CONFIDENCE LIMITS SURROUNDING SCOTLAND POVERTY ESTIMATES

The poverty estimates presented in the Scottish Households Below Average Income publication series are based on a sample survey and therefore there is some degree of statistical error, or uncertainty, around the estimates produced. In other words, when it is reported that 11 per cent of individuals are living in absolute poverty before housing costs, this should be understood not as an exact figure but as a best estimate.

Confidence limits provide a guide to how robust the estimates are. Tables 1 and 2 provide the 90% confidence limits around the key poverty estimates, and graphs 1 to 4 illustrate the data.

Table 1 shows that the best estimate for the number of individuals in absolute poverty after housing costs in 2003/04 was 14%, with a lower confidence limit of 12% and an upper confidence limit of 15%. This means that we can be 90% confident that the percentage of individuals in absolute poverty lies between 12 and 15 percent. Similarly, the lower confidence limit for the number of people in absolute poverty was 960,000 and the upper confidence limit was 1,080,000. So we can be 90% confident that the true number lies between those two figures.

As a rule of thumb, if the confidence limits of an estimate over two years over lap, it should be assumed that the difference between the two years is not significant. In other words, there may not have been a change in the percentage or number of people in poverty between those two years. The width of the confidence limits surrounding the past 3 years poverty estimates, when compared to the magnitude of change between years, suggests that much caution should be exercised when making year on year comparisons.

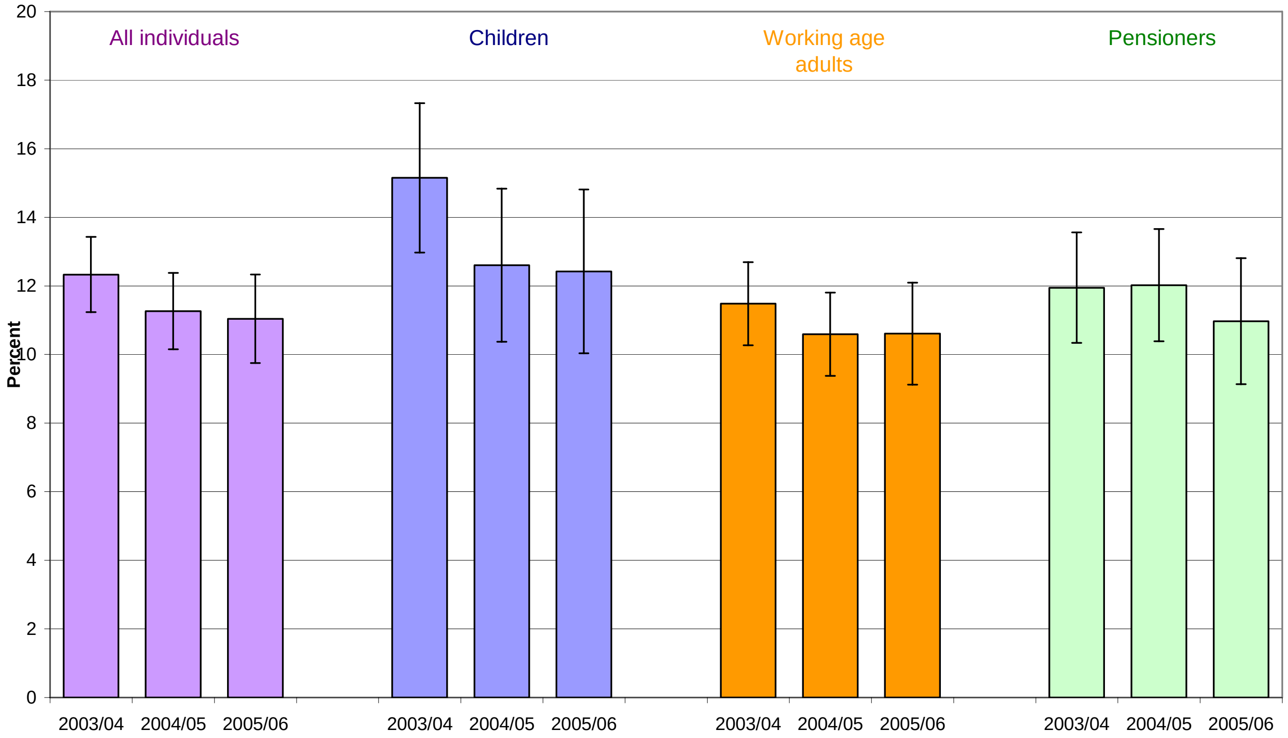
Table 1: Percentage and number of individuals in **absolute** poverty with 90% confidence limits

