Interacting nonresponse and response errors

Håkan L Lindström

SCB

R&D Report
Statistics Sweden
Research - Methods - Development
1991:3

Från trycket	Mars 1991
Producent	Statistiska centralbyrån, utvecklingsavdelningen
Ansvarig utgivare	Åke Lönnqvist
Förfrågningar	Håkan L Lindström, 019-17 66 94

ISSN 0283-8680
Printed in Sweden

INLEDNING

TILL

R & D report : research, methods, development / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1988-2004. – Nr. 1988:1-2004:2.
Häri ingår Abstracts : sammanfattningar av metodrapporter från SCB med egen numrering.

Föregångare:

Metodinformation : preliminär rapport från Statistiska centralbyrån. – Stockholm : Statistiska centralbyrån. – 1984-1986. – Nr 1984:1-1986:8.

U/ADB / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1986-1987. – Nr E24-E26

R & D report : research, methods, development, U/STM / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1987. – Nr 29-41.

Efterföljare:

Research and development : methodology reports from Statistics Sweden. – Stockholm : Statistiska centralbyrån. – 2006-. – Nr 2006:1-.

R & D Report 1991:3. Interacting nonresponse and response errors / Håkan L Lindström.
Digitaliserad av Statistiska centralbyrån (SCB) 2016.

urn:nbn:se:scb-1991-X101OP9103

INTERACTING NONRESPONSE AND RESPONSE ERRORS

Håkan L Lindström, U/STM

CONTENT: Page

Håkan L Lindström
U/STM

SUMMARY:

Fairly often a high response rate in a survey is thought to indicate that nonresponse errors are of negligable size. It is then consistent to believe that steps taken to reduce the nonresponse rate will always reduce the mean square error of an estimator. The present report demonstrates that this is not the case, and stresses the necessity that strategies intended to reduce nonresponse must be checked with regard to measurement errors as well.
Response errors are apt to appear when a respondent is under some pressure to provide an answer. The pressure may consist in a legal obligation to answer, the persuasive power of an interviewer or in just lack of time or interest. It is important to have an idea about the size and distribution of response errors and their interaction with response probabilities in order to protect the quality of the completed survey and use available resources in an efficient way.

Lindström and Lundström (1974) studied nonresponse errors by using a **response probabilities** concept to find out how much the distribution of the nonrespondents would influence the expectation and standard error of the unweighted sample average when the sample was a simple random sample (s r s) from a continuous distribution. They studied a case where each unit was attributed a response probability, which was a function of its value. The bias of the estimator was studied for several types of distributions and at varying levels and distributions of nonresponse. The nonresponse bias fluctuated greatly depending on the variation in response probabilities even when the nonresponse rate was fixed. They also could point out conditions when the nonresponse bias increased although the nonresponse rate dropped.

The present report proceeds along the same lines but includes models to evaluate effects of both nonresponse bias and response errors, allowing the response errors to have both random and nonrandom components. Chapter 2 summarizes and extends the results of the 1974 report. Chapters 3 - 5 add response errors to the model. The proposed models permits us to analyze the influence on the Mean Square Error (MSE) of both the distribution of the respondents and of the response errors that in some way depend on the actions taken to reduce the nonresponse rate.

Although the error functions of this study are modeled with the intention to be realistic approximations with some empirical support, the approach at this stage is mainly theoretical. The models, and still more the calculations, have been used **to point out cases where an unweighted estimator is sensitive to nonresponse errors and changes in the measurement procedure and cases where it is robust.** As there is very little evidence to indicate reasonable values of response errors, empirical studies are needed to quantify the size of the errors involved, before the ideas can be applied to the development of strategies for nonresponse follow-up and data collection that will result in a more efficient use of resources in surveys.

1 THE EMPIRICAL BACKGROUND

1.1 PRACTICES TO REDUCE NONRESPONSE

A low nonresponse rate in a survey is often thought to guarantee small nonresponse bias of the estimators - although this obviously is not always true - and in practically all surveys, procedures are instituted to reduce the nonresponse rate to a minimum.

The design of such procedures is normally a basic and integrated part of the planning work. However, there are cases where the nonresponse mounts to unforeseen levels, and where supplementary efforts have to be introduced during the data collection period. For such cases it might be useful to distinguish between the **planned design** and the **realized design** when the outcome of a survey is described in terms of the distribution of the sample on respondent groups and the combination of data collection techniques used.

Procedures used to persuade *up to then nonrespondents* include renewed and intensified attempts to contact them. Some different actions are reminders by mail or telephone, call-backs with intensified persuasion, use of specially skilled interviewers, use of incentives, use of data collection methods adapted to the nonresponse-prone group, and prolonged data collection period. Experienced researchers knew well that additional efforts and the adaption of the data collection method to special groups of the sample can increase the response rate substantially. But the choice of follow-up method can influence the distribution of the *new respondents* as well as their consent or capacity to give correct answers. There may be a continuous increase of the efforts to reduce nonresponse but it is difficult to observe it. In practice one will rather observe a stepwise procedure as attempts to reduce the nonresponse rate are planned centrally and designed to take place uniformly.

Sometimes it is possible to distinguish between groups of respondents in a survey by the type of efforts and measures chosen or by the time needed to make them respond. When such groups can be distinguished, we will refer to them as **response waves of the sample**. In that way we will avoid confusion with the classification of the respondents by explanatory variables or with poststrata formed in order to improve the estimators by weighting. In many cases it is possible to identify the response waves by the number of **attempts** made to turn sampled units into respondents.

A well known procedure to increase the response rate is to draw a random subsample among the nonrespondents. If the data collection is well monitored, this can be a very useful device. There are only a few cases reported, however, where the follow-up efforts are focused on groups that are expected to contribute extensively to the nonresponse bias. Sometimes the efforts are instead directed towards groups which are thought to be easy to find and persuade to cooperate. If the nonresponse follow-up is not carefully monitored, an increase of the response rate might at worst make the distribution of respondents even more skew and increase the bias. Several examples of the importance of good interviewer training and monitoring are given by Lievesley (1988?).

1.2 PURE NONRESPONSE EFFECTS

As nonresponse reduces the sample size, there will be an increase of the sampling error even if there is nothing systematic in the distribution of nonrespondents. When the nonresponse is systematic its distribution has an influence on both the bias and the variance of the estimator.

It is not difficult to find empirical reasons to analyze the distribution of the respondents and consider its significance. Those who did not respond early may have thought that the survey did not concern them because their activity, interest or knowledge in the studied field was in some way restricted. Several studies have demonstrated that the response waves often have different averages in the survey variables, and that the proneness of a group to cooperate and its average are correlated. Such studies have mostly been done on mail surveys, where it is easy to classify the respondents by response waves, e.g. by those who answered immediately, those who answered after each mailed reminder, and those who answered only in a telephone interview.

A couple of examples will be enough to demonstrate that nonresponse follow-up is important for reducing the bias. All researchers on survey methods can provide similar examples from their own experience.

Table 1 presents a well known example of this type given by Cochran (1963). The respondents in a mail survey of fruit growers were broken down by response wave. The average number of fruit trees was calculated for each wave.

Table 1 AVERAGE BY RESPONSE WAVE

Response wave	Average number of fruit trees	Per cent of population
Response to first mailing	456	10
" second "	382	17
" third "	340	14
Nonrespondents after three mailings	290	59

K-E Kristiansson (1986) found in the Swedish *Survey on Living Conditions* that it was about 7 times as usual to receive social security benefits among the not-at-home's (22.9 %) as among the respondents (3.2%). There was no difference in this respect between the respondents and the other categories of nonrespondents. The benefit sum given to each social security receiver was about 50 per cent higher in the not-at-home group than in the rest of the sample. To allocate the resources to a bias-reducing nonresponse follow-up was obviously an important matter in that survey.

1.3 RESPONSE ERRORS IN DIFFERENT RESPONSE WAVES

Usually one has to be prepared to find response errors already in the answers of the most cooperative respondents. When the data collection continues among those who were not so easy to contact and those who needed more persuasion there are two other sources of response errors that can make their influence felt.

The first of them can appear even when the same data collection method is used in all the attempts. If the tentative nonrespondents are turned into reluctant respondents during the second or later attempt, there is a risk that their answers will be biased and that the response variance will increase. If worst comes to worst, extreme efforts to reduce the nonresponse rate may even increase the MSE by introducing a large bias. Very little is known about the size of this effect. Statistical practitioners are commonly aware of this, but solutions are rarely offered, nor the effects estimated. Groves (1989) presents this problem in more detail and Leslie insisted already in 1972 that high response rates are not always necessary for good estimates.

The other source of response error appears when the statistician after the opening attempt(s) tries to make the nonrespondents participate by changing the data collection method. There is always the risk that a different response behaviour will be introduced in the later response waves, if the same data collection method is not used as in the first. Several studies clearly demonstrate that the choice of measurement method has a substantial impact on the results, especially when it comes to sensitive information. The choice of data collection method, the presence or not of an interviewer, and the time available can influence how the respondents formulate their answers. Postal surveys tend to give higher percentages than interview surveys when individuals are asked about their drinking and smoking habits, diary surveys and interview surveys give different estimates of household expenditure, etc.

