NatCen
National Centre for Social Research

Sampling Issues for a Fifth WERS

Susan Purdon
Director, Survey Methods Centre
35 Northampton Square
London EC1V 0AX

Prepared for Department for Trade and Industry

Dec 2002

Introduction

The sponsors of WERS98 have carried out a consultation exercise on the design and scope of another WERS. As part of that exercise the National Centre for Social Research were commissioned to do a small piece of work on three sample design issues:

(i) the implications of lowering the WERS threshold from 10 to 5 employees
(ii) the sample design options for separate analysis by country
(iii) the sample design options for separate analysis by Government Office region.

This paper describes the conclusions of that work.

The current WERS design

The main features of the current WERS design are:

- there is a minimum establishment size threshold of 10 employees;
- large establishments are greatly over-sampled relative to smaller establishments (which is to allow for separate reporting by size)
- some SIC groups (E, F, H, J and O) are over-sampled relative to other groups
- there is, at present, no over-sampling by country or region.

The total issued sample size for the cross-sectional design in WERS98 was 3192, which after excluding ineligibles and refusals gave an achieved sample size of 2193.

Approach

To compare alternative designs we have made use of the 'effective sample size'. The effective sample size is the simple random sample size that would give equivalent precision to WERS. For instance, as noted above, WERS98 yielded a total achieved sample size of 2193. But the *effective* sample size for estimates based on all establishments is estimated to be just 623. This means that the precision achieved by WERS, with its unequal probability design, could have been achieved with an equal probability simple random sample size of just 623. Most of the 'loss' in effective sample size is attributable to the over-sampling of larger establishments. (There are of course very good reasons why WERS uses an unequal probability design, and we would certainly not recommend a move to simple (equal-probability) random sampling. But the effective sample size allows designs to be compared easily – hence its use here).

Another term used in this paper, related to the effective sample size, is the *design factor*. The design factor is the (multiplicative) increase in standard errors associated with the use of a complex sample design rather than a simple random sample design. For instance, a design factor of 1.2 implies that the standard error is 20% larger than it would be for an equally sized simple random sample. The design factor is calculated as the square root of the sample size divided by the effective sample size.

In making estimates of the effective sample size and/or design factors under different designs we have taken into account the effect of differential sampling rates by size, but not by SIC. The effects of the latter are relatively small, and to take them into account would make the analysis here unnecessarily complicated[1].

Secondly, in calculating effective sample sizes we have assumed that the survey estimate being measured does not differ greatly between size groups (or, for that matter, between countries or regions). This is obviously an over-simplification – for almost all WERS estimates there seems to be some relationship with size. Making this assumption means that for some survey variables the effective sample size will be over-estimated, and for others the effective sample size will be under-estimated. In a very crude sense the assumption of no difference in the survey estimate by size might be interpreted as giving a 'central' estimate of the effective sample size. (The alternative approach would be to estimate the effective sample size more accurately for a small range of WERS variables using WERS98 data, but, as is demonstrated below, this does not alter the conclusions reported here.)

As well as presenting the effective sample size we have also, at the request of the DTI, given some indication of the likely impact of alternative designs on standard errors. In calculating these standard errors a similar assumption of no difference in the survey estimate by size is used.

The formulae used to calculate the effective sample size and the standard errors are given in the appendix.

It should be noted that this simplistic method of estimating the 'average' or 'central' effective sample size or design factor is not the approach used to estimate the published 'average' design factors[2]. The published estimates of 'average' design factors are estimated by calculating the design factors for a large number of variables and taking the average across these. The two methods do give different results, but the simpler method used in this report does give a figure within the central range of design factors used in the published average.

The implications of lowering the WERS threshold from 10 to 5 employees (Table 1)

To assess the implications of lowering the threshold for WERS from 10 to 5, we have assumed that the issued sample size for the 5-9 group will be about 350 (i.e similar to the sample size for the 10-24 group). On the assumption that the rate of ineligibles and refusals for this size group will be similar to those for the 10-24 group, this 350 will yield about 175 productive interviews[3].

---

[1] Analysis suggests that for most survey variables standard errors increase by about 3% if the differential sampling by SIC is taken into account. This is trivial compared to an increase of about 80% when the differential sampling by size is taken into account.

[2] Airey, C. *et al* (1999) "The Workplace Employee Relations Survey (WERS) 1997-8, Technical Report", National Centre for Social Research.