		Percentage			Frequency (thousands)		
		Lower confidence limit	Estimate	Upper confidence limit	Lower confidence limit	Estimate	Upper confidence limit
Before Housing Costs							
All individuals	2003/04	17	19	20	860	920	980
	2004/05	16	17	18	790	860	920
	2005/06	16	18	19	800	880	950
Children	2003/04	21	24	26	220	250	270
	2004/05	18	21	23	190	210	240
	2005/06	18	21	24	190	210	240
Working Age Adults	2003/04	15	16	17	440	480	520
	2004/05	14	15	16	420	460	500
	2005/06	14	15	17	420	470	520
Pensioners	2003/04	20	22	24	180	190	210
	2004/05	18	20	22	170	190	200
	2005/06	18	20	23	170	190	210
After Housing Costs							
All individuals	2003/04	19	21	22	960	1020	1080
	2004/05	18	19	20	900	960	1020
	2005/06	18	20	21	920	990	1060
Children	2003/04	24	27	29	250	280	300
	2004/05	22	25	27	230	250	270
	2005/06	22	24	27	220	250	270
Working Age Adults	2003/04	17	18	20	520	560	600
	2004/05	17	18	20	510	560	600
	2005/06	18	19	21	540	590	640
Pensioners	2003/04	19	21	22	170	180	200
	2004/05	15	16	18	140	150	170
	2005/06	15	16	18	140	150	170

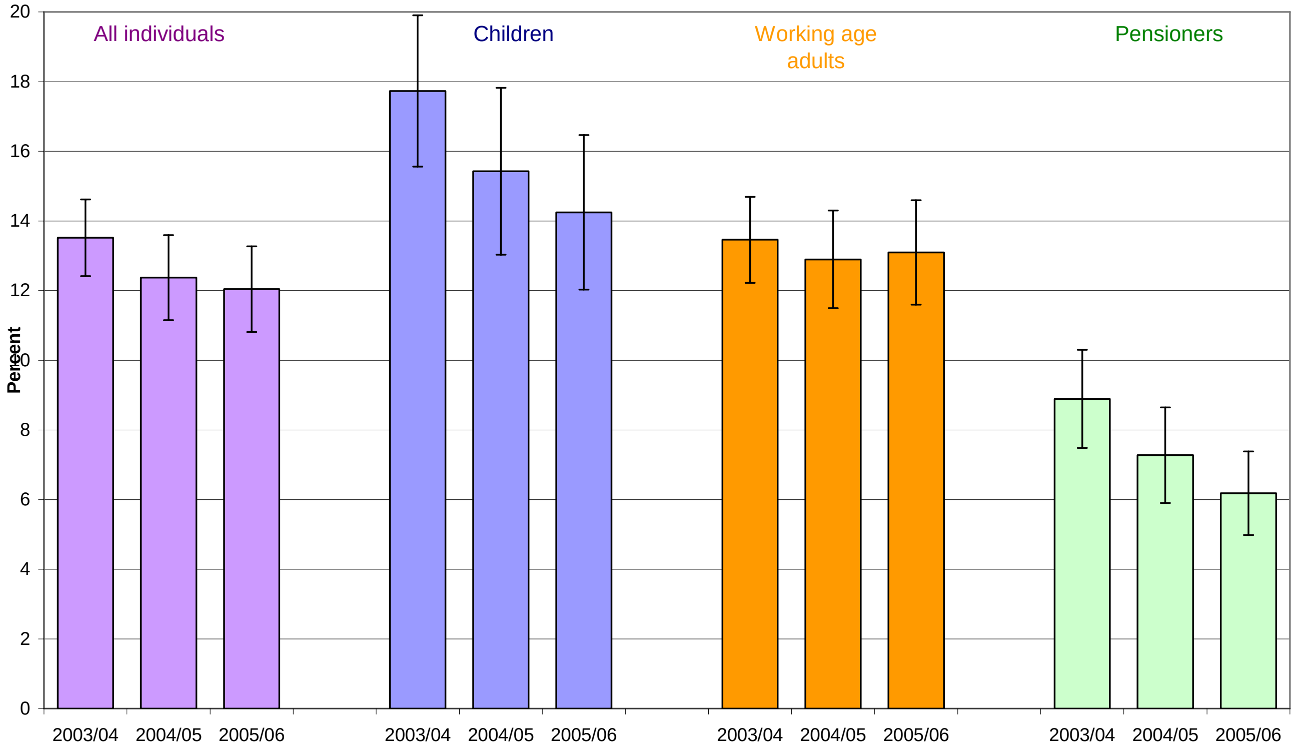
Table 2: Percentage and number of individuals in **relative** poverty with 90% confidence limits

