Chapter 4

Design effects at sampling phase

The 'design effect' is a useful and fairly compact term to indicate the influence of each sampling design on the uncertainty of each particular estimate. This effect can be finally estimated in the proper survey analysis when the data and all required design variables are available. However, the same things should have been tried to predict or anticipate as well as possible when designing and implementing the sampling. This chapter is focused on giving the core instruments for this anticipation. First, the key terms are presented.

The general formula for the design effect, symbolised by *DEFF*, is as follows:

$$DEFF = \text{var}(s) / \text{var}(srs)$$

in which $\text{var}(s)$ = variance estimate of the particular sampling design s.

$\text{var}(srs)$ = variance estimate of the simple random design, or the reference design.

The effect *DEFF* thus is a relative ratio. If it is higher than 1 then the design *s* is less efficient than the *srs* design, otherwise more efficient. The *DEFF* thus is quadratic. If the standard errors are wished to compare the square root is needed. This is called *DEFT* = the square root of DEFF.

The design effect in sampling designing is good to divide into the two components:

- *DEFF* due to clustering = *DEFFc*
  and
- *DEFF* due to varying inclusion probabilities = *DEFFp*.

Their product is the whole design effect

$$DEFF = DEFFc * DEFFp$$

Next we explain how each of these components has been anticipated for the sampling design.

DEFF due to clustering

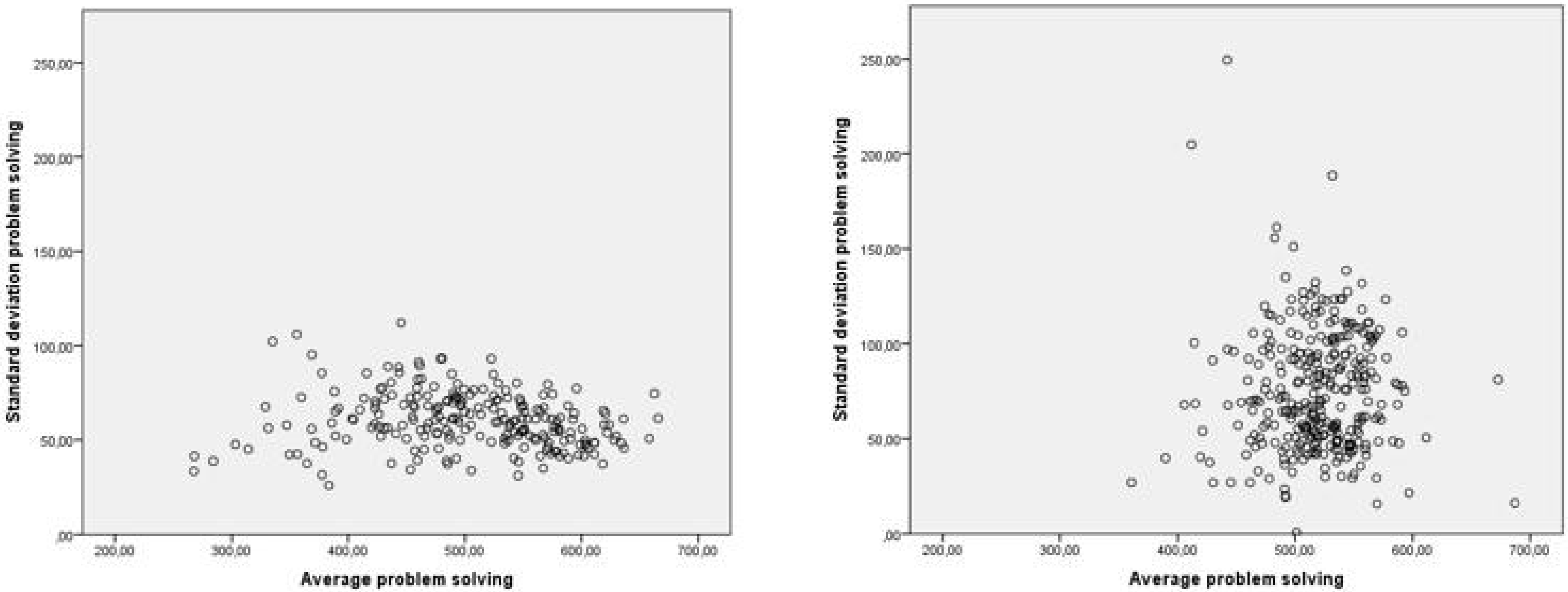

Figure 4.1 Scatter plots between average scores of problem solving and their standard deviations in the Pisa, Germany (left) and Finland (right). Plots are primary sampling units = schools