[3] The 1992 Employer's Recruitment Practice Survey gave a response rate for establishments with 3-10 employees of 69%, compared to 77% overall and 72% in the 11-24 category. So there was no very strong gradient in response rates by size. This tallies with response rate patterns observed in previous WERS for 10+ establishments, and it seems reasonable to assume a

Two possibilities for changing the design have been considered:
Design 2: keep the same overall issued sample size as WERS98 (This is Design 2 of Tables 1a and 1b)
Design 3: boost the overall sample size by the 350 in the 5-9 group and keep the rest of the WERS design the same as in WERS98. This gives a total issued sample size of 3542.
(Design 1 is the current WERS design)

Table 1a sets out the issued and achieved sample sizes by size of establishment. Table 1b gives the effective sample sizes and illustrative standard errors for two survey estimates, one of 50% and one of 10%. (In all the tables 'size' is based on IDBR employment data rather than on interviewer-collected size data.)

Design 3 is clearly better from a statistical viewpoint in that it incurs no loss of precision in the 10+ establishment estimates.

However, if Design (2) is adopted there are a number of implications (Table 1b):

- the effective sample size for 10+ establishment estimates reduces from 623 to about 554 (which we would not consider to be a particularly serious reduction – it increases standard errors for percentages by about 0.1 percentage points);
- the effective sample size for 5+ establishment estimates is about 640 under Design 2, which is slightly greater than the effective sample size for 10+ establishments under the current design (at 623). The implication of this is that 5+ establishment estimates will have very slightly greater precision that 10+ establishment estimates in WERS98. Put another way, the design effects will be slightly lower.

The fact that, relative to WERS98, a slightly larger effective sample size is achieved under a design where the 5-9 establishments are included is perhaps rather surprising. The reason is essentially that, whereas in WERS98 just 11% of the sample covered the 57% of the population in the smallest size category (namely the 10-24 group), under Design 2, 21% of the sample will cover the 76% of the population in the smallest size category (5-24). Thus the 'under-representation' of the smallest establishments is less severe under Design 2 than under the current WERS design. It

---

similar pattern in response if a 5-9 group is introduced. This does not, however, address the question of the ineligibility rate. We have not been able to locate any figures on the number of ineligibles (i.e. businesses that no longer exist) that you would expect to find in an IDBR sample of the 5-9 group. Since most of the losses between the issued and productive samples are likely to be because of ineligibles rather than non-response, in deciding on the size of sample to select it is more important to predict the ineligible rate properly rather than the response rate. If no data on eligiblity rates can be found the best option for WERS might be to take a fairly large sample from the 5-9 group but then to divide this into two groups: a starting sample of, say, 350, and a reserve sample. If the ineligibility rate in the starting sample is higher than anticipated then the whole, or a proportion of, the reserve sample would be issued.

should be stressed here that there would *not* be a gain in the effective sample size if the sample size for the 5-9 group was set much lower than the proposed 350.

All of this suggests that, although Design 3 is a better option than Design 2 if it can be afforded, Design 2 should be acceptable to most WERS users.

It *might* be thought that the conclusions we have reached are sensitive to the way the effective sample size has been calculated. To illustrate that this is not the case we have taken a small range of estimates from WERS98 for 10+ establishments for which design factors were published, and then calculated what the (approximate) design factor would be for 5+ establishments if the total WERS sample size stays the same but 350 are used to cover the 5-9 group (i.e. design 2). In each case the survey estimate for the 5-9 category is varied (three times) to illustrate the impact such variation has on the design factor[4].

The results are shown in Table 1c, the three estimates for the 5-9 group being labelled Variations 1, 2 and 3. For example, the variable '% of establishments with no union present' gave a figure of 53.4% in WERS98 for 10+ establishments. This estimate had a design factor of 1.66. If the percentage for the 5-9 group was 53% then the design factor for 5+ establishments would increase to 1.87; if the percentage for the 5-9 group was just 20% then the design factor for 5+ establishments would stay at 1.66; and if the percentage for the 5-9 group was as high as 90% then the design factor for 5+ establishments would actually decrease to 1.51. But, even though this demonstrates that the design factors will change if the 5-9 group is included, the conclusion appears to be that they will not change dramatically. The same holds for the other variables considered. This is consistent with the conclusion reached above.

The sample design options for separate analysis by country

In considering the options for separate analysis by country and region we have assumed that the minimum size threshold for WERS stays at 10 employees. Our overall conclusions would not change if the lower threshold was adopted.

The sample sizes in Wales and Scotland under the current WERS design are far too low for separate reporting (Design 1C of Table 2a), the effective sample size for Wales being just 26 and for Scotland just 54. The alternatives we have considered are:

Design 2C: Maintaining the same overall sample size as in WERS98 but taking equal samples sizes per country. This has the effect of reducing the effective sample size for England from 543 to just 208, which we do not believe would be acceptable. In addition standard errors for 'all GB' estimates would increase by about 1 percentage point.