The two sources of response errors probably interact. They can both be included in the error models presented later in this report.

There is an abundance of studies that compare measurement methods in independent surveys. The occurrence of differences in estimated results due to the choice of measurement method is so well established, that it is enough to give one example. In a Swedish experiment, Bergman, Häll and Lindström (1980) compared telephone and at-home interviews. For many variables there were no differences between the two independent surveys, but some significant differences were also found. Some of them are reported in table 2. Similar examples are given by Groves (1989).

Table 2 FINDINGS BY DIFFERENT MEASUREMENT METHODS. PERCENTAGE DIFFERENCE

State or activity	Telephone interview	Personal interview	Difference %
* had a lasting illness	23	32	-9
* used some medicine	8	3	5
* did household work	66	59	7
* studied	21	13	8
(response rate	89	80)	

However, like other similar experiments this study does not entirely reflect the stated problem, as it compares the estimates of two independent samples. But it warns that something similar well might happen when different data collection methods are chosen for different response waves.

2 A MODEL FOR RESPONSE PROBABILITIES

2.1 DESIGN AND ESTIMATOR

In a population consisting of the units U_i, (i = 1,2 ... N) a simple random sample (srs) of n units is drawn with replacement (wr). The population average in the variable X is estimated with the **unweighted average among the n_r respondents:**

$$\bar{x}_r = \frac{\sum^{n_r} x_{ri}}{n_r} \qquad (2.1.1)$$

Another way to write this estimator which better reflects the sampling and response model is

$$\bar{x}_r = \frac{\sum^{n} R_i I_i x_i}{\sum R_i I_i} \qquad (2.1.2)$$

R_i and I_i are indicator variables. I_i = 1 if U_i is included in the sample, otherwise it is 0. R_i takes on the value 1 when U_i is included in the sample and responds, otherwise the value is 0. The probability to respond given U_i is denoted $P(R_i=1|U_i)$. **If there is no nonresponse or if all $P(R_i=1|U_i)$ are equal, the unweighted sample average (2.1.1) is an unbiased estimator of the population average.**

When all $P(R_i=1|U_i)$ are not equal, the unweighted sample average (2.1.1) may be biased. For obvious reasons, empirical studies of nonresponse bias are very scarce. Some Swedish studies reviewed and reported by Lindström (1983) give anyhow some idea of its size in different surveys. However, these findings pertain only to a restricted number of variables and domains of study.

If the different response probabilities are known, one can adjust the estimator and eliminate the nonresponse bias. According to Särndahl and Swensson (1985), an unbiased estimate of the population mean can be obtained by use of an expanded Horvitz-Thompson estimator. Conversion of additional nonrespondents into respondents is not needed to reduce the bias. Additional response waves of sampling units merely tend to reduce the standard error.

Håkan L Lindström
U/STM

2.2 MODELLING RESPONSE FUNCTIONS

As response probabilities rarely can be directly evaluated, it is important to use whatever possibilities there are, to find out something about the size of the nonresponse errors. By modelling and simulation we can arrive at a conception of the size of the errors we risk when we use a specific estimator and establish plausible values or at least limits for nonresponse errors in real surveys.

If the response probabilities in some way depend on X, and if $P(R_i=1|U_i)$ can be specified as a function of X - $\mathbf{P(R_i=1|X_i)}$ - then the response probabilities will be easier to model. It is still easier if the individual response probabilities can be approximated with a continuous response probability function $p(r|x)$.

The cases that demand special attention are those where the response probability function is monotone. Monotone and among them linear response probability functions can induce important nonresponse bias. As linear functions often approximate other monotone functions fairly well within an interval, one can get a good idea of when the distribution of the nonresponse threatens the accuracy of the estimator, by concentrating on linear functions in the rest of this study. They have also the advantage that they permit fairly simple analytic solutions for bias and variances.

There are several empirical examples on response frequency distributions to support the idea that there are not so few cases where the application of monotone response probability functions would be relevant. Two are presented below:

Table 3 shows some results from the 1968 Swedish Family Expenditure Survey. Household size was established for all the sampled units and response rates were calculated for the households by number of members. It is not very surprising to find that the response rates rise by household size. The more household members, the greater the possibility to find someone at home.

Table 3 RESPONSE RATE BY HOUSEHOLD SIZE

Household size	Response rate
1 member	0.64
2	0.71
3	0.74
4≤	0.84

In table 4 the nonresponse rates go up together with income class. The example originates from the American National Bureau of Economic Research and is reported by Dalenius (undated reprint of an early report). Compared to Swedish experiences, the difference in response rate between the extreme classes is very great.

Table 4 NONRESPONSE RATE BY INCOME

Income class $	Nonresponse %
0 - 500	1.0
500 - 1 000	1.1
1 000 - 1 500	4.5
1 500 - 2 000	6.3
2 000 - 2 500	8.2
2 500 - 3 000	9.5
3 000 - 5 000	12.3
5 000 - 10 000	17.0
10 000 and more	35.0

Nonmonotone response probability functions $p(r \mid x)$ should of course not be totally disregarded. Those which are symmetric around the centre of the interval on which X is defined tend to induce only minor nonresponse bias but may have an influence on the variance. Those which are independent of the values of the observations will only reduce the sample size and make it a random variable.

2.3 ONE WAVE OF RESPONDENTS - BIAS, VARIANCE AND MEAN SQUARE ERROR

To study the approximate behaviour of the estimator (2.1.1) in the presence of linear response functions, we make a number of assumptions:

* The distribution of the variable X can be approximated by a continuous frequency function, f(x),
* f(x) is defined on the finite and closed interval (0, K)
* The average is μ, the variance σ^2 and the third central moment is μ_3.
* The response probability depends on X and can be approximated with a continuous response function $p(r \mid x)$.
* Each unit is drawn and responds independently of all other units.

Lindström and Lundström (1974) have shown that when the response probability function is linear, i.e. when $p(r|x) = a + bx$, on a restricted interval (0, K), the expected response rate P is

$$P = \int p(r|x)f(x)dx = \mathbf{a} + \mathbf{b}\,\boldsymbol{\mu} \qquad (2.3.1)$$

and as soon as $P(n_r = 0)$ can be neglected, the expectation of the unweighted sample average for the respondents is

$$E(\bar{x}_r) = \mu + b\sigma^2/P = \mu + \mathbf{B} = \boldsymbol{\mu}'. \qquad (2.3.2)$$

All integrals in this report are evaluated on the interval (0, K) which is not pointed out in every single case. When $p(r|x) = a + bx$ is specified for a calculation one must check the values of a and b. In model (2.3.1), **a** can only take on a value between 0 and 1, and $\|b\|$ must not exceed 1/K. b is also restricted by the value that **a** and **P** takes on as $0 \le p(r|x) \le 1$ in the interval (0,K). The limits of **b** are:

$$\max\left(-\frac{1-P}{\mu}, \frac{-P}{K-\mu}\right) \le b \le \min\left(\frac{P}{\mu}, \frac{1-P}{K-\mu}\right) \qquad (2.3.3)$$

The calculations can be standardized for a couple of interesting distribution types if we define D as the difference in response probability between the extreme X-values and express **b** as

$$b = \frac{\mathbf{P(R|K) - P(R|0)}}{K} = \frac{\mathbf{D}}{K}; \qquad (2.3.4)$$

To terms of the order n^{-1}, the average variance is approximated by:

$$E_{n_r}\{Var(\bar{x}_r|n_r)\} \approx \frac{1}{Np}\left[\sigma^2 + \frac{b\mu_3}{P} - B^2\right] = \frac{1}{nP}\sigma_r^2; \qquad (2.3.5)$$

When the nonresponse bias $B(\bar{x}_r) = b\sigma^2/P$ **is denoted B** the mean square error of $\bar{x}_r$ is:

$$E_{n_r}\{MSE(\bar{x}_r|n_r)\} \approx (b\sigma^2/P)^2 + \frac{1}{nP}\left[\sigma^2 + \frac{b\mu_3}{P} - B^2\right] = B^2 + \frac{\sigma^2}{nP} + \frac{b\mu_3}{np^2} - \frac{B^2}{nP}$$

$$= B^2\left(1 - \frac{1}{nP}\right) + \frac{B}{nP}\left(\frac{P}{b} + \frac{\mu_3}{\sigma^2}\right); \qquad (2.3.6)$$

When b = 0 there is no bias and the group of respondents can be looked upon as a random subsample of the total sample. When $\|b\|$ increases but P is fixed, $\|B\|$ increases as well. So will $MSE[x_{nP|r}]$ do in most cases of practical importance. If b is not very small, B^2 will soon dominate (2.3.6) when n increases.

The aim of the 1974 study was to find out how much the level and distribution of nonresponse made $MSE[\bar{x}_{nP}]$ exceed σ^2/n - *the variance of the unbiased estimator when there is no nonresponse.* As formula (2.3.6) for the MSE is somewhat complex, numerical calculations under varied conditions were useful to illustrate its implications.

