R&D Report 1998:5
Research – Methods – Development

Estimation from Order πps Samples with Non-Response

Bengt Rosén & Pär Lundqvist

SCB
Statistiska centralbyrån Statistics 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 1998:5. Estimation from order πps samples with non-response / Bengt Rosén; Pär Lundqvist.
Digitaliserad av Statistiska centralbyrån (SCB) 2016.

urn:nbn:se:scb-1998-X101OP9805

R&D Report 1998:5

Research – Methods – Development

Estimation from Order πps Samples with Non-Response

Bengt Rosén & Pär Lundqvist

R&D Report 1998:5
Research - Methods - Development

Estimation from Order πps Samples with Non-Response

Från trycket	Oktober 1998
Producent	Statistiska centralbyrån, *Statistics Sweden,* metodenheten Box 24300, *SE*-104 51 STOCKHOLM
Utgivare	Lars Lyberg

Förfrågningar	Bengt Rosén bengt.rosen@scb.se telefon 08-783 44 90 telefax 08-667 77 88
	Pär Lundqvist par.lundqvist@scb.se telefon 08-783 49 06 telefax 08-667 77 88

ISSN 0283-8680

Printed in Sweden
SCB-Tryck, Örebro 1998

October 1998

Estimation from Order πps Samples with Non-Response

Bengt Rosén & Pär Lundqvist

Abstract

An instrument for utilization of auxiliary information in sample surveys is πps sampling, i.e. to sample without replacement with inclusion probabilities proportional to given size measures. Rosén (1997b) introduced a novel class of πps schemes, called order πps schemes, and advised procedures for point and variance estimation. These procedures cover, however, only full response situations. Since some portion of non-response is rule rather than exception in practical surveys, it is desirable to have disposal of estimation procedures which work also under non-response. Such procedures is the topic of this paper.

Section 4, which is the core of the paper, presents the non-response adjusted estimation procedures, for point estimation and confidence intervals, which thereafter are given theoretical as well as simulation based justifications. The simulation results are presented in Section 5 and the theoretical justifications in Sections 6-8.

The tentative conclusion are as follows. Provided, of course, that response mechanism and response model comply with each other, the non-response adjusted estimation procedures work satisfactorily under quite general conditions. In particular, point estimator biases are negligible for all non-response rates already for quite small sample sizes. Good performance of the confidence intervals requires larger sample sizes, and depends on the type of relation between the size variable and the study variable.

CONTENTS

Estimation from Order πps Samples with Non - Response

1 Introduction and outline

An instrument for utilization of auxiliary information in sample surveys is πps sampling, i.e. to sample without replacement with inclusion probabilities proportional to given size measures. Rosén (1997b) introduced a novel class of πps schemes, called ***order πps schemes***, and advised procedures for point and variance estimation. The procedures cover, however, only full response situations. Since some portion of non - response is rule rather than exception in practical surveys, it is desirable to have disposal of estimation procedures which work also under non - response. Such procedures is the topic of this paper, which is organized as follows.

Section 2 briefly reviews some basic concepts, notably that of order πps sampling. Section 3 specifies the response models that are considered. Section 4, which is the core of the paper, presents procedures for non - response adjusted estimation. Thereafter the procedures are given theoretical as well as simulation justifications. In the belief that the latter are the most convincing they are presented first, in Section 5. The theoretical justifications are given in Sections 6, 7 and 8. The broad conclusion from the simulation findings is that the non - response adjusted estimation procedures work quite satisfactorily, provided that response mechanism and response model comply with each other.

The theoretical justifications are based on a limit result for an auxiliary notion, called ***quota random retain order sampling***, which is introduced in Section 6, where also the limit result is formulated. The proof is deferred to Section 8. Section 7 provides the bridge from quota random retain order sampling to order πps sampling with non - response. Since the estimation procedures are based on limit results they are approximate to some extent in "finite" situations. The simulation results in Section 5 provide information on approximation goodness.

Throughout the paper. P, E, V and $\hat{V}$ denote probability, expectation, variance and variance estimator. $N(\mu, \sigma^2)$ stands for the normal distribution with mean μ and variance σ^2.

2 Basic concepts, notably order πps sampling

$U = (1, 2, ..., N)$ denotes a finite population and $\mathbf{y} = (y_1, y_2, ..., y_N)$ a variable on it. The corresponding y - total over the population is;

$$\tau(\mathbf{y}) = \sum_{i=1}^{N} y_i . \tag{2.1}$$

A sampling frame in the form of a list with items that one - to - one correspond with the units in U is used. We let U refer to frame as well as population. For completeness we recall the Rosén (1997b) definition of order πps sampling. To that end we first present the more general notion "order sampling" from Rosén (1997a).

DEFINITION 2.1 Order sampling: To each unit i in a population $U = (1, 2, ..., N)$ is associated a probability distribution $F_i(t)$ with density $f_i(t)$, $0 \le t < \infty$. ***Order sampling*** with ***sample size*** $n, n \le N$, and ***order distributions*** $\mathbf{F} = (F_1, F_2, ..., F_N)$, denoted $OS(n; \mathbf{F})$, is conducted as follows. Independent ***ranking variables*** $Q_1, Q_2, ..., Q_N$ with distributions $F_1, F_2, ..., F_N$ are realized. The units with the ***n smallest*** Q - values constitute the sample.

Order πps schemes comprise a subclass of the OS(n; **F**) schemes, as follows.

DEFINITION 2.2 Order πps sampling: H(t) is a probability distribution with density h(t), $t \geq 0$, n is a positive integer and $\underline{\lambda} = (\lambda_1, .., \lambda_N)$ real numbers which satisfy;

$$0 < \lambda_i < 1, \quad i = 1, 2, \ldots, N, \qquad \sum_{i=1}^{N} \lambda_i = n. \tag{2.2}$$

Order πps sampling with ***sample size*** n, ***shape distribution*** H and ***target inclusion probabilities*** $\underline{\lambda}$ is OS(n; **F**) with order distributions that satisfy either, and hence both, of the following equivalent conditions (i) and (ii). H^{-1} denotes inverse function.

(ii) The ranking variables Q have the structure: $Q_i = H^{-1}(Z_i)/H^{-1}(\lambda_i)$, $i = 1, 2, \ldots, N$, where $Z_1, \ldots, Z_N$ are independent random variables uniformly distributed on [0, 1].

(i) The order distributions $\mathbf{F} = (F_1, F_2, \ldots, F_N)$ are;

$$F_i(t) = H(t \cdot H^{-1}(\lambda_i)) \quad \text{with density} \quad f_i(t) = H^{-1}(\lambda_i) \cdot h(t \cdot H^{-1}(\lambda_i)), \quad i = 1, 2, \ldots, N. \tag{2.3}$$

A general order πps scheme is referred to by the notation OSπps(n; H; $\underline{\lambda}$). Particular OSπps schemes are named by their shape distributions.

In spite of the fact that OSπps schemes are very simple to implement, their definition admittedly looks intricate when first met. The chief reason why OSπps schemes are of practical interest is the following result, which is discussed at length in Rosén (1998).