Design 3C: Maintaining the same overall sample size but setting the issued sample size at 500 in Wales and Scotland. This still leads to quite a large reduction in the effective sample size for England (to 428), but is better than Design 2C. The impact of

[4] A pessimistic design factor of 1.2 for the 5-9 group is assumed to allow for the disproportionate sampling by industry.

this design would be to increase standard errors for GB estimates by about 0.2 percentage points.

Design 4C: Increasing the overall sample size to give the same effective sample size for England as in WERS98 but giving issued sample sizes of 500 in Wales and Scotland. This would mean increasing the overall issued sample size from 3192 to 3783. Under this design standard errors for GB estimates would reduce very slightly (relative to WERS98).

In conclusion, of the three designs considered we would recommend something along the lines of Design 4C, but believe that Design 3C would be a reasonable compromise solution.

The sample design options for separate analysis by region

The current sample design gives issued sample sizes that vary quite considerably by government office region (GOR) from just 128 in the NE to 514 in London. Arguably only the very largest of these are large enough to allow for separate reporting, and even for these the effective sample size is fairly small at around 90 (Table 3a).

We have considered two options for re-distributing the sample by GOR, both of which keep to the current issued sample size of 3192:

Design 2R: Setting the issued sample size in each GOR at 290
Design 3R: Setting the minimum issued sample size per GOR at 250 (and re-distributing the rest of the sample pro-rata).

Neither of these options is particularly damaging to 'all GB' estimates, with standard errors increasing by about 0.2 percentage points at worst. But, neither of these options gives effective sample sizes per GOR that are large enough for separate reporting. (The minimum issued sample size of 250 gives an effective sample size for estimates based on all establishments (10+) of just 53. For estimates based on 25+ sized establishments the effective sample size would be larger at about 90, but this would still be too small for many purposes.)

The conclusion appears to be that if regional analysis of WERS data is a requirement then the total sample size for WERS would need to be increased quite considerably (perhaps to about 5500, to give 500 per region. This would give an effective sample size per region of about 100).

Conclusions

In summary:

(i) Reducing the threshold for WERS to 5 employees seems perfectly feasible, although the best way to do it would be to boost the overall sample size to include this group. This way the standard errors for estimates based on larger establishments will not be increased.

(ii) The issued sample size for Wales and Scotland could be increased to about 500 without a great loss of precision in England or GB estimates.

(iii) Changing the sample design to allow for separate reporting by region does not appear to be feasible with the current sample size. Regional estimates would require a considerably larger overall sample size.

Table 1a: Alternative designs for including establishments with 5-9 employees

Establishment size	Population numbers	Design 1: Current WERS design		Design 2: Same sample size as WERS98 but with 350 in 5-9 category		Design3: Current WERS design but with additional 350 in 5-9 category	
		Issued sample size	Achieved sample size	Issued sample size	Achieved sample size	Issued sample size	Achieved sample size
5-9	311360	0	-	350	175	350	175
10-24	233150	362	182	322	162	362	182
25-49	91250	603	406	537	362	603	406
50-99	44135	566	375	504	334	566	375
100-199	20435	562	398	500	354	562	398
200-499	11260	626	476	557	424	626	476
500+	3840	473	356	421	317	473	356
Total sample size		*3192*	*2193*	*3192*	*2127*	*3542*	*2368*
Sample size for 10+		*3192*	*2193*	*2842*	*1952*	*3192*	*2193*
Sample size for 25+		*2830*	*2011*	*2520*	*1790*	*2830*	*2011*

Table 1b: The effect of alternative designs on the effective sample size and on standard errors

	Design 1: Current WERS design	Design 2: Same sample size as WERS98 but with 350 in 5-9 category	Design3: Current WERS design but with additional 350 in 5-9 category
Total achieved sample size	2193	2127	2368
Achieved sample size for 10+	2193	1952	2193
Achieved sample size for 25+	2011	1790	2011
Effective sample size for estimates based on establishments with 5+ employees	-	640	670
Effective sample size for estimates based on establishments with 10+ employees	623	554	623
Effective sample size for estimates based on establishments with 25+ employees	1107	986	1107
Estimates based on establishments with 5+ employees			
Approx. standard error for an estimate of 50%	-	2.0%	1.9%
Approx. standard error for an estimate of 10%	-	1.2%	1.2%
Estimates based on establishments with 10+ employees			
Approx. standard error for an estimate of 50%	2.0%	2.1%	2.0%
Approx. standard error for an estimate of 10%	1.2%	1.3%	1.2%
Estimates based on establishments with 25+ employees			
Approx. standard error for an estimate of 50%	1.5%	1.6%	1.5%
Approx. standard error for an estimate of 10%	0.9%	1.0%	0.9%

Table 1c: Estimated impact on design factors of including a 5-9 group (Design 2)