It is convenient to write

$Var(\bar{x}_{nP|r})$ for $E_{n_r}\{Var(\bar{x}_r|n_r)\}$ and $MSE(\bar{x}_{nP|r})$ for $E_{n_r}\{MSE(\bar{x}_r)|n_r)\}$

For the comparison of different realized designs, it will be useful to introduce a standardized measure. Here we use the square root of the ratio

$$(\mathbf{NRE}[\bar{x}_{nP|r}])^2 = \mathbf{MSE}[\bar{x}_{nP|r}]/\mathbf{Var}[\bar{x}] =$$

$$n\left(\frac{B}{\sigma}\right)^2 + \frac{1}{P}\left(1 + \frac{\sigma_r^2 - \sigma^2}{\sigma^2}\right) = n\alpha + \beta(1 + \Gamma) \; ; \qquad (2.3.7)$$

and term it the **NONRESPONSE EFFECT** (NRE) in analogy with the concept design effect. Partitioning $(NRE)^2$ as in (2.3.7) demonstrates three different effects of nonresponse:

α the nonresponse bias depending on the distribution of the nonresponse. $n\alpha$ will soon come to dominate NRE when n increases.

β the reduction of the sample size from n to nP

Γ the impact on the size of the unit variance

The NRE is quickly dominated by $n\alpha$ of formula (2.3.7) which mainly depends on n and b. Even small increases in $\|b\|$ will lead to an increase in NRE as well. The principal strategy to follow up nonresponse must then be to make b as close to zero as possible as also Γ of (2.3.7) is close to zero when b is.

2.4 SPECIAL DISTRIBUTIONS - ONE WAVE OF RESPONDENTS

When $\{NRE[\bar{x}_{nP|r}]\}^2 = MSE[\bar{x}_{nP|r}]/Var[\bar{x}]$ is calculated for the rectangular and triangular distribution types the results can be nicely condensed as **this ratio is independent of K.**

For the rectangular distribution:

$f(x) = 1/K$, $\mu = K/2$, $\sigma^2 = K^2/12$ and $\mu_3 = 0$;

Substitution in (2.3.7) and rearrangement of the formula gives

$$(NRE[\bar{x}_{nP|r}])^2 = n\left[\frac{D^2}{12P^2}\right] + \frac{1}{P}\left[1 - \frac{D^2}{12P^2}\right] \qquad (2.4.1)$$

For the triangular distribution:

$f(x) = 2(K-x)/K^2$, $\mu = K/3$, $\sigma^2 = K^2/18$ and $\mu_3 = K^3/135$;

the same operations result in :

$$(NRE[\bar{x}_{nP|r}])^2 = n\left[\frac{D^2}{15P^2}\right] + \frac{1}{P}\left[1 + \frac{D}{7P} - \frac{D^2}{15P^2}\right] \qquad (2.4.2)$$

For both the rectangular and the triangular distribution, the NRE turns out to be independent of both the position and the length of the interval on which X is defined. This makes the results useful for a larger class of distributions.

Different types of results have been calculated to demonstrate the effects of varying n and the parameters a and b of the linear response probability function. Calculations have only been made for reasonable values of D and P in order to restrict the number of table and diagrams. The average response rate, P, is kept in the interval 0.65 - 0.95. The differences in response rate between the extreme groups of a sample are allowed to differ with at most 30 per cent,

i.e. : $\| D \| = \| P(r|K)-P(r|0) \| \leq 0.30$

which might be a little more than necessary as differences larger than 20 per cent rarely appear in practical work.

The calculations presented in this report are:

* The nonresponse effect according to (2.4.1) and (2.4.2) and its components have been calculated for varying D and fixed values of n and P (n = 100 and P = 0.80). The results are presented in tables 5 and 6.

* The effect on NRE of letting n increase when P is fixed and D varies is shown in pictures 1 and 2 for n = 100, 400, 1000.

* The size of the bias when a and P varies but D is fixed (D = -0.20) is given in table 7.

* The sample size, n, where the 95 per cent confidence statement is violated when D varies and n and P are fixed (n = 100, P = 0.80) is given in table 8.

Table 5 NONRESPONSE EFFECT FOR THE TRIANGULAR DISTRIBUTION **WHEN n=100 and P = 0.80.**

D	$B/\sigma = \sqrt{\alpha}$	variance increase $\beta(1+\Gamma)$	$n\alpha$	NRE $\sqrt{n\alpha+\beta(1+\Gamma)}$
-0.30	- 0,10	1.17	0.94	1.45
-0.25	- 0,08	1.19	0.65	1.36
-0.20	- 0.06	1.20	0.42	1.27
-0.15	- 0.05	1.21	0.23	1.20
-0.10	- 0.03	1.23	0.10	1.15
-0.05	0.02	1.24	0.03	1.12
0.00	0	1.25	0	1.12
0.05	0.02	1.26	0.03	1.13
0.10	0.03	1.27	0.10	1.17
0.15	0.05	1.28	0.23	1.23
0.20	0.06	1.29	0.42	1.31
0.25	0.08	1.30	0.65	1.40
0.30	0.10	1.31	0.94	1.50

($\mu/K = 1/3$, $\beta = 1/P = 1.25$)

Table 5, as will table 6, shows that $n\alpha$, the nonresponse bias component, varies markedly even at a fixed response level, while the variance component $\beta(1+\Gamma)$ varies much less. Even for the small sample size n = 100, $n\alpha$ is almost as large as $\beta(1+\Gamma)$ for the extreme values of D in tables 5 and 6.

Table 6 NONRESPONSE EFFECT FOR THE RECTANGULAR DISTRIBUTION **WHEN n=100 and P = 0.80.**

D	$B/\sigma \approx \sqrt{\alpha}$	variance increase $\beta(1+\Gamma)$	$n\alpha$	NRE $\sqrt{n\alpha+\beta(1+\Gamma)}$
-0.30	- 0.11	1.24	1.17	1.55
-0.25	- 0.09	1.24	0.81	1.43
-0.20	- 0.07	1.24	0.52	1.33
-0.15	- 0.05	1.25	0.29	1.24
-0.10	- 0.04	1.25	0.13	1.17
-0.05	0.02	1.25	0.03	1.13
0.00	0	1.25	0	1.12
0.05	0.02	1.25	0.03	1.13
0.10	0.04	1.25	0.13	1.17
0.15	0.05	1.25	0.29	1.24
0.20	0.07	1.24	0.52	1.33
0.25	0.09	1.24	0.81	1.43
0.30	0.11	1.24	1.17	1.55

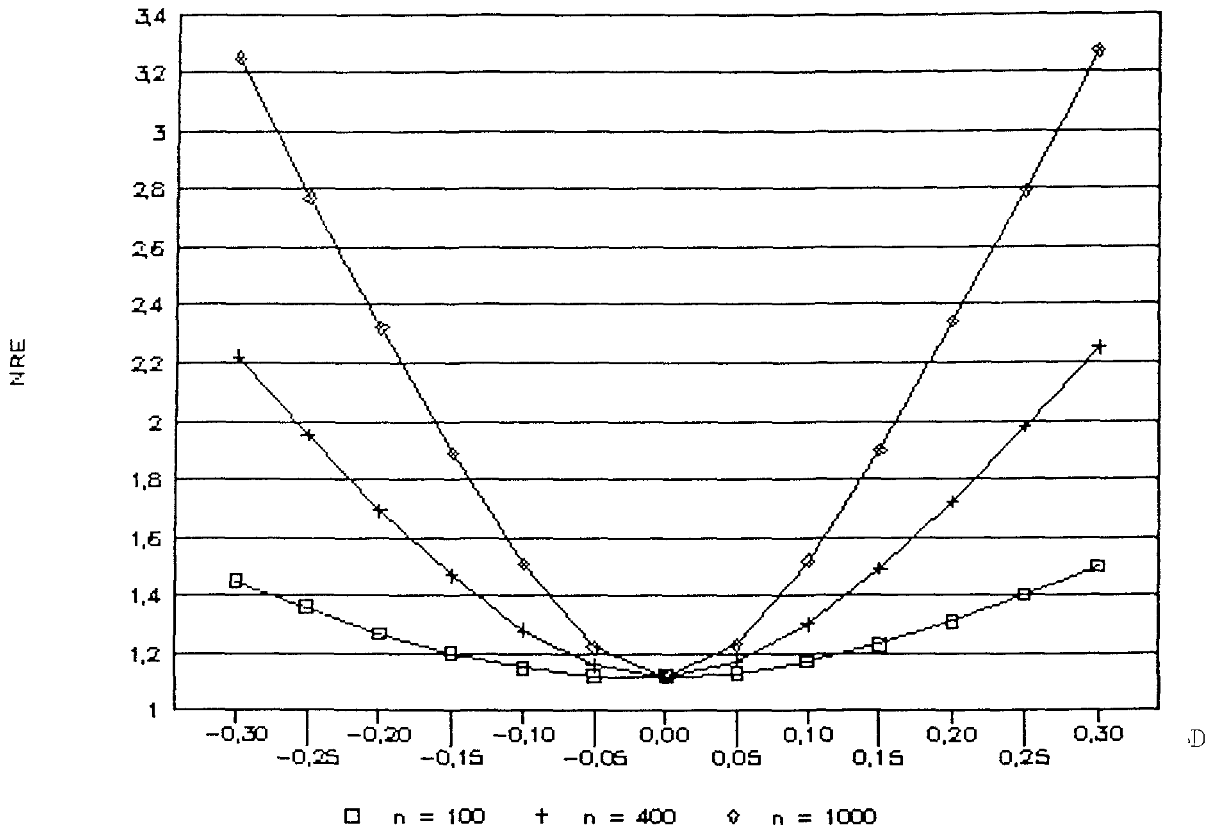
($\mu/K = 1/2$, $\beta = 1/P = 1.25$)

The results shown in tables 5 and 6 represent conditions that might appear in many fairly well-conducted studies. When response rates are lower, as assumed in the calculations of table 7 the NRE can be much larger. Given the slope, D/K, the bias component $n\alpha$ will increase in importance with decreasing P-level and decrease with increasing P-level.