APPROXIMATE INCLUSION PROBABILITIES FOR OSπps: For OSπps(n; H; $\underline{\lambda}$) holds with good approximation, whatever the shape distribution H is;

$$\text{The sample inclusion probability for unit } i \approx \lambda_i, \quad i = 1, 2, \ldots, N. \tag{2.4}$$

Result (2.4) implies that OSπps is a means for sampling with inclusion probabilities (approximately) equal to given target inclusion probabilities $\underline{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_N)$. The λ:s are typically generated as follows. The sampling frame comprises a ***size variable*** $\mathbf{s} = (s_1, s_2, \ldots, s_N)$. The corresponding ***normalized size variable*** is obtained by scaling the **s** - values so that their sum becomes 1;

$$\mathbf{s}^* = (s_1^*, s_2^*, \ldots, s_N^*) = (s_1, s_2, \ldots, s_N) / \sum_{i=1}^{N} s_i. \tag{2.5}$$

Then, for a given sample size n, the λ:s are defined as;

$$\lambda_i = n \cdot s^*, \quad i = 1, 2, \ldots, N. \tag{2.6}$$

In (2.5) it may happen that one or more λ exceeds 1. If so, some special measure has to be taken; Introduction of a "take all" stratum, or something else. We presume that $\lambda_i < 1$.

In line with Rosén (1997a & b) we will pay special attention to the following OSπps schemes.

Uniform OSπps: $H(t) = t$, $h(t) = 1$, $0 \leq t \leq 1$, $H(t) = 1$, $h(t) = 0$, $1 \leq t \leq \infty$,

$$\text{with inverse } H^{-1}(\lambda_i) = \lambda_i, \quad i = 1, 2, \ldots, N. \tag{2.7}$$

Exponential OSπps: $H(t) = 1 - e^{-t}$, $h(t) = e^{-t}$, $0 \leq t < \infty$,

$$\text{with inverse } H^{-1}(\lambda_i) = -\log(1 - \lambda_i), \quad i = 1, 2, \ldots, N. \tag{2.8}$$

Pareto OSπps: $H(t) = t/(1 + t)$, $h(t) = 1/(1 + t)^2$, $0 \leq t < \infty$,

$$\text{with inverse } H^{-1}(\lambda_i) = \lambda_i / (1 - \lambda_i), \quad i = 1, 2, \ldots, N. \tag{2.9}$$

3 Response modeling

In situations where non - response occurs during data collection a prerequisite for unbiased inference is that the inference is based on a good model of the factual ***response mechanism***. We confine to the following simple and well-known response mechanism model.

RESPONSE MECHANISM 3.1 : To each unit i in U is associated a random 0-1-variable R_i, called ***response indicator***, which specifies if unit i responds ($R_i = 1$) or not ($R_i = 0$) if addressed. The corresponding ***response propensities*** are;

$$\beta_i = P(R_i = 1), \quad i = 1, 2, \ldots, N. \tag{3.1}$$

The following assumptions (i) and (ii) are made.

(i) $R_1, R_2, \ldots, R_N$ are mutually independent. (3.2)

In a sample survey context with sample inclusion indicators $I_1, I_2, \ldots, I_N$, (i.e. $I_i = 1$ if unit i is sampled and $I_i = 0$ otherwise);

(ii) The stochastic vectors $(I_1, I_2, \ldots, I_N)$ and $(R_1, R_2, \ldots, R_N)$ are independent. (3.3)

We say that ***over-all uniform response propensities*** are at hand if;

$$\beta_1 = \beta_2 = \ldots = \beta_N = \beta. \tag{3.4}$$

We say that ***group-wise uniform response propensities*** are at hand if U partitions into disjoint ***response homogeneity groups*** $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_G$ with response propensities $\beta^{(g)}$, and;

$$\beta_i = \beta^{(g)}, \quad i \in \mathcal{G}_g, \quad g = 1, 2, \ldots, G. \tag{3.5}$$

The case with over-all uniform response propensities can of course be seen as a special case of group-wise uniform response propensities, with the entire population as the sole response homogeneity group.

When non-response occurs, assumptions about the response mechanism must be introduced in the estimation step. Such assumptions are called the ***response model***. We confine to response models which comply with the above response mechanism model, to the effect that the response model specifies a set $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_G$ of surmised response homogeneity groups. Note that specification of values for response propensities is part of the response mechanism model, but not of the response model. It goes without saying that the success of non - response adjustment depends on the agreement between response mechanism and response model.

4 Non - response adjusted estimation for OSπps samples

In this section we formulate results about non-response adjusted estimation for OSπps. Theoretical justification of the results is deferred to Sections 6, 7 and 8, while the next Section 5 provides simulation based justification. We provide background for the adjusted procedures by stating estimation formulas from Rosén (1997b) for the full response case.

4.1 Estimation under full response

ESTIMATION PROCEDURE 4.1 : An OSπps sample with sample size n, shape distribution H with density h and target inclusion probabilities $\underline{\lambda} = (\lambda_1, \lambda_2.., \lambda_N)$ is drawn from $U = (1, 2, ..., N)$. The values of a variable $\mathbf{y} = (y_1, y_2, ..., y_N)$ are observed for all sampled units. Then, the following holds under general conditions.

a) Consistent estimation of the total $\tau(\mathbf{y})$ is given by;

$$\hat{\tau}(\mathbf{y})_{OS} = \sum_{i \in \text{Sample}} y_i / \lambda_i \,. \tag{4.1}$$

b) A confidence interval for $\tau(\mathbf{y})$ with approximate confidence level $1-\alpha$ is given by;

$$\hat{\tau}(\mathbf{y})_{OS} \pm \delta_{\alpha/2} \cdot \sqrt{\hat{V}[\hat{\tau}(\mathbf{y})_{OS}]}\ , \tag{4.2}$$

where $\delta_{\alpha/2}$ is the $(1-\alpha/2)$-fractile in the standard normal distribution, and;

$$\hat{V}[\hat{\tau}(\mathbf{y})_{OS}] = \frac{n}{n-1} \cdot \sum_{i \in \text{Sample}} \left(\frac{y_i}{\lambda_i} - \sum_{j \in \text{Sample}} \frac{y_j \cdot a_j}{\lambda_j} \Big/ \sum_{j \in \text{Sample}} a_j \right)^2 \cdot (1-\lambda_i), \tag{4.3}$$

with

$$a_i = h(H^{-1}(\lambda_i)) \cdot H^{-1}(\lambda_i)/\lambda_i, \quad i = 1,2,\ldots,N. \tag{4.4}$$

Also, $\hat{V}[\hat{\tau}(\mathbf{y})_{OS}]$ yields consistent estimation of the estimator variance $V[\hat{\tau}(\mathbf{y})_{OS}]$.

The a_i-values for the schemes (2.7)-(2.9) are stated in Lemmas 3.1 and 4.1 in Rosén (1997b). They are listed below.

For ***uniform*** OSπps: $a_i = 1, \quad i = 1,2,\ldots,N.$ (4.5)

For ***exponential*** OSπps: $a_i = -\ln(1-\lambda_i) \cdot (1-\lambda_i)/\lambda_i, \quad i = 1,2,\ldots,N.$ (4.6)

For ***Pareto*** OSπps: $a_i = 1-\lambda_i, \quad i = 1,2,\ldots,N.$ (4.7)

4.2 Estimation when non-response is modeled by over-all uniform response propensities

The non-response adjusted estimation procedure can broadly be described as follows. Modify "sampled" to "responding" and make the corresponding modifications of the target inclusion probabilities. Then use the full response estimation formulas (4.1) and (4.3). Under the response model with over-all uniform response propensities, the modifications are as follows.

Sample is exchanged for "responding sampled units". (4.8)

Sample size is exchanged for "number of responding units". (4.9)

λ_i is exchanged for λ'_i given by (2.6) with n = "number of responding units". (4.10)

The precise result is stated below.