	Percentage in 1998 for 10+ establishments	Design factor in WERS98 for 10+ establishments	Variation 1		Variation 2		Variation 3	
			Assumed percentage for 5-9 establishments	Estimated design factor for 5+ establishments	Assumed percentage for 5-9 establishments	Estimated design factor for 5+ establishments	Assumed percentage for 5-9 establishments	Estimated design factor for 5+ establishments
% with no union present	53.4%	1.66	53%	1.87	20%	1.66	90%	1.51
% accredited as an Investor in People	33.9%	1.82	34%	1.93	10%	1.68	50%	1.94
% with no part-time employees	17.9%	1.88	18%	1.96	10%	1.84	50%	1.95
% with formal procedures for dealing with discipline	87.6%	2.23	88%	2.12	50%	2.05	20%	1.60

Table 2a: Alternative designs for separate reporting by country

Country	Population numbers	Design 1C: Current WERS design			Design 2C: Re-distribution of current sample size to give equal numbers per country		
		Issued sample size	Achieved sample size	Effective sample size	Issued sample size	Achieved sample size	Effective sample size
England	351410	2783	1912	543	1064	731	208
Wales	17245	129	89	26	1064	731	212
Scotland	35370	280	192	54	1064	731	206
All		*3192*	*2193*	*623*	*3192*	*2193*	*271*

Country	Population numbers	Design 3C: Current sample size but with minimum of 500 per country			Design 4C: Increased sample size to give minimum of 500 in Wales and Scotland but with no loss of precision for England estimates		
		Issued sample size	Achieved sample size	Effective sample size	Issued sample size	Achieved sample size	Effective sample size
England	351410	2192	1506	428	2783	1912	543
Wales	17245	500	344	100	500	344	100
Scotland	35370	500	344	96	500	344	96
All		*3192*	*2193*	*536*	*3783*	*2599*	*671*

Table 2b: The effect of alternative designs for country on 'all country' standard errors

	Design 1C: Current WERS design	Design 2C: Re-distribution of current sample size to give equal numbers per country	Design 3C: Current sample size but with minimum of 500 per country	Design 4C: Increased sample size to give minimum of 500 in Wales and Scotland but with no loss of precision for England estimates
Total sample size (10+)	2193	2193	2193	2599
Effective sample size	623	271	536	671
Approx. standard error for an estimate of 50%	2.0%	3.0%	2.2%	1.9%
Approx. standard error for an estimate of 10%	1.2%	1.8%	1.3%	1.2%

Table 3a: Alternative designs for separate reporting by GOR

Region	Population numbers	Design 1R: Current WERS design			Design 2R: Re-distribution of current sample size to give equal numbers per region			Design 3R: Current sample size but with minimum of 250 per region		
		Issued sample size	Achieved sample size	Effective sample size	Issued sample size	Achieved sample size	Effective sample size	Issued sample size	Achieved sample size	Effective sample size
NE	15840	128	88	24	290	199	60	250	172	52
NW	45990	362	249	71	290	199	61	324	223	68
Y&H	34150	266	183	53	290	199	61	250	172	53
EM	29215	244	168	45	290	199	62	250	172	54
WM	36910	295	203	59	290	199	61	250	172	53
East	37270	285	196	57	290	199	63	250	172	54
London	58835	514	353	93	290	199	56	461	317	89
SE	58505	454	312	89	290	199	61	407	280	86
SW	34695	256	176	52	290	199	63	250	172	54
Wales	17245	129	89	26	290	199	62	250	172	54
Scotland	35370	280	192	54	290	199	60	250	172	52
All		*3192*	2193	623	*3190*	2192	585	*3192*	2193	637

Table 3b: The effect of alternative designs for region on 'all country' standard errors

	Design 1R: Current WERS design	Design 2R: Re-distribution of current sample size to give equal numbers per region	Design 3R: Current sample size but with minimum of 250 per region
Total sample size (10+)	2193	2192	2193
Effective sample size	623	585	637
Approx. standard error for an estimate of 50%	2.0%	2.1%	2.0%
Approx. standard error for an estimate of 10%	1.2%	1.2%	1.2%

Appendix: Calculation of the effective sample size and the approximate standard error

The effective sample size (eff) is calculated using the formula

$$eff = \frac{\left(\sum_h n_h w_h\right)^2}{\sum_h n_h w_h^2}$$

where n_h is the achieved sample size in stratum h, and

w_h is the weight applied to the sample members of stratum h.

The weight is calculated as equal to the population count divided by the issued sample size.

The standard error around a percentage p is estimated as

$$s.e.(p) = \sqrt{\frac{p(100-p)}{eff}}.$$

This is reasonably accurate as long as p does not vary considerably between strata. If p does vary between strata then the standard error using this formula will sometimes be too small and sometimes too large.