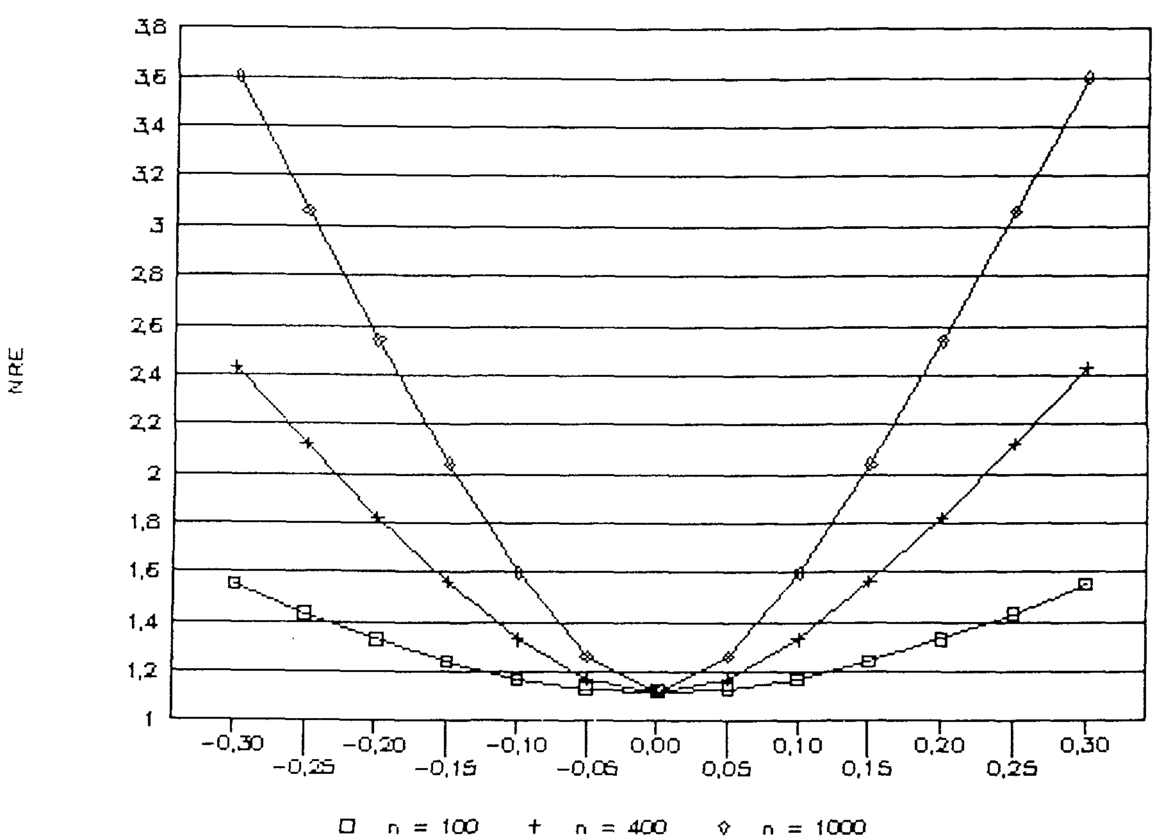
Compared to factor β, factor Γ had a very small impact on the size of the variance component of $(NRE)^2$ for all studied levels of D. The deviations of $\beta(1+\Gamma)$ from β varied between -0.08 and 0.06 for the triangular distribution and were almost negligable for the rectangular distribution.

It is obvious from the formulas that very soon $n\alpha$ - the squared nonresponse bias factor - will dominate NRE, when n increases and $\|D\|$ deviates from 0. How fast this happens is shown in pictures 1 and 2, where NRE has been calculated for the sample sizes 100, 400, and 1000 for both types of distributions at the same P level. All types of comparisons found in this report demonstrate the rectangular distribution to be more sensitive to the effects of linear response probability than the triangular. A distribution concentrated around the average will tend to be more robust than a more widely spread distribution.

Picture 1 NONRESPONSE EFFECT FOR DIFFERENT SAMPLE SIZES - TRIANGULAR DISTRIBUTION, P = 0.80



Picture 2 NONRESPONSE EFFECT FOR DIFFERENT SAMPLE SIZES - RECTANGULAR DISTRIBUTION, P = 0.80

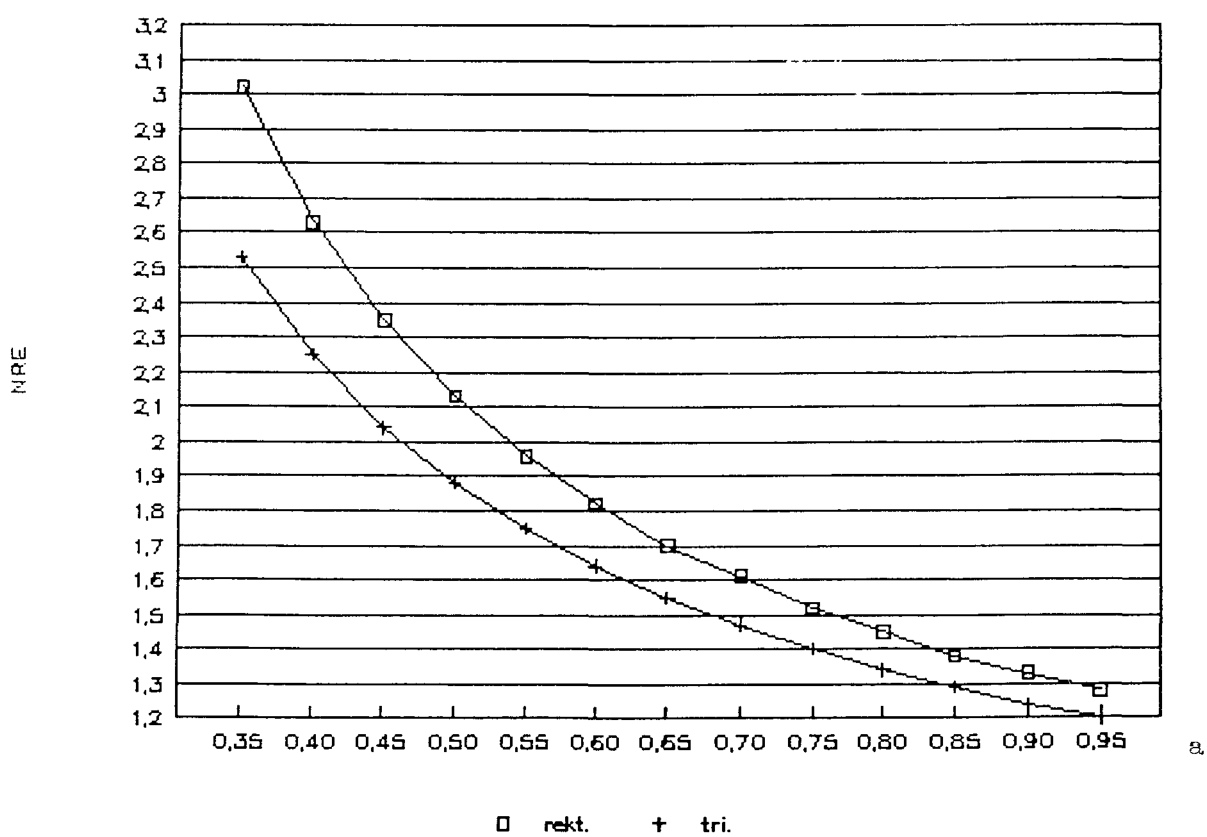


How fast NRE increases when **a** decreases is shown in table 7 and picture 3. It is also seen in the formulas (2.4.1-2). The lower the response rate, the more important is it that the respondent group is as similar to a simple random sample as possible in order to avoid large NRE.