		Percentage			Frequency (thousands)		
		Lower confidence limit	Estimate	Upper confidence limit	Lower confidence limit	Estimate	Upper confidence limit
Before Housing Costs							
All individuals	2003/04	11	12	13	560	610	670
	2004/05	10	11	12	510	560	620
	2005/06	10	11	12	490	550	620
Children	2003/04	13	15	17	130	160	180
	2004/05	10	13	15	110	130	150
	2005/06	10	12	15	100	130	150
Working Age Adults	2003/04	10	11	13	310	350	380
	2004/05	9	11	12	290	320	360
	2005/06	9	11	12	280	330	370
Pensioners	2003/04	10	12	14	90	110	120
	2004/05	10	12	14	100	110	130
	2005/06	9	11	13	90	100	120
After Housing Costs							
All individuals	2003/04	12	14	15	620	670	720
	2004/05	11	12	14	560	620	680
	2005/06	11	12	13	540	600	660
Children	2003/04	16	18	20	160	180	210
	2004/05	13	15	18	130	160	180
	2005/06	12	14	16	120	140	170
Working Age Adults	2003/04	12	13	15	370	410	440
	2004/05	11	13	14	350	390	440
	2005/06	12	13	15	360	400	450
Pensioners	2003/04	7	9	10	70	80	90
	2004/05	6	7	9	50	70	80
	2005/06	5	6	7	50	60	70

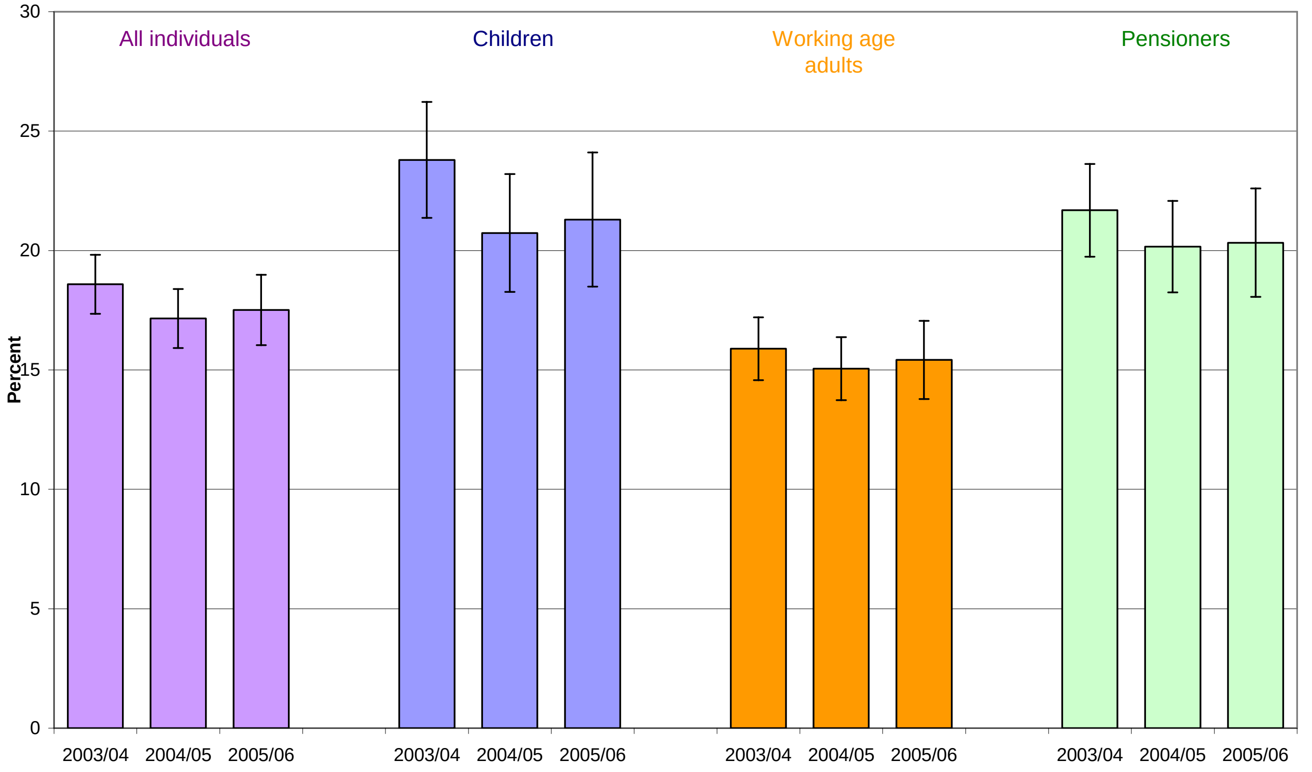
Graph 1: 90% Confidence intervals around the **absolute** poverty estimates **before housing costs**