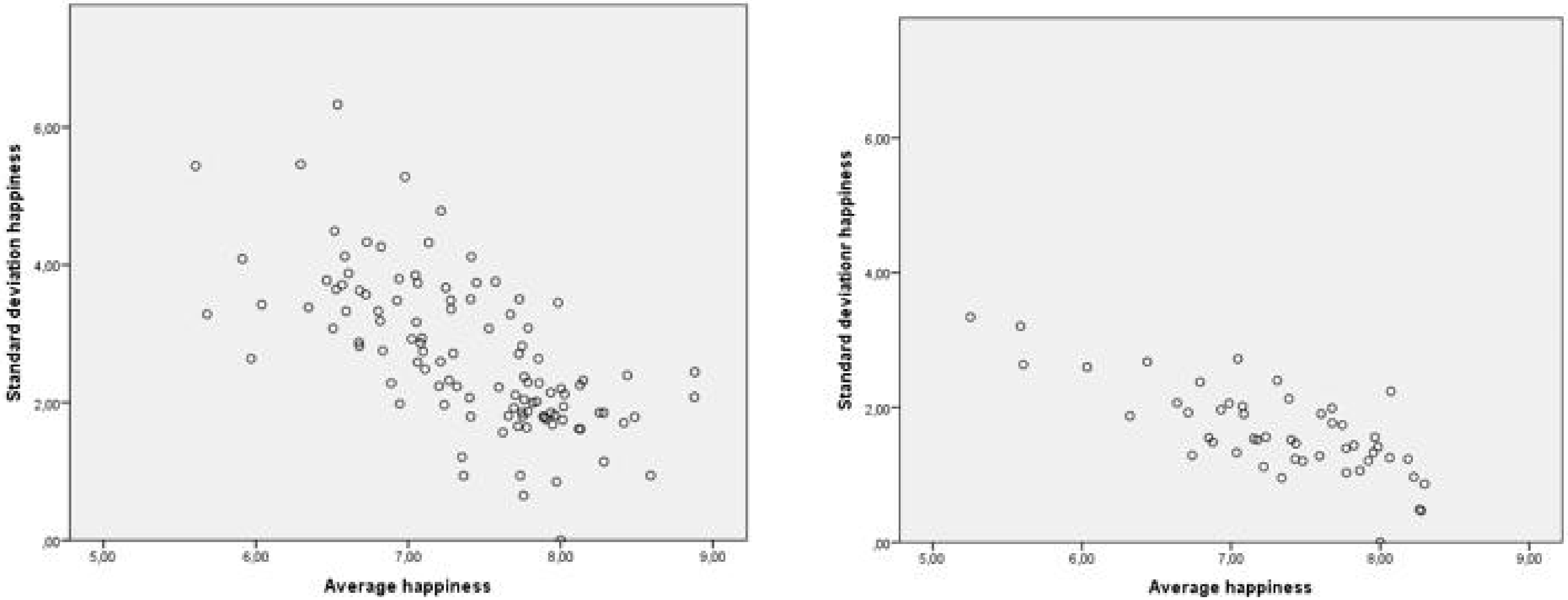

Figure 4.2 Scatter plots between average happiness and standard deviation of happiness in stratum 1 (left) and stratum 2 (right). Plots are primary sampling units = small areas.

Indicator for homogeneity versus inhomogeneity
Is The intraclass correlation (rho) that can be estimated most easily using linear mixed model; rho = (0, 1)

Table 4.1. Intraclass correlations in the cases of Figures 4.1. and 4.2.