ESTIMATION PROCEDURE 4.2: An OSπps sample with size n, shape distribution H with density h and target inclusion probabilities $\underline{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_N)$ is selected from the population $U = (1,2,\ldots,N)$. The a_i are according to (4.4). Non-response occurs during data collection, and the response model presumes over-all uniform response propensities. Set;

$\mathfrak{R}$sample = the set of responding sampled units, (4.11)

n' = number of responding units, (4.12)

$$\lambda'_i = (n'/n) \cdot \lambda_i = n' \cdot s_i^*, \quad i = 1,2,\ldots,N. \tag{4.13}$$

Then the following holds under general conditions.

a. Consistent estimation of the total $\tau(\mathbf{y})$ is given by;

$$\hat{\tau}(\mathbf{y})_{OSA} = \sum_{i \in \mathfrak{R}\text{sample}} y_i/\lambda'_i . \tag{4.14}$$

b. A confidence interval for $\tau(\mathbf{y})$ with approximate confidence level $1-\alpha$ is given by;

$$\hat{\tau}(\mathbf{y})_{OSA} \pm \delta_{\alpha/2} \cdot \sqrt{\hat{V}[\hat{\tau}(\mathbf{y})_{OSA}]}\ , \tag{4.15}$$

where $\delta_{\alpha/2}$ is the $(1-\alpha/2)$-fractile in the standard normal distribution, and;

$$\hat{V}[\hat{\tau}(y)_{OSA}] = \frac{n'}{n'-1} \cdot \sum_{i\in\Re\text{sample}} \left(\frac{y_i}{\lambda'_i} - \sum_{j\in\Re\text{sample}} \frac{y_j \cdot a_j}{\lambda'_j} \Big/ \sum_{j\in\Re\text{sample}} a_j \right)^2 \cdot (1-\lambda'_i) \,. \tag{4.16}$$

Remark 4.1 : Computation of the variance estimator (4.16) is simplified by expansion of the square. After some straightforward algebra this yields;

$$\hat{V}[\hat{\tau}(y)_{OSA}] = \frac{n'}{n'-1} \cdot (A - 2 \cdot B \cdot D/K + C \cdot D^2/K^2)\,, \quad \text{where} \tag{4.17}$$

$$A = \sum_{i\in\Re\text{sample}} (y_i/\lambda'_i)^2 \cdot (1-\lambda'_i)\,, \quad B = \sum_{i\in\Re\text{sample}} (y_i/\lambda'_i) \cdot (1-\lambda'_i)\,, \quad C = \sum_{i\in\Re\text{sample}} (1-\lambda'_i)\,, \tag{4.18}$$

$$D = \sum_{i\in\Re\text{sample}} (y_i/\lambda'_i) \cdot a_i\,, \quad K = \sum_{i\in\Re\text{sample}} a_i\,. \quad ¤ \tag{4.19}$$

Remark 4.2 : The variance estimator (4.16) emanates from (6.16) and (6.17), which are based on Theorem 8.1. In (4.16) the target inclusion probabilities appear in adjusted versions λ' while the a_i are based on the original λ : s. A natural wonder is what happens if also adjusted a_i-values are used, leading to the following modified estimator;

$$\hat{V}'[\hat{\tau}(y)_{OSA}] = \frac{n'}{n'-1} \cdot \sum_{i\in\Re\text{sample}} \left(\frac{y_i}{\lambda'_i} - \sum_{j\in\Re\text{sample}} \frac{y_j \cdot a'_j}{\lambda'_j} \Big/ \sum_{j\in\Re\text{sample}} a'_j \right)^2 \cdot (1-\lambda'_i)\,, \tag{4.20}$$

where a'_i is given by the following modification of (4.4);

$$a'_i = h(H^{-1}(\lambda'_i)) \cdot H^{-1}(\lambda'_i)/\lambda'_i\,, \quad i = 1,2,\ldots,N. \tag{4.21}$$

For the schemes (2.7) - (2.9), (4.21) becomes as follows.

For ***uniform*** OSπps: $a'_i = 1, \quad i = 1,2,\ldots,N.$ (4.22)

For ***exponential*** OSπps: $a'_i = -\ln(1-\lambda'_i) \cdot (1-\lambda'_i)/\lambda'_i$, if $\lambda'_i < 1$,
$= 0$ if $\lambda'_i = 1$, $i = 1,2,\ldots,N.$ (4.23)

For ***Pareto*** OSπps: $a'_i = 1-\lambda'_i, \quad i = 1,2,\ldots,N.$ (4.24)

To distinguish the two variance estimators, the (4.16) estimator is said to be of ***Type 1*** and the (4.20) estimator of ***Type 2***. ¤

Remarks 4.3 : (i) Under the present response model, over - all uniform response propensities, $\lambda'_i < 1$ always holds. The case $\lambda'_i = 1$ in (4.23) is introduced for later use.

(ii) Note that $\hat{V}' = \hat{V}$ for uniform OSπps, since then $a'_i = a_i$.

(iii) As regards computation of $\hat{V}'$, the procedure in Remark 4.1 can be used after the following modifications of D and K;

$$D = \sum_{i\in\Re\text{sample}} (y_i/\lambda'_i) \cdot a'_i\,, \quad K = \sum_{i\in\Re\text{sample}} a'_i\,. \quad ¤ \tag{4.25}$$

4.3 Estimation when non-response is modeled by group-wise uniform response propensities

Also when the response model comprises more than one response homogeneity group the adjusted estimation procedure follows the previous lines. First n and λ_i are modified, here group-wise, then the full response Procedure 4.1 is applied. The precise formulation is given below.

ESTIMATION PROCEDURE 4.3: An OSπps(n;H;$\underline{\lambda}$) sample is drawn from U=(1,...,N). The a_i are according to (4.4). Non-response occurs during data collection, and is modeled by group-wise uniform response propensities with homogeneity groups $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_G$. Set;

$$n_g = \text{number of sampled units from group } \mathcal{G}_g,\ g=1,2,\ldots,G, \tag{4.26}$$

$$n'_g = \text{number of responding units from group } \mathcal{G}_g,\ g=1,2,\ldots,G, \tag{4.27}$$

$$\mathfrak{R}_g\text{sample} = \text{the set of responding sampled units from group } \mathcal{G}_g,\ g=1,2,\ldots,G, \tag{4.28}$$

$$\mathfrak{R}\text{sample} = \mathfrak{R}_1\text{sample} \cup \mathfrak{R}_2\text{sample} \cup \ldots \cup \mathfrak{R}_G\text{sample}, \tag{4.29}$$

$$\lambda'_i = \min(n'_{g(i)} \cdot \lambda_i \Big/ \sum_{j \in \mathcal{G}_{g(i)}} \lambda_j,\ 1) = \min(n'_{g(i)} \cdot s_i \Big/ \sum_{j \in \mathcal{G}_{g(i)}} s_j,\ 1),\quad i=1,2,\ldots,N, \tag{4.30}$$

where $g_{(i)}$ is the index for the $\mathcal{G}$-group to which unit i belongs.

