error of an estimate is given by:

$$se = 1000 \times \text{square root}(M_T \times G/1000),$$

where  $M_T$  is the number of thousands of the estimate, and  $G$  is the average grossing factor of the area.

The confidence interval is then 1.96 x standard error.

- ii) for rates, the approximate standard error of  $p$ , where  $p$  is a percentage, rather than a proportion of 1, is given by:

$$se = 100 \times \text{square root} ((p/100) \times (1-(p/100))/(n-1)),$$

where  $n$  is the number of observations from which the denominator of the rate is derived.

The confidence interval is then 1.96 x standard error.

However, these formulae are for standard errors for survey estimates from simple random samples. This is not the case for the LFS, because data are presented for individuals but the household is the sampling unit. Therefore, a design factor should be applied to the standard errors (and hence to confidence intervals) to take account of clustering within households. Different variables will have different patterns of clustering within households. For example the clustering of employment status is likely to be very different from the clustering of ethnic group. There will be much clustering of ethnic group within households, and so a high design factor should be applied for estimates which include ethnicity.

In the confidence intervals published on Nomis, a design factor has been included for estimates which include ethnicity, but not for any other variables. Design factors have not been calculated for these latest data, but a table of design factors for some variables is included in Volume 1 of the LFS User Guide. These have been calculated from quarterly LFS data and are reproduced in Table 1 below. Although they have not been calculated from these latest data, they should give users some indication of the magnitude by which the simple random sample standard error underestimates the true standard error, and users should take this into account when assessing the reliability of the LFS estimates published here.

Table 1. Approximate Design Factors for a selection of LFS variables

Characteristic	Design Factor
<b>All people 16+ by Economic Activity:</b>	
Economically active	1.17
All in employment	1.17
Employees	1.16
Self-employed	1.07
ILO unemployment	1.04
Economically inactive	1.17
<b>Men 16+ by Economic Activity :</b>	
Economically active	1.02
All in employment	1.02
Employees	1.02
Self-employed	1.02
ILO employment	1.01
Economically inactive	1.02
<b>Women 16+ by Economic Activity:</b>	
Economically active	1.02
All in employment	1.02
Employees	1.02
Self-employed	1.00
ILO unemployment	1.02
Economically inactive	1.02
<b>All people 16+ in employment by type of employment</b>	
Full-time employees	1.02
Part-time employees	0.96
Full-time self-employed	1.03
Part-time self-employed	1.02
Government employment and training programmes	1.00
Unpaid family worker	1.04
<b>Men 16+ in employment by type of employment</b>	
Full-time employees	1.02
Part-time employees	1.01
Full-time self-employed	1.01
Part-time self-employed	1.01
Government employment and training programmes	1.00
Unpaid family worker	1.00
<b>Women 16+ in employment by type of employment</b>	
Full-time employees	1.01
Part-time employees	1.01
Full-time self-employed	1.01
Part-time self-employed	1.00
Government employment and training programmes	1.00
Unpaid family worker	1.01

<b>All aged 16+ in employment by:</b>	
Agriculture & Fisheries	1.16
Energy & Water	1.02
Manufacturing	1.05
Construction	1.03
Distribution	1.06
Transport and Communication	1.04
Banking, Finance, Insurance	1.06
Public Admin, Health, Education	1.04
Other Services	1.05
<b>Men aged 16+ in employment by:</b>	
Agriculture & Fisheries	1.05
Energy & Water	1.00
Manufacturing	1.01
Construction	1.02
Distribution	1.02
Transport & Communication	1.01
Banking, Finance, Insurance	1.01
Public Admin, Health, Education	1.00
Other Services	1.01
<b>Women aged 16+ in employment by:</b>	
Agriculture	1.01
Energy & Water	0.99
Manufacturing	1.01
Construction	1.00
Distribution	1.01
Transport & Communication	1.01
Banking, Finance, Insurance	1.01
Public Admin, Health, Education	1.01
Other Services	1.01
<b>Economically active people 16+ who were ILO unemployed:</b>	
All persons	1.04
Men 16+	1.01
Women 16+	1.02
<b>Economically active people 16-17 who were ILO unemployed:</b>	
All persons 16-17	1.02
Men 16-17	1.02
Women 16-17	1.00
<b>Economically active people 18-24 who were ILO unemployed:</b>	
All persons 18-24	1.03
Men 18-24	1.03
Women 18-24	1.02
<b>Economically active people 25-34 who were ILO unemployed:</b>	
All persons 25-34	1.02
Men 25-34	1.00
Women 25-34	1.00

<b>Economically active people 35-49 who were ILO unemployed:</b>	
All persons 35-49	1.01
Men 35-49	1.00
Women 35-49	1.00
<b>Economically active people 50-59/64 who were ILO unemployed:</b>	
All persons 50-59/64	1.02
Men 50-64	1.00
Women 50-59	1.00
<b>Economically active people 60+/65+ who were ILO unemployed:</b>	
All persons 60+/65+	1.00
Men 65+	1.00
Women 60+	1.00
<b>People aged 16+ by ethnic origin: White</b>	
All	1.38
Men	1.13
Women	1.10
<b>People aged 16+ by ethnic origin: Black Carribean</b>	
All	1.25
Men	1.09
Women	1.09
<b>People aged 16+ by sex and ethnic origin: Indian</b>	
All	1.53
Men	1.15
Women	1.20
<b>People aged 16+ by sex and ethnic origin: Pakistani</b>	
All	1.65
Men	1.27
Women	1.25