Table 7 NRE WHEN **a** (and P) VARIES FOR CONSTANT D = - 0,2

	Triangular		Rectangular	
a	P	NRE	P	NRE
0.35	0.28	2.53	0,25	3.02
0.40	0.33	2.25	0.30	2.63
0.45	0.38	2.04	0.35	2.35
0.50	0.43	1.88	0.40	2.13
0.55	0.48	1.75	0.45	1.96
0.60	0.53	1.64	0.50	1.82
0.65	0.58	1.55	0.55	1.70
0.70	0.63	1.47	0.60	1.61
0.75	0.68	1.40	0.65	1.52
0.80	0.73	1.34	0.70	1.45
0.85	0.78	1.29	0.75	1.38
0.90	0.83	1.24	0.80	1.33
0.95	0.88	1.20	0.85	1.28

Picture 3 NRE WHEN **a** (and P) VARIES FOR CONSTANT D = - 0,2



There are several ways to express the consequences of systematic non-response bias. A common rule requires that the bias must not exceed 20 per cent of the standard error. If this condition is fulfilled, the confidence level is only slightly reduced - at most to 94.5 per cent. By application of formula (2.3.7) one can calculate the critical sample size that makes the relative size of a certain bias too large for the 95 per cent confidence level to be valid any more. The condition is:

$$n\left(\frac{B}{\sigma}\right)^2 \leq (0.20)^2 \left(\frac{1}{P}\ \frac{\sigma_r^2}{\sigma^2}\right). \qquad (2.4.3)$$

The values of nP for which condition (2.4.3) are violated are reported in table 8 for the distributions under study.

Table 8 CRITICAL SAMPLE SIZES nP WHEN $B^2 \geq 0.20\ \sigma_r^2/nP$ AND THE RESPONSE RATE IS P = 0.80.

D	Distribution type: Rectangular	Triangular
-0.30	4	5
-0.25	6	7
-0.20	10	12
-0.15	17	21
-0.10	38	47
-0.05	154	190
0.00	-	-
0.05	154	194
0.10	38	49
0.15	17	22
0.20	10	12
0.25	6	8
0.30	4	6

Table 8 shows that even a weak correlation between response probabilities and the variable values will reduce the intended 95 per cent confidence level for small sample sizes when confidence intervals are calculated without regard for the nonresponse bias.

The main interest of this study is focused on the NRE, but the non-response bias should not be completely overlooked. The relative nonresponse bias:

$RB[\bar{x}_r] = B[\bar{x}_r]/\mu$ is easily calculated. For $0 \leq X \leq K$ it is

* D/6P for the uniform distribution and
* D/5P for the triangular one.

In many surveys the response rate turns out to be $0.65 \leq P \leq 0.85$ or higher. Where observations on response rates were made, the relation $\|D\| \leq 0.20$ (or at most 0.30) was usually fulfilled. Then the ratio $\|D\|/P$ will take on values smaller than 1/2, and not even a very skew distribution of the non-response will make an estimator hopelessly biased. The relative nonresponse bias, RB, remains of limited size. For the triangular and rectangular distributions the limit for $\|RB\|$ will be approximately 0.1 under these conditions.

2.5 TWO OR MORE RESPONSE WAVES - RESPONSE PROBABILITIES

The response probabilities of response wave no j are denoted:

$$p_j(r|x) \qquad (2.5.1)$$

$p_1(r|x)$ is the probability to be in the first response wave, which consists of those who provide answers in the first data collection attempt, without any reminder. $p_2(r|x)$ is the probability to be in the second response wave, which implies an answer in the first extra effort, etc. j can take on the values 1,2,..r where r indicates the last response wave. There is also a probability $p_{j=nr}(r|x)$ to remain nonrespondent after all the attempts of the data collection agency.

P_j is the expected relative size of response wave j; (2.5.2)

$P_{j\leq r}$ is the expected relative size of the response waves 1 to r together. (2.5.3)

$W_j = P_j / P_{j\leq r}$ is the expected relative size of response wave j in the group of all respondents. (2.5.4)

Summing over respondents and non-respondents one gets:

$$P_{j\leq nr} = \sum_{1}^{nr} P_j = 1;$$

In the linear case the response probability function for response wave 1 is:

$$p_1(r|x) = a_1 + b_1x \; ; \qquad (2.5.5)$$

The increase in total response probability if attempt number j is made is

$$p_j(r|x) = a_j + b_jx - (a_{j-1} + b_{j-1}x) = (a_j - a_{j-1}) + (b_j - b_{j-1})x \qquad (2.5.6)$$

It will be easier to write

$$a_{dj} = (a_j - a_{j-1}) \text{ and similarly } b_{dj} = (b_j - b_{j-1}) \qquad (2.5.7)$$

It is necessary that $a_j \geq a_{j-1}$ and $a_j + b_jx \geq a_{j-1} + b_{j-1}x$ for $0 \leq X \leq K$.

The total probability to respond in wave r or before is

$$p_{j\leq r}(r|x) = a_r + b_rx \; ; \qquad (2.5.8)$$

2.6 TWO OR MORE RESPONSE WAVES - EXPECTATION AND BIAS

The principles followed to index the probabilities of separate and cumulated waves are also applied on averages.

The expected average of response wave j then proves to be :

$$E(\bar{x}_j) = \mu + \frac{\sigma^2(b_j - b_{j-1})}{P_j} = \mu + B_j = \mu_j'; \qquad (2.6.1)$$

The model permits averages in the response waves to be different. If the wave j (j = 1,2 ..r) of respondents is of the relative size P_j, the average of the observations up to wave j will be

$$\bar{x}_{j\leq r} = \Sigma[P_j\bar{x}_j]/P_{j\leq r} \; . \qquad (2.6.2)$$

From here on summation is from 1 to r, if not otherwise denoted. Combining (2.6.1) and (2.6.2) and taking expectation gives

$$E(\bar{x}_{j\leq r}) = \mu + \frac{\sigma^2 b_r}{P_{j\leq r}} = \mu + B_{j\leq r} = \mu_{j\leq r}' \; ; \qquad (2.6.3)$$

which of course is the same as if there was one response wave with the response probability function $a_r + b_rx$. If the final group of nonrespondents could be included as well when the average would be estimated without bias, $E[x_{j\leq nr}] = \mu$, because $a_{nr} = 1$ and $b_{nr} = 0$ and (2.6.4)

It is easy to calculate the conditions when an attempt to increase the response rate also reduces the bias. When the response rate goes up from P_1 to

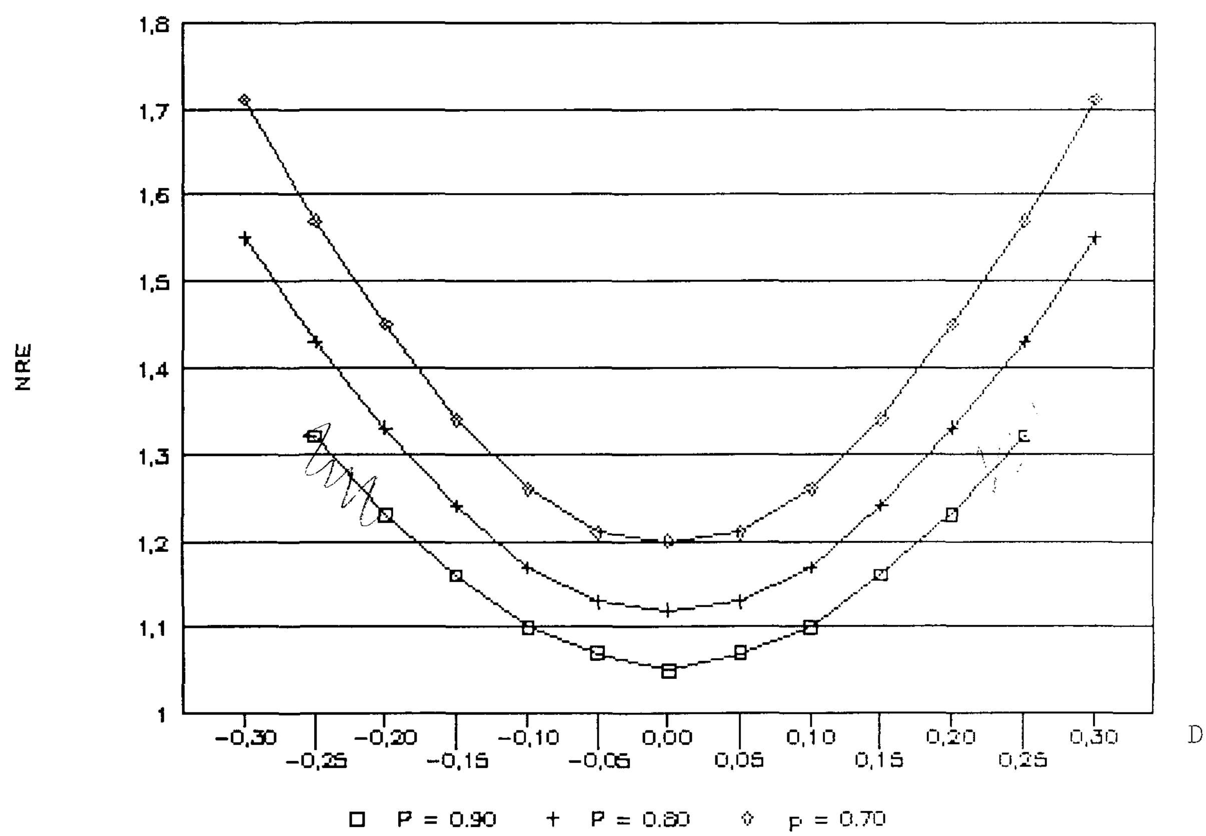
P_2 and at the same time the slope changes to b_2 from b_1, there is only a reduction of the nonresponse bias when

$$\|b_2\sigma^2/P_2\| \leq \|b_1\sigma^2/P_1\| \quad \text{or} \quad \|b_2/P_2\| \leq \|b_1/P_1\| \qquad (2.6.5)$$