Then the following holds under general conditions.

a. Consistent estimation of the total $\tau(\mathbf{y})$ is given by;

$$\hat{\tau}(\mathbf{y})_{OSGA} = \sum_{i \in \mathfrak{R}\text{sample}} y_i / \lambda'_i . \tag{4.31}$$

b. A confidence interval for $\tau(\mathbf{y})$ with approximate confidence level 1-α is given by;

$$\hat{\tau}(\mathbf{y})_{OSGA} \pm \delta_{\alpha/2} \cdot \sqrt{\hat{V}[\hat{\tau}(\mathbf{y})_{OSGA}]}, \tag{4.32}$$

where $\delta_{\alpha/2}$ is the (1 - α/2)-fractile in the standard normal distribution, and;

$$\hat{V}[\hat{\tau}(\mathbf{y})_{OSGA}] = \sum_{g=1}^{G} \frac{n'_g}{n'_g - 1} \cdot \sum_{i \in \mathfrak{R}_g\text{sample}} \left(\frac{y_i}{\lambda'_i} - \sum_{j \in \mathfrak{R}_g\text{sample}} \frac{y_j \cdot a_j}{\lambda'_j} \Big/ \sum_{j \in \mathfrak{R}_g\text{sample}} a_j \right)^2 \cdot (1 - \lambda'_i). \tag{4.33}$$

Remark 4.4: The λ-modification in (4.30) can be concretized as follows. First the size values are normalized on each $\mathcal{G}$-group. Thereafter the λ':s are computed by group-wise application of the modified version of (2.6). Note that this operation requires "frame knowledge" of $\mathcal{G}$-belonging. The min(·, 1) operation in (4.30) is introduced for the following reason. Without this, admittedly somewhat ad hoc, rule it can happen, although only in exceptional cases, that one or more λ'_i becomes greater than 1. This operation also explains formula (4.23). ¤

***Remark 4.5*:** In analogy with what is said in Remark 4.1, computation of the variance estimator (4.33) is simplified by expansion of the squares. This leads to the formula;

$$\hat{V}[\hat{\tau}(\mathbf{y})_{OSGA}] = \sum_{g=1}^{G} \frac{n'_g}{n'_g - 1} (A_g - 2 \cdot B_g \cdot D_g / K_g + C_g \cdot D_g^2 / K_g^2), \tag{4.34}$$

where $A_g, B_g, \ldots, K_g$ are the $\mathfrak{R}_g$sample analogues of the quantities in (4.18) and (4.19). ¤

***Remark 4.6*:** In Remark 4.2 we introduced a modified version of the estimator (4.16). The corresponding modification of the variance estimator (4.33) is;

$$\hat{V}[\hat{\tau}(\mathbf{y})_{OSGA}] = \sum_{g=1}^{G} \frac{n'_g}{n'_g - 1} \cdot \sum_{i \in \mathfrak{R}_g\text{sample}} \left(\frac{y_i}{\lambda'_i} - \sum_{j \in \mathfrak{R}_g\text{sample}} \frac{y_j \cdot a'_j}{\lambda'_j} \Big/ \sum_{j \in \mathfrak{R}_g\text{sample}} a'_j \right)^2 \cdot (1 - \lambda'_i), \tag{4.35}$$

where a' is defined by (4.21). As in Remark 4.2, we distinguish the two variance estimators by calling (4.33) the ***Type 1*** estimator and (4.35) the ***Type 2*** estimator. ¤

4.4 Approximate formulas for estimator variances

The theme in this section is theoretical estimator variances. Even if such variances are unimportant in the estimation phase, they are of interest in the survey planning phase and also in a method evaluation of the present type. As stated in Procedure 4.1, in the full response case the variance estimator (4.3) provides consistent estimation of the corresponding theoretical variance. An asymptotically correct approximation formula from Rosén (1997b) is repeated below. The novel part concerns analogous asymptotically correct approximation formulas for theoretical estimator variance under non-response. There is a difference, though, between the cases without and with non-response, which is discussed first.

4.4.1 The non-response adjusted estimation procedures are conditional inference procedures

When non - response occurs, the number of responding units is random. As a consequence Estimation Procedures 4.2 and 4.3 are, which is discussed in more detail in Section 7, conditional inference procedures, with conditioning on the number of responding units in the different response homogeneity groups. In particular, (4.16) and (4.33) are consistent estimators of the corresponding conditional estimator variances. Derivation of exact formulas for these conditional variances is unfeasible, but asymptotically correct approximate formulas are given. Justifications are given in Sections 6 and 7.

4.4.2 Results on approximate conditional estimator variances

As before we start by giving background by stating results for the full response case, fetched from Approximation Result 3.1 in Rosén (1997b)

APPROXIMATE VARIANCE FOR THE ESTIMATOR (4.1): Let assumptions and notation be as in Procedure 4.1. Then, under general conditions, (4.36) provides an asymptotically correct approximation of the variance of $\hat{\tau}(\mathbf{y})_{OS}$;

$$V[\hat{\tau}(\mathbf{y})_{OS}] \approx \frac{N}{N-1} \cdot \sum_{i=1}^{N} \left(\frac{y_i}{\lambda_i} - \sum_{j=1}^{N} \frac{y_j \cdot \alpha_j}{\lambda_j} \Big/ \sum_{j=1}^{N} \alpha_j \right)^2 \cdot \lambda_i \cdot (1-\lambda_i), \quad \text{where} \tag{4.36}$$

$$\alpha_i = h(H^{-1}(\lambda_i)) \cdot H^{-1}(\lambda_i), \quad i = 1,2,\ldots,N. \tag{4.37}$$

The α_i-values for the schemes in (2.7)-(2.9), which are given in Lemmas 3.1 and 4.1 in Rosén (1997b), are as stated below.

For ***uniform*** OSπps: $\alpha_i = \lambda_i, \quad i = 1,2,\ldots,N.$ (4.38)

For ***exponential*** OSπps: $\alpha_i = -\ln(1-\lambda_i) \cdot (1-\lambda_i), \quad i = 1,2,\ldots,N.$ (4.39)

For ***Pareto*** OSπps: $\alpha_i = \lambda_i \cdot (1-\lambda_i), \quad i = 1,2,\ldots,N.$ (4.40)

We now turn to theoretical conditional variances corresponding to the adjusted estimators. As stated earlier, conditioning is made on the number of responding units, n' in the over-all case and $(n'_1, n'_2, \ldots, n'_G)$ in the group case.

APPROXIMATE CONDITIONAL VARIANCE FOR THE ESTIMATOR (4.14): Let assumptions and notation be as in Procedure 4.2, and α_i be according to (4.38)-(4.40). Then, an asymptotically correct approximation of the conditional variance $V[\hat{\tau}(\mathbf{y})_{OSA} | n']$ is;

$$V[\hat{\tau}(\mathbf{y})_{OSA} | n']_{appr} = \frac{N}{N-1} \cdot \sum_{i=1}^{N} \left(\frac{y_i}{\lambda'_i} - \sum_{j=1}^{N} \frac{y_j \cdot \alpha_j}{\lambda'_j} \Big/ \sum_{j=1}^{N} \alpha_j \right)^2 \cdot \lambda'_i \cdot (1-\lambda'_i), \tag{4.41}$$

APPROXIMATE CONDITIONAL VARIANCE FOR THE ESTIMATOR (4.31) : Let assumptions and notation be as in Procedure 4.3, and α_i as in (4.38)-(4.40). Then, an asymptotically correct approximation of the conditional variance $V[\hat{\tau}(y)_{OSGA} | n_1,..,n_G]$ is, where N_g denotes the number of units in group $\mathcal{G}_g$;

$$V[\hat{\tau}(y)_{OSGA} | n_1',\dots,n_G']_{appr} = \sum_{g=1}^{G} \frac{N_g}{N_g - 1} \cdot \sum_{i \in \mathcal{G}_g} \left(\frac{y_i}{\lambda_i'} - \sum_{j=1}^{N} \frac{y_j \cdot \alpha_j}{\lambda_j'} \Big/ \sum_{j=1}^{N} \alpha_j \right)^2 \cdot \lambda_i' \cdot (1 - \lambda_i'), \tag{4.42}$$

5 Simulation study of the non-response adjusted procedures

5.1 Introduction

The estimation procedures in Section 4 are derived by limit considerations, which are carried out in Sections 6 - 8. Since practical survey situations are "finite", procedures based on asymptotic results are always afflicted with some amount of approximation error when used in practice. In the present context, ***point estimator bias*** and ***confidence level bias*** are of particular interest. It is important to have an idea of the magnitude of the biases due to the approximations. The natural, not to say the only, feasible way to study approximation goodness is by Monte Carlo simulations, which is the theme of this section. We start by giving some background from the full response case.