Variable	**Intraclass correlation**
Pisa Problem solving for Germany	0.617
Pisa Problem solving for Finland	0.141
Happines of stratum 1 in the test data	0.006
Happines of stratum 2 in the test data	0.050

DEFFc = 1 + (b-1) rho in which *b*=average net cluster size

Table 4.2. Intraclass correlations and the DEFF's due to clustering for some variables in the ESS countries of two-stage cluster design, in round 5.

Variable	Rho	DEFFc
Robbery rate	0.0294	1.43
Robbery fear	0.0198	1.29
Opinion: mothers should stay more at home, not to work	0.0287	1.42
Opinion: talented students should be more awarded	0.0028	1.04
Happiness	0.0005	1.01

DEFF due to varying inclusion probabilities = *DEFFp*.

In the sampling stage, the *DEFFp's* are reasonable to anticipate using the expected weights in this formula

$$DEFFp = 1+cv(w)^2$$

in which *cv(w)* *is* the coefficient of variation of the weight *w*. The first approximation of the weight *w* could be the

Next two pages: examples of the 2012 PISA

Country code 3-character	Weight	Mean	Coeff of Variation	Sum	DEFFp
Macao-China	PISA STUDENT WEIGHT	1,0	1,7	5366	1,00
Korea	PISA STUDENT WEIGHT	119,9	13,1	603632	1,02
France	PISA STUDENT WEIGHT	151,7	15,1	699779	1,02
Turkey	PISA STUDENT WEIGHT	178,8	16,7	866681	1,03
Sweden	PISA STUDENT WEIGHT	20,0	18,4	94936	1,03
Norway	PISA STUDENT WEIGHT	12,7	19,2	59432	1,04
Japan	PISA STUDENT WEIGHT	177,6	20,2	1128179	1,04
Bulgaria	PISA STUDENT WEIGHT	10,3	20,5	54255	1,04
Serbia	PISA STUDENT WEIGHT	14,5	21,9	67934	1,05
Shanghai-China	PISA STUDENT WEIGHT	16,4	22,7	84965	1,05
Germany	PISA STUDENT WEIGHT	151,4	22,8	756907	1,05
Malaysia	PISA STUDENT WEIGHT	83,1	22,9	432080	1,05
Croatia	PISA STUDENT WEIGHT	9,1	24,6	45506	1,06
Singapore	PISA STUDENT WEIGHT	9,2	25,0	51088	1,06
Hong Kong-China	PISA STUDENT WEIGHT	15,1	25,2	70636	1,06
Poland	PISA STUDENT WEIGHT	82,4	27,1	379545	1,07
Israel	PISA STUDENT WEIGHT	21,4	27,5	107990	1,08
Uruguay	PISA STUDENT WEIGHT	7,5	28,7	39771	1,08
Chinese Taipei	PISA STUDENT WEIGHT	48,4	30,3	292542	1,09
United Kingdom	PISA STUDENT WEIGHT	138,5	31,8	579422	1,10
Ireland	PISA STUDENT WEIGHT	10,8	32,3	54010	1,10
Russian Federation	PISA STUDENT WEIGHT	224,5	34,8	1174528	1,12
Slovak Republic	PISA STUDENT WEIGHT	11,7	36,7	54636	1,13

Austria	PISA	STUDENT WEIGHT	17,3	36,9	82225	1,14
Belgium	PISA	STUDENT WEIGHT	13,7	38,9	117889	1,15
Montenegro	PISA	STUDENT WEIGHT	1,6	40,5	7714	1,16
United States of America	PISA	STUDENT WEIGHT	710,9	41,2	3538783	1,17
Hungary	PISA	STUDENT WEIGHT	19,0	43,3	91179	1,19
Estonia	PISA	STUDENT WEIGHT	2,4	44,0	11627	1,19
Netherlands	PISA	STUDENT WEIGHT	44,0	47,3	196262	1,22
Portugal	PISA	STUDENT WEIGHT	16,8	51,2	96034	1,26
Chile	PISA	STUDENT WEIGHT	33,4	52,4	229159	1,27
Australia	PISA	STUDENT WEIGHT	17,3	58,3	250711	1,34
Italy	PISA	STUDENT WEIGHT	95,0	64,4	521902	1,41
Czech Republic	PISA	STUDENT WEIGHT	15,4	71,2	82250	1,51
United Arab Emirates	PISA	STUDENT WEIGHT	3,5	71,4	40612	1,51
Brazil	PISA	STUDENT WEIGHT	435,3	72,4	2397036	1,52
Finland	PISA	STUDENT WEIGHT	6,8	79,3	60047	1,63
Slovenia	PISA	STUDENT WEIGHT	3,1	80,6	18418	1,65
Denmark	PISA	STUDENT WEIGHT	8,8	81,7	65642	1,67
Spain	PISA	STUDENT WEIGHT	36,4	101,1	370862	2,02
Colombia	PISA	STUDENT WEIGHT	61,7	101,9	559674	2,04
Canada	PISA	STUDENT WEIGHT	16,2	111,1	347987	2,23