Formula (2.6.5) says that the nonresponse bias is reduced only when the proportional increase of the response rate is larger then the proportional increase of the slope of the response function. For example, if $P_1 = 0.6$ and $b_1 = 0.10$, an increase of the response rate with 0.15 to $P_2 = 0.75$ will lead to an increased positive bias when $b_2 > 0.10*0.75/0.60 = 0.125$

Picture 4 gives the ratio b/P for $- 0.30 \leq b \leq 0.30$ with one curve for each of P = 0.70, 0.80 and 0,90. From the diagram we can see when an increase in response rate from P_1 to P_2 also reduces the nonresponse bias and when it does not. This happens in interval where the curve for a higher response rate is above the curve for a lower response rate.

Picture 4 NRE FOR THE RECTANGULAR DISTRIBUTION WHEN n = 100 and D VARIES



When P = 0.90 $\|a\|$ can not take on values higher than 0.20.

3 SIMULTANEOUS RESPONSE AND NONRESPONSE ERRORS.

The realized survey design that is modelled here is in outline:

There exists a "preferred" data collection procedure that is well implemented and in good control. For all practical purposes, the ideal would be to measure the variable X in this way for the whole sample, but this procedure can only be carried out among the sampled units that constitute the first response wave.

When the next data collection attempt is made, the attitude to the survey among those who will make up the second response wave may differ from that of those who constituted the first wave. One may also have to modify or or change the data collection method which also may generate measurement errors. The observed variable Y deviates from the ideal variable X systematically, randomly or in both ways. In particular, we might expect a systematic effect when the respondents for some reason or other are reluctant to answer. A proportional or random effect might be more likely if they are pressed for time or uninterested and prone to give quick but uncommitted and approximate answers.

The model disregards some complications in order to make the results more perspicuous. The model can certainly be elaborated in several ways with regard to sampling design, relaxation of the conditions of independence, follow up strategy, use of supplementary information in the estimation and nonresponse compensation. For instance, in many cases a random subsample of the nonrespondents from the first attempt is selected for the second attempt. This is frequently the case when the second attempt is more expensive than the first. We avoid such complications by including all the nonrespondents of the first attempt in the second and following attempt. An elaboration is meaningful and useful mainly in the context of a specific survey that can provide empirical information on the distribution of the errors.

4 A DETERMINISTIC MODEL

4.1 GENERAL ASSUMPTIONS

As a start, let us look at a simple but flexible model for the new variable Y_j of response wave j: $Y_j = \varepsilon_j + \delta_j X$; $j = 1,2, ..r$; (4.1.1)

where ε_j is a constant systematic error and
δ_j a proportional measurement effect
The observations are independent of each other

Given X and the response wave this model for the measurement process generates a deterministic response error. There are some situations where this is not a totally unrealistic approximation. For example, all respondents in wave j, might prefer to exaggerate their income in order to impress the interviewers. In that case $\varepsilon_j \geq 0$ and $\delta_j \geq 1$. If the question was asked by a tax collector one would not be surprised if both $\delta_j < 1$ and $\varepsilon_j \leq 0$.

This deterministic model will serve as a first approach to calculate analytic results and gain some understanding of the process. The relative simplicity of the model permits us to sort the different effects and to get at least some idea of their size and interaction before studying more flexible models by means of simulations.

4.2 EXPECTATION AND BIAS

As $\bar{x}_1 = \bar{y}_1$, the unweighted estimate of the observations

$$\bar{y}_{j\leq r} = \frac{\sum\sum y_{ji}}{\sum n_j} = \sum \frac{n_j}{\sum n_j} \bar{y}_j ; \qquad (4.2.1)$$

will have the approximate expectation

$$E[\bar{y}_{j\leq r}] \approx \frac{\sum \int (\varepsilon_j + \delta_j x) p(r_j | x) f(x) dx}{\sum \int p(r_j | x) f(x)\, dx} ; \qquad (4.2.2)$$

In the case of linear response probability functions the denominator of $E[\bar{y}_{j\leq r}]$

is $\sum \int (a_r + b_r x) f(x) dx = P_{j\leq r} = a_r + b_r \mu.$ (4.2.3)

and the numerator of is $\sum \int (\varepsilon_j + \delta_j x)(a_{dj} + b_{dj} x) f(x) dx$ (4.2.4)

so the expected average is approximately

$$E[\bar{y}_{j\leq r}] \approx \frac{\sum \int (\varepsilon_j + \delta_j x)(a_{dj} + b_{dj} x) f(x) dx}{a_r + b_r \mu} =$$

$$\frac{\sum [\varepsilon_j a_{dj} + (\varepsilon_j b_{dj} + \delta_j a_{dj})\mu + \delta_j b_{dj}(\mu^2 + \sigma^2)]}{a_r + b_r \mu} =$$

$$\frac{\sum (a_{dj} + b_{dj}\mu)[\varepsilon_j + \delta_j(\mu + B_j)]}{a_r + b_r \mu} = \sum W_j[\varepsilon_j + \delta_j(\mu + B_j)] =$$

$$\sum W_j[\varepsilon_j + \delta_j \mu'_j] = \sum W_j \mu''_j = \mu'' ; \qquad (4.2.5)$$

The bias of (4.2.1) is approximately

$$\text{Bias}(\bar{y}_{j\le r}) \approx \frac{\sum P_j[\varepsilon_j + \delta_j(\mu + B_j)]}{P_{j\le r}} - \mu =$$

$$\sum W_j[\varepsilon_j + (\delta_j-1)\mu + \delta_j B_j] = \sum W_j[\varepsilon_j + (\delta_j-1)(\mu + B_j) + B_j] =$$

$$\sum W_j\varepsilon_j + \sum W_j(\delta_j-1)(\mu_j') + B_{j\le r} = A + B + C\,. \qquad (4.2.6)$$

The rearrangement in formula (4.2.6) illustrates that the total bias under *the linear models is a sum of three additive components which in turn* express:

- A the weighted average of systematic errors - the constant response bias
- B the weighted proportional response bias
- C the nonresponse bias in all the waves together.

We can expect that in most cases $\| \mu \| \ge \|B_j\|$, and $(\delta_j-1)\|(\mu + B_j)\|$ can then be expected to have the same sign as (δ_j-1). A and C can both take on negative or positive values. The effects of the three components can obviously reinforce each other but also counteract or even eliminate each other.

4.3 THE VARIANCE AND MSE

The variance of the unweighted estimator (4.2.1) conditioned on the expected distribution of respondents $\{E(n_r)\} = \{nP_j\}$ is

$$\text{Var}[\bar{y}_{nP|r}] = \frac{\sum \int (\varepsilon_j + \delta_j x - \mu'')^2 p(r_j|x)f(x)dx}{nP_{j\le r} * \sum \int p(r_j|x)f(x)dx} =$$

$$\frac{1}{n(P_{j\le r})^2}\left[\sum \int (\varepsilon_j + \delta_j x - \mu_j'')^2 p(r_j|x)f(x)dx + \sum(\mu''_j - \mu'')^2 \int p(r_j|x)f(x)dx\right] =$$

$$[\sum W_j(\sigma_j^2 + (\mu''_j - \mu'')^2)]/nP_{j\le r} \qquad (4.3.1)$$

The unit variance, σ_j^2, in wave j in case of linear response and error functions is:

$$\sigma_j^2 = \int [\varepsilon_j + \delta_j x - (\varepsilon_j + \delta_j\mu + \delta_j B_j)]^2(a_{dj} + b_{dj}x)f(x)dx/P_j =$$

$$\delta_j^2 \int [x - (\mu + B_j)]^2(a_{dj} + b_{dj}x)f(x)dx/P_j =$$

$$\delta_j^2[\sigma^2 + \frac{b_{dj}\mu_3}{P_j} - B_j^2] \qquad (4.3.2)$$

Combining formulas (4.3.1) and (4.3.2) gives

$$Var[\bar{y}_{nP|r}] = \frac{1}{nP_{j\leq r}} \sum w_j \, [\delta_j^2(\sigma^2 + \frac{b_{dj}\mu_3}{P_j} - B_j^2) + (\mu_j'' - \mu'')^2] \qquad (4.3.3)$$

To make it easier to identify the effect of added measurement errors, we can also after some formula manipulation write the variance:

$$Var[\bar{y}_{nP|r}] = Var[\bar{x}_{nP|r}] + \frac{\sum w_j(\delta_j^2 - 1)\sigma_j^2}{nP_{j\leq r}} + \sum w_j[(\mu_j'' - \mu'')^2 - (\mu_j' - \mu')^2]/nP_{j\leq r} \qquad (4.3.4)$$

$Var[\bar{x}_{nP|r}]$ in (4.3.4) is the same variance as in (2.3.5) if the r response waves are regarded as one wave only.