5.1.1 On the performance of OSπps schemes under full response

Approximation goodness in the full response case is studied in Rosén (1997b) where, i.a. uniform, exponential and Pareto OSπps are considered. The findings can be summarized as follows. For all three schemes the point estimator bias is negligible already for very small sample sizes (say $n \geq 10$). For confidence levels the broad lines are as follows. Biases are "bearable" for fairly small sample sizes (say $n \geq 20$), "moderate" when sample size increases (say $n \geq 50$) and negligible for large sample size (say $n \geq 100$). However, a comprehensive picture is more complex, as always when goodness of the normal distribution approximation is involved. There is no simple rule of thumb for small confidence level bias in terms of just sample size.

Rosén (1997b) also compares the estimation precision for a number of πps schemes, including the present OSπps schemes, with the following chief result. Among πps schemes which admit objective assessment of sampling errors, Pareto OSπps is the superior one. It never performs worse than uniform and exponential OSπps, but sometimes considerably better. Its degree of superiority varies from situation to situation, though. Crucial factors in that context are the sampling rate and the relation between the study variable **y** and the size variable **s**. For the last factor, a fairly rough classification was introduced in Rosén (1997b). The relation between the study and size variables is said to have ***linear***, increasing ***convex*** or increasing ***concave trend***, depending on the shape of the curve for regression of **y** on **s**. In broad terms the following holds. As regards estimation precision the three OSπps schemes perform very similarly in situations with linear trend, but Pareto OSπps is superior in situations with convex or concave trend, the more superior the larger the sampling rate is.

5.1.2 Main aim of the present study of non-response adjusted estimation

Even if three OSπps schemes are considered in this paper, comparison of their over - all performances under non - response is not a chief aim, but to see how well the adjusted estimation procedures work for each of them. Background for this is that Statistics Sweden surveys use uniform as well as Pareto OSπps. Point estimator bias and confidence level bias are taken as the central performance measures. Under non - response additional potentially crucial factors enter the picture, notably;

- The response mechanism,
- The response propensities,
- The agreement between response mechanism and response model.

5.2 Organization of the simulation study

5.2.1 Broad lines

The broad lines in the simulation study are stated below. Details are given afterwards.

Step 1 Population generation: A population size, N, was decided on and value pairs $\{(y_i, s_i);\ i=1,2,\ldots,N\}$ for a study variable **y** and a size variable **s** were generated.

A generated population was used unaltered throughout the subsequent steps.

Step 2 Generation of responders: A response mechanism was decided on, and used to divide the population into responders and non-responders.

Step 3 Sample selection: A sample size, n, was decided on, and uniform, exponential and Pareto OSπps samples of size n were drawn.

Steps 2 and 3 yielded ***responding samples***. For fixed combinations of population, response mechanism and sample size, 3000 independent runs of (Step 1 + Step 2) were made.

Step 4 Estimates from responding samples: A response model, i.e. a set of surmised response homogeneity groups, was decided on. Then estimates from observation data for the responding samples were computed, using Estimation Procedures 4.2 and 4.3. Also approximate conditional estimator variances were computed.

The estimates from the 3000 responding samples were stored in a separate file.

Step 5 Compilation of performance measures: From the data in the "estimate store file" various means and standard deviations were computed, to provide measures for judging the performance of the adjusted estimation procedures.

5.2.2 Step 1: Population generation

Population values $\{(y_i, s_i);\ i=1,2,\ldots,N\}$ were generated by model (5.1) below, which is the same as the one used in Rosén (1997b);

$$s_i = i,\quad y_i = c\cdot(s_i^{\gamma} + \sigma\cdot L_i\cdot\sqrt{s_i^{\gamma}}),\quad c>0,\ \sigma\geq 0,\ \text{the } L_i \text{ being iid } N(0,1),\ i=1,2,\ldots,N. \quad (5.1)$$

The parameter γ determines the shape of the size-variable trend. For γ close to 1 the trend is linear, for γ greater than 1 it is convex and it is concave for γ less than 1. The parameter σ determines how much the **y**-values scatter around the trend curve.

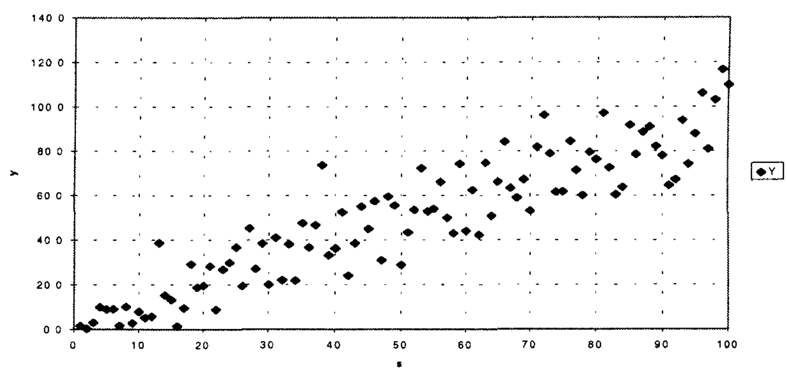
Four populations, named A, B, C and D, were considered. They were generated by the model (5.1) with parameters as stated below.

Population A: $N=100,\ \gamma=1,\ \sigma=2.$ Has linear trend.
Population B: $N=100,\ \gamma=1.5,\ \sigma=2.$ Has convex trend.
Population C: $N=100,\ \gamma=0.5,\ \sigma=0.5.$ Has concave trend.
Population D: $N=500,\ \gamma=0.5,\ \sigma=0.5.$ Has concave trend.

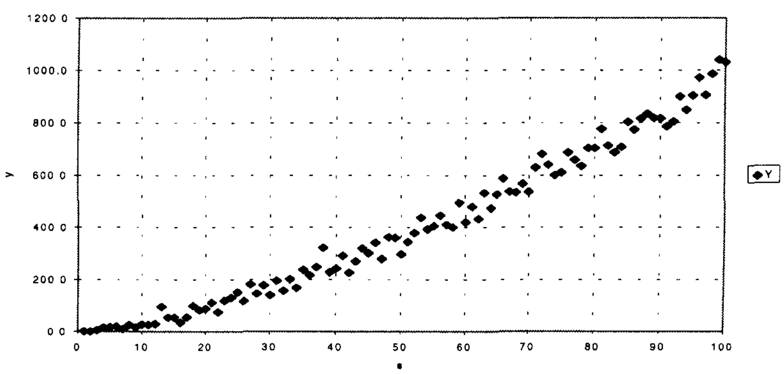
As seen, the populations range over the three trend shapes. Populations A, B and C have size N = 100, which certainly is a very small population size in a practical context. The main reason for considering such small populations was to keep computation times down. Moreover, if approximations in fact are quite good for small population and sample sizes, they are (reasonably) only better for larger ones. However, population D, which has the same γ and σ as population C but N = 500, was introduced to allow for larger sample sizes than possible for C.

The populations are illustrated graphically in the figures below.