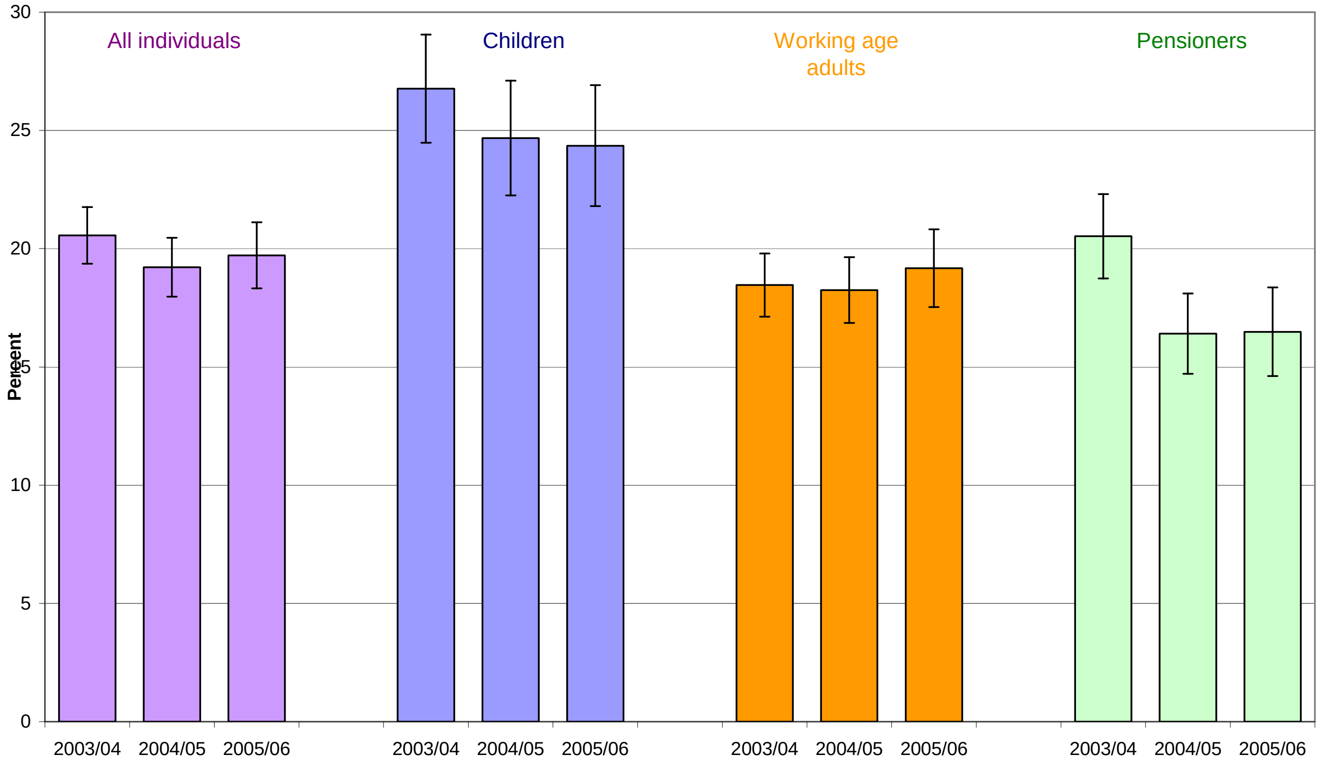
Graph 2: 90% Confidence intervals around the **absolute** poverty estimates **after housing costs**