The Mean Square Error is:

$$MSE[\bar{y}_{nP|r}] = \frac{1}{nP_{j\leq r}} \sum w_j[\delta_j^2(\sigma^2 + \frac{b_{dj}\mu_3}{P_j} - B_j^2) + (\mu''_j - \mu'')^2]$$

$$+ [\sum W_j[\varepsilon_j + (\delta_j - 1)\mu + \delta_j B_j]]^2 \, ; \qquad (4.3.5)$$

The square root of the ratio $MSE[\bar{y}_{j\leq r}|\{nP_j\}]/Var(\bar{x})$ will be referred to as the **Nonresponse and Measurement Error Effect (NRME)** in this report.

As we cannot be sure that the NRME will decrease when more sampling units become respondents, it is important to find out at what stage of the data collection procedure we reach the minimum of

$$MSE[\bar{y}_{nP|r}] = E[\bar{y}_{j\leq r} - \mu]^2 \qquad (4.3.6)$$

or at least if there is a response wave after which it is little use to make more attempts. There are also time lost and extra costs to consider.

4.4 TWO RESPONSE WAVES

In an attempt to make it easier to see how each of the error sources influences bias (4.2.6) and the MSE (4.3.5), one can look at the case with just two response waves and no response errors in wave 1. When $r = 2$, $b_{d1} = b_1$, $\varepsilon_1 = 0$ and $\delta_1 = 1$ the bias is by application of (4.2.6)

$$B(\bar{y}_{j\le2}|\{nP_j\}) \approx W_2\varepsilon_2 + W_2(\delta_2 - 1)(\mu'_2) + B_{j\le2} = A + B + C; \qquad (4.4.1)$$

A the weighted response bias in wave two
B the weighted proportional response variation in wave two
C the nonresponse bias in the two waves together.

In some cases the result is simplified. When $\delta_2=1$ or if $P_2/b_{d2} = \sigma^2/\mu$ the bias will be $W_2\varepsilon_2 + BP_{j\le2}$. When ε_2 and b_2 have the same sign the errors will cumulate, if not there is a tendency for the measurement bias to cancel the nonresponse bias. When $\delta_2=1$ and $\varepsilon_2 = 0$ (4.4.1) is reduced to the total nonresponse bias after the two response waves.

The between-variance component is in the case of two response waves

$$\sum w_j(\mu_j'' - \mu'')^2 = W_1W_2(\mu_1'' - \mu_2'')^2 = W_1W_2[\mu + B_1 - \varepsilon_2 - \delta_2(\mu + B_2)]^2 =$$

$$W_1W_2[B_1 - B_2 - \varepsilon_2 - (\delta_2 - 1)(\mu + B_2)]^2. \qquad (4.4.2)$$

The variance for the unweighted average of the observations is

$$Var[\bar{y}_{nP|2}] = \frac{1}{nP_{j\le2}} [\sigma^2(W_1 + W_2\delta_2^2) + \mu_3(b_1 + b_{d2}\delta_2^2)/P_{j\le2}$$

$$- (W_1B_1^2 + W_2\delta_2^2B_2^2) + W_1W_2[\mu(1 - \delta_2) - \varepsilon_2 + B_1 - \delta_2B_2]^2 =$$

$$Var[\bar{x}_{nP|2}] + \frac{W_2(\delta_2^2 - 1)\sigma_2^2 + W_1W_2[(\mu_1'' - \mu_2'')^2 - (B_1 - B_2)^2]}{nP_{j\le2}} \qquad (4.4.3)$$

By dividing the mean square error with the variance for a sample without nonresponse we get the squared nonresponse and measurement error effect.

$$[NRME(\bar{y}_{nP|2})]^2 =$$

$$[(W_1 + W_2\delta_2^2) + \frac{\mu_3}{P_{j\le2}}(b_1 + b_{d2}\delta_2^2) - (W_1B_1^2 + W_2\delta_2^2B_2^2]/P_{r\le2} +$$

$$W_1W_2[\mu(1-\delta_2) - \varepsilon_2+B_1 - \delta_2B_2]^2/P_{r\le2} + n[W_2\varepsilon_2 + (\delta_2-1)(W_2\mu-\sigma^2b_{d2}) + \sigma^2b_2]^2/\sigma^2$$
$$(4.4.4)$$

Not even in this very simplified case was it possible to rewrite the formula so far that the results can be easily expressed in tables and diagrams.

5 SIMULATION RESULTS

Although the model presented in chapter 4 gave some information on the effects of response and nonresponse errors, it is not sufficiently realistic for most applications. In applied studies, the modelling must be flexible enough to conform to the empirical results. Among other things, it must include the possibility of random response errors.

In order to study a few selected situations, some SAS simulations on PC have been made. The conditions were formulated to approach realistic and not unusual situations. The results should be regarded as a first step in an analysis of the sensitivity of the estimator (4.2.1) to both nonresponse and response errors.

The conditions for the simulations are:

1 There are two response waves, both with linear response functions.
2 The observations are mutually independent.
3 There are no response errors in the first response wave.
4 In the second response wave, $Y_2 = X + \varepsilon_2$. ε_2 is normally distributed with $\mu_\varepsilon =$ ~~(1+~~u_2~~)~~X and $\sigma_\varepsilon = v_2 X$. ε_2 expresses the measurement errors in the observations among the respondents during wave two. The errors may depend on X. However, Y_2 is not permitted to take values outside the interval (0, K). When $Y_2 < 0$, then 0 is substituted, and when $Y_2 > K$, then K is substituted.
5 The variable X has a rectangular distribution. (The triangular distribution tended to be less sensitive, at least to nonresponse errors.

$\mu_\varepsilon = u_2 X$

For all the simulations, the sample size is n = 100 . The number of replications is m = 1000 which seems large enough to give an understanding of the behaviour of the estimator (4.2.1). This number of repetitions may, however, not be large enough for precise information on a particular design.

In each replication, s = 1 ... 1000, of the sample the unweighted average (4.2.1) and NRME($\bar{y}_{j\le 2}$) was estimated together with the bias component of NRME($\bar{y}_{j\le 2}$) for the combined first and second response wave. The same estimation was made for the first response wave separately. These results were in good agreement of the theoretical results of section 2.3.

The estimators of the bias and of $((\text{NRME}_s(\bar{y}_{j\leq 2}))^2$ in replikation no s (s = 1. 2. ...1000) are:

$$\hat{B}_s(\bar{y}_{j\leq 2}) = \bar{y}_{j\leq 2,s} - \mu_x \quad (5.1) \qquad \text{and} \qquad (\bar{y}_{j\leq 2,s} - \mu_x)^2/\sigma_x^2/n \quad (5.2)$$

The estimators of the average over all the 1000 replications are

$$\hat{B}(\bar{y}_{j\leq 2}) = \frac{\sum_s \hat{B}_s(y_{j\leq 2})}{m} \quad \text{and} \qquad (5.3)$$

$$\widehat{\text{NRME}}(\bar{y}_{j\leq 2}) = \sqrt{\frac{\sum_s(\bar{y}_{j\leq 2,s} - \mu_x)^2/\sigma_x^2/n}{m}} \qquad (5.4)$$

Standard errors of the estimators (5.3) and (5.4) were calculated. They were very small and varied but little around 0.003.

In the simulations the sample size of each response wave is a random variable which contributes to the variance. NRME(y) includes the random variation in the size of the response waves and differs in that respect from the theoretical MSEs of sections 3 and 4 which were conditioned on the expected sample size.

The situations illustrated by the simulation are :

In the first response wave **P_1 = 0,65, b_1 = 0.10 and a_1 = 0.60.**

In the second attempt the response rate is increased to **P_2 = 0.85, so the second response wave constitutes 20 per cent of the entire sample**. This is done in one of three ways:

i. With reduced slope; **b_2 = 0.00, a_2 = 0.85;**
In this case the differences in response rates are succesfully evened out. The nonresponse bias disappears.

ii. With the same slope; **b_2 = 0.10, a_2 = 0.80;**
Here the increase is same within the entire range of X. The nonresponse bias is somewhat reduced. The precision of x is improved but not necessarily the precision of y.

iii. With increased slope; **b_2 = 0.20, a_2 = 0.75;**
The largest reduction in nonresponse rate comes in the groups that already are prone to respond. The nonresponse bias will increase. The precision of x is improved but not necessarily the precision of y.