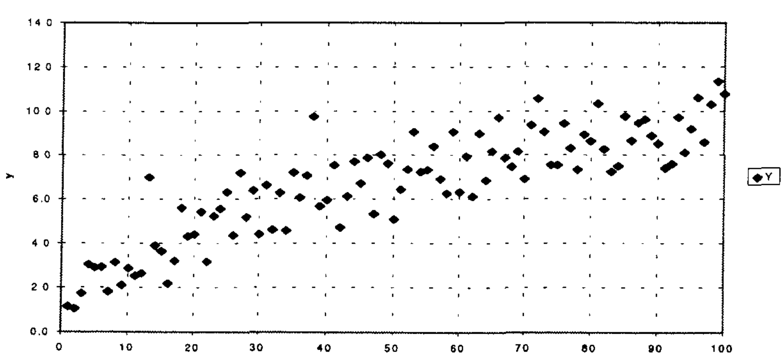
Population A



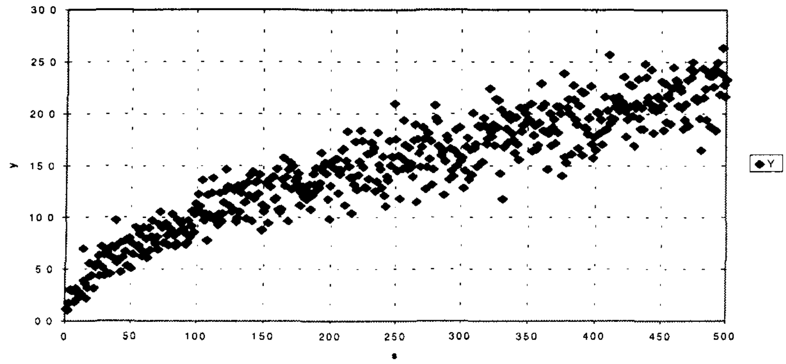
Population B



Population C



Population D



5.2.3 Step 2: Generation of responders

Responders were generated by realizing independent R - values in Response Mechanism 3.1. Models with over - all as well as group - wise uniform response propensities were employed, in the sequel referred to as ***Response Mechanisms*** **O** and **G** respectively.

For Response Mechanism O the following response propensities were used: β =70, 80, 90 and 100 %. Hence, the simulations also covered situations with full response.

For Response Mechanism G the following response homogeneity groups were employed;

$$\textbf{Group 1}: \{1 \le i \le 50\}, \quad \textbf{Group 2}: \{50 < i \le 75\}, \quad \textbf{Group 3}: \{75 < i \le 100\}. \tag{5.2}$$

Two sets of group - wise response propensities were employed;

a) $\beta_1 = 0.7$, $\beta_2 = 0.8$ and $\beta_3 = 0.9$. Response propensities increase with size. (5.3)

b) $\beta_1 = 0.9$, $\beta_2 = 0.8$ and $\beta_3 = 0.7$. Response propensities decrease with size. (5.4)

5.2.4 Step 3: Sample selection

Sample sizes were determined by prescription of sampling rates, which were chosen to be: 10, 20, 30, 40 and 50 %. As a consequence, for N = 100 the sample sizes were n = 10, 20, 30, 40 and 50 and for N = 500 they were n = 50, 100, 150, 200 and 250.

Uniform, exponential and Pareto OSπps samples were drawn by the procedures in Definitions 2.1 and 2.2. The same Z - variables [see (ii), Def. 2.2] were used for all three schemes.

Steps 2 and 3 were carried out with independent random numbers Z and R.

5.2.5 Step 4: Estimates from responding samples

The employed ***response models*** were "no grouping" and "grouping with the three $\mathcal{G}$ - groups in (5.2)", referred to as ***Response Model I*** and ***Response Model II***.

For each responding sample the estimator $\hat{\tau}(y)$ and its variance were estimated. Associated confidence intervals were computed, and it was checked if they contained the true $\tau(y)$ - value or not. Moreover were computed approximate conditional variances (which do not depend on the entire responding sample, though, only on the number of responding units).

Details of the estimation procedures are specified below, with labeling after response model.

Estimation procedure I

λ - modification: By (4.13).
Point estimator for $\tau(\mathbf{y})$: According to (4.14).
Variance estimators: According to (4.16), Type 1, as well as (4.20), Type 2.
Confidence interval: By (4.15)+(4.16), with approx. 95% confidence level, $\delta_{\alpha/2}=1.96$.
Approximate conditional estimator variance: According to (4.41).

Estimation procedure II

λ - modification: By (4.30).
Point estimator for $\tau(\mathbf{y})$: According to formula (4.31).
Variance estimators: According to (4.33) as well as (4.35).
Confidence interval: By (4.32)+(4.33), with approx. 95% confidence level, $\delta_{\alpha/2}=1.96$.
Approximate conditional estimator variance: According to (4.42).

5.2.7 Step 5: Summarizing performance measures

The final step comprised computation of various averages based on the estimates generated in the simulation runs, to yield performance measures for the adjusted estimation procedures.

Conditional versus unconditional performance measures

As stated in Section 4.4.1, the non - response adjusted estimation procedures are conditional inference procedures, with conditioning after number(s) of responding units. Accordingly the most informative type of evaluation would be by ***conditional performance measures***, computed by partitioning samples after the number of responding units before averaging. In contrast, ***unconditional performance measures*** are computed by averaging over all samples. In the sequel we confine to unconditional performance measures, though, for two main reasons. One is that presentation of conditional material is very space demanding. Another that reliable conditional means require very large numbers of simulation runs. Evaluations based on conditional means will be presented elsewhere. We want to mention, though, that we have computed conditional performance measures in some situations, and that they only support the subsequent conclusion drawn from the unconditional ones.

Before proceeding we give some justification of the fact that unconditional measures do yield information on the performances. First note the formulas : $E(E(X \mid W)) = E(X)$ and $V(X) = E[V(X \mid W)] + V[E(X \mid W)]$, which hold for any random variables X and W. If $E(X \mid W)$ = constant = $E(X)$, the last formula reduces to $V(X) = E[V(X \mid W)]$. With $X = \hat{\tau}(\mathbf{y})$ and $W = n'$ (= number of responding units, for simplicity we confine to over - all uniform response propensities) we get, where (5.6) holds provided that $\hat{\tau}(\mathbf{y})_{OSA}$ is (at least nearly) conditionally unbiased, i.e. if $E(\hat{\tau}(\mathbf{y})_{OSA} \mid n') \approx \tau(\mathbf{y})$;

$$E(\hat{\tau}(\mathbf{y})_{OSA}) = E[E(\hat{\tau}(\mathbf{y}) \mid n')] = \sum_q E(\hat{\tau}(\mathbf{y})_{OSA} \mid n' = q) \cdot P(n' = q), \tag{5.5}$$

$$V[\hat{\tau}(\mathbf{y})_{OSA}] \approx E[V[\hat{\tau}(\mathbf{y}) \mid n'] = \sum_q V[\hat{\tau}(\mathbf{y})_{OSA} \mid n' = q] \cdot P(n' = q). \tag{5.6}$$

From (5.5) and (5.6) the following is seen. If $\tau(\mathbf{y})$ agrees well with the average of $\hat{\tau}(\mathbf{y})_{OSA}$ over all generated samples, i.e. the empirical version of $E(\hat{\tau}(\mathbf{y})_{OSA})$, (5.5) tells that the majority of conditional means $E(\hat{\tau}(\mathbf{y})_{OSA} \mid n' = q)$ can neither all be upwards nor downwards biased. However, the possibility that there may be cancellation effects in the summation in (5.5) is not ruled out, but we believe that this is unlikely (and this belief is supported by the conditional means which have been studied). Hence, we take agreement of $\tau(\mathbf{y})$ and the empirical version of $E(\hat{\tau}(\mathbf{y})_{OSA})$ as a strong indication that also $E(\hat{\tau}(\mathbf{y})_{OSA} \mid n' = q)$ agrees with $\tau(\mathbf{y})$ for the majority of q:s. Analogously, agreement between the empirical version of $V[\hat{\tau}(y)_{OSA}]$ and the average of $V[\hat{\tau}(y)_{OSA}\,|n'=q]_{appr}$ over all samples is an indication that $V[\hat{\tau}(y)_{OSA}\,|n'=q]_{appr}$ yields good approximation of $V[\hat{\tau}(y)_{OSA}\,|n'=q]$ for the majority of q:s

Performance measures

Since averages are based on as many as 3 000 simulation runs, we are in the sequel a bit sloppy to distinguish between "empirical" and "true".