Graph 3: 90% Confidence intervals around the **relative** poverty estimates **before housing costs**



Graph 4: 90% Confidence intervals around the **relative** poverty estimates **after housing costs**



METHOD

The methodology used to calculate the confidence intervals presented above was based on Binder and Kovacevic (1995) who provided a method for calculating confidence intervals around fractions of a median. This method was adjusted to allow for complications in the HBAI sampling strategy and methodology used to estimate numbers of people in poverty. The following explanation of the methodology is taken from 'Estimating uncertainty in measures of poverty via the Bootstrap' by Libby Cox.

Let M be the median, $L = 0.6 * M$ be the low-income threshold and $p = p(L)$ be the proportion of the sample falling below L. Then

1 $$Var(p) \approx (k^2/4 + p(1-p) - kp)/n,$$

where n is the number of households in the sample and k is defined as

2 $$k = 0.6 * f(L)/f(M)$$

and f is the probability density function of household income, such that

3 $$Prob(y < y_0) = \int_{-\infty}^{y_0} f(y)dy.$$

The density functions, $f(L)$ and $f(M)$, can be estimated indirectly as follows:

4 $$Var(M) \approx \left[1/\left(4 * f(M)^2\right)\right]/n$$

$$f(M) \approx 1/\left(2 * se(M) * \sqrt{n}\right).$$

Similarly,

5 $$Var(L) \approx p(1-p)/f(L)^2$$

$$f(L) \approx \frac{\sqrt{p(1-p)}}{se(L) * \sqrt{n}}.$$

In order to calculate the standard error of M, the confidence interval around the median can be estimated by treating the median as a percentile, M =0.5, and the standard error derived.

6 $$CI(M) = 1.96 * \sqrt{\frac{p(M) * (1 - p(M))}{n}}$$

$$se(M) = CI(M)/(2 * 1.96)$$

The same process can be carried out to calculate confidence interval around the proportion of people below the low-income threshold and hence the standard error and density function of L.

7 $$CI(L) = 1.96 * \sqrt{\frac{p(L)*(1-p(L))}{n}}$$

$$se(L) = CI(L)/(2*1.96)$$

Applying these results to equation 1, gives an estimate of the variance of $p(L)$, the proportion of households below the low-income threshold, given a simple random sample. This can then be compared to the variance of $p(L)$ obtained when L is assumed to be fixed. This can be calculated using

8 $$Var(p) = \frac{p(L)*(1-p(L))}{n}.$$

Using the results from equations 1 and 8, the ratio between the variance of p when L is allowed to vary and when L is fixed provides an estimate of the inflation due to the uncertainty in the threshold L.

9 $$VR = \frac{Var(p)}{Var(\widehat{p})}$$

Next, the variance of the proportion of a subgroup of the population in households below a fixed income threshold, $\breve{p}$, is calculated, e.g., the number of children living in households with low incomes in Scotland.

Let

$$\breve{p} = \frac{no.of\ poor\ people\, in\, subgroup}{no.of\ people\, in\, subgroup}.$$

For each household define

$$c_i = no.of\ subgroup\, in\, household\ i$$
$$d_i = c_i \qquad if\ poor$$
$$d_i = 0 \qquad otherwise$$
$$z_i = d_i - \breve{p} * c_i$$

Then calculate the simple variance of z_i across the subgroup and approximate the variance of $\breve{p}$, given that $\breve{n}$ is the number of households in the subgroup sample, i.e. the number of Scottish households, and m is the number of subgroup individuals in the subgroup sample, i.e. children living in Scottish households.

10 $$Var(z_i) = \frac{\sum_{i=1}^{n} (z_i - \bar{z})^2}{\breve{n}}$$

$$Var(\breve{p}) \approx Var(z) * \breve{n} / m^2 \ ,$$

Having estimated $Var(\breve{p})$, the variance for the subgroup of interests, given by equation 6 above, this should now be inflated using the variance ratio for uncertainty around the threshold given by 5. It should also be inflated to account for the complex sample design of the FRS using a design effect. The design effect was calculated to be 1.1, and hence the resultant variance and confidence interval for the statistics of interest can be derived by

11 $$Var(p) = Var(\breve{p}) * VR * 1.1$$

$$CI(p) = p \pm 1.96 * \sqrt{Var(p)}$$