The calculated bias and NRE of wave 1 and 2 when there are no response errors are given in table 9 as a basis for the evaluation of the added effect of response errors.

Table 9 CALCULATED NRE at n =100 FOR THE RECTANGULAR DISTRIBUTION WHEN P_2 = 0,85 and P_1 =0.65, $u_2 = 0$, $v_2 = 0$.

				$NRE(\bar{x}_{j\leq 2})$	$B(\bar{x}_{j\leq 2})/K$
a_1 =	0.60	b_2 =	0.10	1.32	0.01
a_2 =	0.85	b_2 =	0	1.08	0.00
	0.80		0.10	1.14	0.01
	0.75		0.20	1.28	0.02

Table 9 gives a new example on the importance to monitor the nonresponse follow up. After the second wave when the response rate P is 85 per cent NRE varies between 1.08 and 1.28. In the last case the NRE is only slightly below the NRE after the first wave when the response rate was only 65 per cent. The bias has even grown larger.

The respondents of the second wave are now thought to produce a systematic response error of the size u_2X and a random response error which is normally distributed around $X + u_2X$ and has the standard deviation v_2X. Table 10 gives some results for selected combinations of u_2 and v_2.

The values of u_2 and v_2 have been chosen somewhat arbitrarily but with the intention that they shall be of resonable size. One can easily go on like this and study the combination of errors one believes to be relevant in each specific application.

Response errors mainly explain the difference between the values in table 9 and those in table 10. The change in size of B/K and NRME depends on $p_2(r|X) = a_{d2} + b_{d2}x$ and the distribution of the response error. A minor part of the difference may derive from the variation in sample size among the replications.

Table 9 SIMULATED NRME at n = 100 FOR THE RECTANGULAR DISTRIBUTION WHEN P_2 = 0,85 and P_1 =0.65 (a_1 = 0.60, b_1 = 0.10).

		$NRME(\bar{y}_{j\leq 2})$	$B(\bar{y}_{j\leq 2})/K$
i. $u_2 = 0.0$	$v_2 = 0.2$:		
a_2 = 0.85	b_2 = 0	1.07	0.00
0.80	0.10	1.13	0.01
0.75	0.20	1.25	0.02
ii. $u_2 = 0.1$	$v_2 = 0.2$:		
a_2 = 0.85	b_2 = 0	1.27	0.02
0.80	0.10	1.52	0.03
0.75	0.20	1.75	0.04
iii. $u_2 = -0.1$	$v_2 = 0.2$:		
a_2 = 0.85	b_2 = 0	1.38	- 0.02
0.80	0.10	1.20	- 0.01
0.75	0.20	1.09	0.00
iv. $u_2 = -0.2$	$v_2 = 0.1$:		
a_2 = 0.85	b_2 = 0	1.81	- 0.04
0.80	0.10	1.63	- 0.03
0.75	0.20	1.38	- 0.02

The reported examples indicate cases where combinations of errors result in serious increase of bias and NRME as well as combinations that do not.

* The introduction of a only random response error ($u_2 = 0$, $v_2 = 0.2$) in the second response wave **(case i)** has no visible effect on either on bias or on NRME compared to the case of table 9 with only response errors.
* When a proportional systematic response error ($u_2 = 0.1$, $v_2 = 0.2$) is added **(case ii)** the relative bias is doubled and the NRME becomes so large that it exceeds the NRE of the first response wave for b_2= 0.1 and 0.2.
* In **case iii** ($u_2 = -0.1$, $v_2 = 0.2$) the proportional response bias is negative. Response and nonresponse errors counteract each other. When b_2= 0.2 there seems to be no bias at all but if b_2= 0 the NRME is larger than the NRE of the first wave.

* If the negative proportional systematic error is allowed to grow still larger as in **case iv** (u_2 = - 0.2, v_2 = 0.1), the bias returns and the NRME is very large and largest when $b_2 = 0$.

The presence of a proportional systematic response error seems to be much larger threat to the accuracy of the estimator than the presence of random response error.

As $\mu/K = 0.50$ for the rectangular distribution, the observed deviations $\|B/K\|/(\mu/K)$ never exceed 10 per cent. (It would of course be easy to make them do so by increasing the systematic part of the errors.) The observations on the limits of the size of the relative bias are however fully in line with the empirical observations reported by Lindström (1983).

6 A COST MODEL

Costs should be allocated in such a way that a minimum NRME is attained for given cost. A simple model for data collection and data processing in a survey is:

$$C = C_0 + n\{c_1 + (1 - P_1)c_2 + \quad (1 - P_r)c_{nr}\}$$

$$+ \; n\{c'_1P_1 + c'_2P_2 + ... \; + c'_rP_r\} \; ; \qquad (6.1)$$

where

- C_0 is a cost independent of the sample size.
- c_1 *is the per unit contact cost for all sampled units.*
- c_j is the added per unit contact cost in attempt j for those who are not in the pre-j response waves.
- c'_j is the data collection and data processing cost for a unit in response wave j.

The cost per response waves is more obvious and design easier to compare if the formula is rearranged to:

$$C - C_0 = n\{P_1(c_1 + c'_1) + P_2(c_1 + c_2 + c'_1) + ... + P_{nr}(c_1 + .. + c_r)\} \; ; \quad (6.2)$$

Consider a situation, where one can make a choice between

Design A: one attempt (estimator $\bar{x}_1$) with the preliminary sample size **n**,
Design B: two attempts (estimator $\bar{y}_{j\leq 2}$), with two response waves with the preliminary sample size **m**.

Design A and B are equally expensive when the relation between the sample sizes **m** and **n** is

$$n = m\left(1 + \frac{(1 - P_1)c_2 + P_2c'_2}{c_1 + P_1c'_1}\right) \; ; \qquad (6.3)$$

Usually the per unit cost for both contacting and data collection will increase with each attempt and in many cases n is larger than m. Having found out what sizes of m and n one can afford, one should then choose the design that has the smaller NRME. If also larger response errors threaten to appear in later response waves, it is not self-evident that it is a wise decision to make every possible attempt to reduce the non-response rate.

An application of this reasoning could be done on the last few responding units in many surveys. It is often very expensive to collect their answers. It will also delay the presentation of the results. If the observations are unreliable as well, there is no advantage at all in collecting them.

Håkan L Lindström
U/STM

7 REFERENCES

Bergman L.R., Häll L.and Lindström H.L. (1980); En studie för att uppskatta bortfallsfel i undersökningarna av levnadsförhållanden. (A study of nonresponse errors in the Surveys of Living Conditions), Report, October 1980, Statistics Sweden.

Cassel C.-M., Särndahl C.-E. and Wretman J.H. (1983); Some uses of statistical models in connection with the nonresponse problem, Incomplete data in sample surveys, Vol 3, 143-172, Academic Press.

Dahlenius T., Bortfallsproblemet vid statistiska undersökningar (The nonresponse problem in sample surveys), Undated reprint from Den svenska marknaden.

Groves R. M. (1989), Survey errors and survey costs, John Wiley and sons, New York.

Leslie L.L. (1972), Are high response rates essential to valid surveys?, Social Science Research 1, p 323-334.

Lievesley D. (1988), Unit Non-response in Interview Surveys. Report, Social and Comunity Planning Research.

Lindström H.L. and Lundström S. (1974), A method to discuss the magnitude of the nonresponse error, Statistisk Tidskrift 1974:6, pp 505-520, Statistics Sweden.

Lindström H.L. (1983), Non-Response Errors in Sample Surveys, Urval no 16, Statistics Sweden.

Särndahl C.-E. and Swensson B. (1985), A general view of estimation for two phases of selection. Promemorior från SCB no 20, Statistics Sweden.

R & D Reports är en för U/ADB och U/STM gemensam publikationsserie, som fr o m 1988-01-01 ersätter de tidigare "gula" och "gröna" serierna. I serien ingår även **Abstracts** (sammanfattning av metodrapporter från SCB).

R & D Reports Statistics Sweden are published by the Department of Research & Development within Statistics Sweden. Reports dealing with statistical methods have green (grön) covers. Reports dealing with EDP methods have yellow (gul) covers. In addition, abstracts are published three times a year (light brown /beige/ covers).

Reports published earlier during 1991:

1991:1 (grön)	Computing elementary aggregates in tee Swedish consumer price index **(Jörgen Dalén)**
1991:2 (grön)	Översikt av estimatorer för stora och små redovisningsgrupper **(Sixten Lundström)**

Kvarvarande **beige** och **gröna** exemplar av ovanstående promemorior kan rekvireras från Inga-Lill Pettersson, U/LEDN, SCB, 115 81 STOCKHOLM, eller per telefon 08-783 49 56.

Kvarvarande **gula** exemplar kan rekvireras från Ingvar Andersson, U/ADB, SCB, 115 81 STOCKHOLM, eller per telefon 08-783 41 47.