For point estimator bias

$$\textit{Relative point estimator bias} = \left(\tau(y) - \frac{1}{3000}\cdot\sum_{u=1}^{3000}\hat{\tau}(y)_u\right)/\,\tau(y)\cdot \tag{5.7}$$

For estimator precision

$$\textit{True estimator standard deviation} = \sqrt{\frac{1}{3000-1}\cdot\sum_{u=1}^{3000}(\hat{\tau}(y)_u - \overline{\hat{\tau}(y)})^2}\cdot \tag{5.8}$$

For variance estimators

$$\textit{Stdev. based on average variance estimates of Type 1} = \sqrt{\frac{1}{3000}\cdot\sum_{u=1}^{3000}\hat{V}[\hat{\tau}(y)]_u}\,. \tag{5.9}$$

$$\textit{Stdev. based on average variance estimates of Type 2} = \sqrt{\frac{1}{3000}\cdot\sum_{u=1}^{3000}\hat{V}'[\hat{\tau}(y)]_u}\,. \tag{5.10}$$

For the approximate conditional variance formulas

$$\textit{Stdev. based on average appr. conditional variances} = \sqrt{\frac{1}{3000}\cdot\sum_{u=1}^{3000}V[\hat{\tau}(y)\,|\,n']^{(u)}_{appr}}\,. \tag{5.11}$$

For confidence levels

$$\textit{Confidence level} = \frac{1}{3000}\cdot\sum_{u=1}^{3000}\mathrm{COVER}_u\,, \tag{5.12}$$

where COVER_u indicates if the confidence interval for the u:th responding sample covers $\tau(y)$.

5.3 Numerical simulation results and conclusions from them

5.3.1 Introduction

This section presents and discusses numerical results from the simulations. Aiming at easy reading, conclusions are presented first. When looking at the tables the reader may then agree or disagree in the conclusions. The numerical material is organized as follows. For the combinations AO, AG, BO, BG, CO, CG and DO of a population (A, B, C or D) and a response mechanism (O or G) a triple of tables, numbered .1, .2 and .3, is presented. The .1 tables concern point estimator bias, .2 tables estimator variances and .3 tables confidence levels.

The G - tables contain the classification ***adequate / non - adequate*** with the following meaning. Estimates are adequate if response mechanism and response model have the same homogeneity groups, otherwise non - adequate. Note that O - tables comprise only adequate estimates.

5.3.2 Concerning point estimator bias

From the O.1 tables and the adequate parts of the G.1 tables the following is seen. Under adequate response model, the non - response adjusted estimation procedures have negligible point estimator bias, relative biases are of order promille. This holds for all considered response rates and sample sizes and for all three OSπps schemes. The tables also show negligible point estimator bias under full response, in agreement with the findings in Rosén (1997b).

As can be expected matters differ, though, when the response model is non-adequate. The G-tables show that point estimator bias then is at hand, even if not of dramatically large.

Conclusion : Provided that the response model is adequate, the following holds for all three OSπps schemes under any of the two response models. Point estimator bias is negligible, for all considered non-response rates and already for very small sample sizes.

5.3.3 Concerning estimator variances

The O.2 and G.2 tables comprise information which allows evaluation with regards to different aspects of estimator variability.

The (a) rows tell true standard deviations for the point estimator.

The (b) and (c) rows show average point estimator standard deviations based on the variance estimators $\hat{V}$ and $\hat{V}'$ respectively.

The (d) row shows unconditional average point estimator standard deviations based on the approximate conditional variance formulas.

Comparison of the variance estimators $\hat{V}$ and $\hat{V}'$

From rows (b) and (c) is seen that the alternative variance estimators $\hat{V}$ and $\hat{V}'$, in (4.16) and (4.20) respectively in (4.33) and (4.35), lead to almost identical estimates. This holds for all three OSπps schemes, whether the response model is adequate or not, and for all considered non-response rates and sample sizes.

Conclusion: In practice it does not matter which of $\hat{V}$ and $\hat{V}'$ that is used.

Performance of the variance estimators

Comparison of row (a) with (b) and/or (c) in the O.2 tables and the adequate parts of the G.2 tables yields the following, for all three OSπps schemes. When response mechanism and response model have over-all uniform response propensities the variance estimators are fairly close to being unbiased for all considered response rates, at least when $n \geq 20$. Under grouped response mechanism with adequate response model, larger sample sizes are needed for the same conclusion, say $n \geq 30$. However, performance of variance estimators is not so interesting per se, the performance of the corresponding confidence intervals is more crucial from a practical point of view.

Tables AG.2, BG.2 and CG.2 also show that the variance estimators can be seriously biased when the response model is non-adequate.

Conclusion : The following holds for all three OSπps schemes and all considered response rates. Under adequate response model the variance estimators work quite satisfactorily, at least for "not too small" sample sizes. Under non-adequate response model the variance estimators may be seriously biased, though.

On point estimator precision

Section 5.1.1 summarized the findings in Rosén (1997b) on estimation precision under full response as follows. Pareto OSπps never performs notably worse than uniform and exponential OSπps, but sometimes considerably better. In situations with linear size-variable trend the three schemes perform very similarly, but Pareto OSπps is better in situations with curved trend, and is the more superior the larger the sampling rate is. The (a) rows in the tables show that this picture holds true under non-response as well, at least in O-situations but less pronounced in G-situations.

Inevitable, though, point estimates become less reliable the higher the non - response rate is, which is seen in all O tables. Table 1 below, based on the O - situations, exhibit the ***relative estimator variance increase due to non - response***, with full response point estimator variance as nominal value. We have also listed the corresponding variance increases for simple random sampling (SRS), using the formula: $\beta^{-1} \cdot (1 - \beta \cdot n/N)/(1 - n/N)$.

Table 1. Relative increase of estimator variances due to non - response

Situation	Resp. rate	Sampling rate 10%				20%				30%				40%				50%			
	β (%)	Uni	Exp	Par	SRS	Uni	Exp	Par	SRS	Uni	Exp	Par	SRS	Uni	Exp	Par	SRS	Uni	Exp	Par	SRS
AO	70	52	51	49	48	58	59	58	54	63	64	64	61	70	70	70	71	74	76	78	86
	80	30	29	28	28	32	32	32	31	37	37	37	36	39	39	40	42	42	42	45	50
	90	15	15	14	12	16	16	16	14	18	18	18	16	17	16	16	19	16	15	17	22
BO	70	58	57	55	48	59	59	63	54	64	67	68	61	79	86	88	71	100	125	130	86
	80	35	34	33	28	37	37	40	31	39	40	41	36	48	53	55	42	60	73	74	50
	90	18	17	16	12	17	18	20	14	19	20	21	16	24	26	28	19	27	31	31	22
CO	70	41	40	42	48	54	56	58	54	61	62	60	61	61	66	67	71	64	70	74	86
	80	26	25	34	28	31	34	36	31	37	38	36	36	36	39	41	42	38	42	44	50
	90	12	11	13	12	16	19	20	14	21	21	20	16	16	18	19	19	17	17	20	22
DO	70	44	43	44	48	45	45	44	54	46	45	46	61	54	56	56	71	55	61	62	86
	80	26	25	26	28	28	28	27	31	28	28	28	36	33	34	33	42	33	36	36	50
	90	12	12	12	12	11	10	10	14	9	9	10	16	15	15	14	19	14	15	15	22

The following conclusions are drawn from Table 1. (*i*) With exception for situations with very high sampling rates and pronounced curved size - variable trend, deterioration of point estimator precision due to non - response follows very much the same pattern as for simple random sampling. (ii) The point estimator precision for Pareto OSπps deteriorates slightly faster than for the other OSπps schemes. However, point estimator precision for Pareto OSπps never falls notably below that for the other two OSπps schemes.