The concept 'Effective sample size' (neff) is a useful term to benchmark

The *neff* corresponds to the sample size in which case the micro data of respondents could give the same accuracy as the simple random design (*srs*) gives. Thus if the net sample data can really be interpreted to be drawn from a target population with srs, we do not need even sampling weights for getting good accuracy estimates. Unfortunately, this is not the case in real life but it is still good to compare the achieved data relative to *srs* data.

Table 4.3 Sampling design summary for determining gross sample size if the effective target sampling size is 1500

Operation	**Example calculation** (average-based, the figures may vary by stratum, cluster and another domain)
1. Target for the effective sample size (*neff)*	**1500**
2. Anticipated missingness due to ineligibility (7%)	1500/0.93 = 1613
3. Anticipated missingness due to unit nonresponse (35%)	1613/0.65 = 2481
4. Anticipated Design Effect due to clustering (*DEFFc*) including anticipated intra-class correlation (=0,025), if the average gross cluster size = 8 then the average net cluster size =8*0.93*0.65= 4.83	*DEFFc* = 1+(4.83-1)*.025 = 1.096 2481*1.096= 2719
5. Anticipated due to varying inclusion probabilities (*DEFF*p) (calculated for anticipated respondents if possible)	*DEFFp* = 1.2 2719*1.2 = 3263
6. Anticipated and realized minimum Gross Sample Size Anticipated Net Sample Size	**3263** 1973 =3263*0.65*0.93

How to decide the sample size and allocate the sample into strata?

The effective target sample size is already decided in Table 4.3 that corresponds to the minimum requirement of the ESS. This decision is based on many criteria. One is that the accuracy of each country estimate would be at about the same level. This gives good opportunities to compare the results between countries. Since the minimum requirement of the respondents is 1500, appropriate estimates between many population groups or domains are possible, but these might be limited if the number of the respondents is not reasonable.

The ESS is a general social survey without any specific target at domain level. If a survey has such the target, the best solution is to use stratification and to allocate each stratum so that the accuracy target is achieved. This can be called **minimum allocation**. The other common allocations are:

- **Proportional** that thus was recommended for the ESS since round 6 but might be changed since this does not accept to take into account the response rate anticipation; it is well known that response rates are lower in urban areas than in rural.

- **Equal** allocation in which case the *neff* of each stratum is equal; the ESS requirement by country thus corresponds to this allocation.

- **Neyman-Tschuprow** allocation in which the gross sample size (or neff) n_h of each stratum *h* is obtained with the formula

$$n_h = n \frac{N_h S_h}{\sum_h N_h S_h}.$$

Here *n* = the desired sample size of the whole sample, n_h = the respective gross sample size of stratum *h*, N_h = the size of the frame population in stratum *h*, and S_h = anticipated standard deviation of the study variable *y*. This standard deviation can be taken from

Various design effects for 'problem solving' in the 2012 PISA, selected countries

Country code 3-character	DEFF Stratifi-cation	DEFF Weights	DEFF Clusters	DEFF All
DEU	1	0.97	13.45	12.11
EST	0.937	1.19	7.98	6.35
FIN	0.972	1.38	5.61	5.25
JPN	0.978	1.04	14.28	14.31
KOR	0.919	1.02	11.94	9.40
NOR	0.999	1.06	6.85	6.78
RUS	1	1.21	11.46	13.38
SWE	0.985	1.06	6.22	6.30
USA	1	1.26	10.51	13.56