Conclusion: Also under non - response holds that Pareto OSπps never has notably worse point estimator precision the other two OSπps schemes, but sometimes noteworthy better. With increasing non - response its edge over the other two schemes gradually lowers, though.

Performance of the approximate conditional variance formulas

Rows (d) and (a) in the tables show that what is said above about "good approximate unbiasedness" for the variance estimators also holds for "good approximation" for the approximate conditional variance formulas, yielding the conlusion below.

Conclusion : For all three OSπps schemes the following holds for all considered response rates. Under adequate response model the approximate conditional variance formulas work fairly satisfactorily, at least for "not too small" sample sizes. Under non - adequate response model the formulas may be quite misleading, though.

5.3.4 On confidence levels

Thinking that true levels even for our most commonly employed asymptotic result based confidence intervals in fact are not as good as we want to believe, the reaction to the confidence levels exhibited here is as follows. For populations A and B, with linear respectively convex size - variable trend, true confidence levels lie surprisingly close to the nominal one already for fairly small sample sizes. This holds irrespective of response rate. Some confidence level bias is at hand, though, in the direction that true confidence levels lie below nominal ones.

The picture for population C is, however, not so good. The agreement between the true and nominal confidence level for a confidence interval based on asymptotic considerations depends on two main factors; the accuracy of the variance estimate and the goodness of the normal distribution approximation. The distribution of $\hat{\tau}(\mathbf{y})$ is for populations of type C, with a concave size-variable trend, more skewed (with tail to the right) than for populations of types A and B, for reasons that we do not quite understand. Hence, good normal distribution approximations requires larger sample sizes in situations with concave trend. To allow for larger sample sizes than admitted by population C we introduced population D, which is of the same shape type as C ($\gamma=0.5$, $\sigma=0.5$) but has larger size, $N=500$. For population D the sample sizes n = 50, 100, 150, 200 and 250 were used. From Tables DO is seen that with increased sample sizes the true confidence levels come considerably closer to the nominal 95%.

Under non-adequate response model the confidence intervals may be quite misleading,though.

Conclusion : Provided that the response model is adequate the following holds. Under linear and convex size-variable trend the non-response adjusted confidence intervals have reasonably good confidence levels already for fairly small sample sizes. Under concave size-variable trend larger sample sizes are required to avoid skewed point estimator distribution with an adverse effect on confidence levels. Irrespective of trend shape the following holds. (i) The response rate has little influence on the confidence levels. (ii) The bias direction is that true confidence levels are lower than the nominal one.

With non-adequate response model the confidence intervals may be severely misleading.

5.3.5 Tentative over-all conclusions

The main conclusion from the simulation findings are, at least tentatively. Under adequate response model the non-response adjusted estimation procedures work quite satisfactorily for all three OSπps schemes. In particular, point estimator biases are negligible for all (reasonable) non-response rates already for quite small sample sizes. The behavior of the confidence intervals is more complex. Response rate has little importance in this context, though, the shape of the size-variable trend seems to be the crucial factor. Under linear and convex trends true and nominal confidence levels agree quite well even for fairly small sample sizes. Under concave trend considerably larger sample sizes are require for small confidence level bias.

Under non - adequate response model the adjusted estimation procedures may be quite misleading. However, this is not particular to OSπps sampling, it holds for any sampling scheme.

5.3.6 Numerical results

Tables AO

Population A, Response Mechanism O, Estimation Procedure I.

Table AO.1 Relative point estimator bias (in %)

Resp. prop. β (%)	Sampling rate 10%			20%			30%			40%			50%		
	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare
70	0.02	0.04	0.02	0.15	0.15	0.13	0.06	0.07	0.05	0.07	0.06	0.04	0.03	0.03	0.03
80	-0.05	-0.05	-0.06	0.11	0.11	0.09	-0.00	0.01	-0.02	0.02	0.01	-0.01	0.01	-0.00	0.00
90	0.01	0.02	0.01	0.14	0.14	0.13	0.03	0.03	0.03	0.05	0.04	0.02	0.01	0.00	0.00
100	-0.04	-0.05	-0.06	0.03	0.02	0.05	-0.02	-0.02	-0.01	-0.01	0.01	0.02	-0.01	0.00	0.01

Table AO.2 Estimator standard deviations and standard deviation estimation

(a) = True point estimator standard deviations by (5.8)
(b) = Average estimated standard deviation by (4.16) + (5.9)
(c) = Average estimated standard deviations by (4.20) + (5.10)
(d) = Average approximate theoretical standard deviations by (4.41) + (5.11)

Resp. prop. β (%)		Sampling rate 10%			20%			30%			40%			50%		
		Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare
70	(a)	497.4	498.0	497.4	339.3	340.7	340.8	264.8	265.1	266.0	218.3	218.6	219.0	183.5	183.9	184.8
	(b)	500.9	501.2	502.0	337.1	337.7	338.3	263.3	263.8	263.9	218.0	217.7	218.2	185.7	185.5	186.3
	(c)	500.9	501.3	502.0	337.1	337.7	338.2	263.3	263.8	263.9	218.0	217.7	218.0	185.7	185.5	185.7
	(d)	503.9	503.6	503.5	338.6	338.5	338.5	264.9	265.0	265.0	219.1	219.1	219.1	186.8	186.8	186.7
80	(a)	460.0	460.5	460.6	310.3	311.3	312.0	242.4	242.5	242.5	197.4	197.7	198.5	165.6	165.5	166.5
	(b)	459.4	459.4	459.8	309.1	309.7	310.4	240.6	241.0	241.0	197.7	197.4	197.9	166.8	166.6	167.1
	(c)	459.4	459.4	459.8	309.1	309.7	310.3	240.6	241.0	241.0	197.7	197.5	197.8	166.8	166.7	166.8
	(d)	462.7	462.6	462.5	311.0	311.0	311.0	242.2	242.2	242.2	198.9	199.0	199.0	168.0	168.0	168.0
90	(a)	433.3	434.2	433.8	290.7	291.2	291.9	225.2	225.6	225.3	180.7	180.8	181.2	149.5	148.8	149.7
	(b)	428.3	428.3	428.7	287.2	287.6	288.1	222.4	222.7	222.8	181.2	180.9	181.3	151.3	150.9	151.1
	(c)	428.3	428.3	428.7	287.2	287.6	288.1	222.4	222.8	222.8	181.2	181.0	181.2	151.3	151.0	151.0
	(d)	431.0	430.9	431.0	289.3	289.3	289.3	223.7	223.7	223.7	182.2	182.2	182.2	152.2	152.2	152.2
100	(a)	404.1	405.3	407.1	270.1	270.5	271.6	207.4	207.3	207.6	167.3	167.9	168.0	139.0	138.8	138.4
	(b)	403.4	403.9	403.6	268.2	268.5	268.6	206.5	206.4	206.4	166.6	166.4	166.3	137.6	136.9	136.8
	(c)	403.4	403.9	403.6	268.2	268.5	268.6	206.5	206.4	206.4	166.6	166.4	166.3	137.6	136.9	136.8
	(d)	403.2	403.2	403.2	270.0	270.0	270.0	207.4	207.4	207.3	167.5	167.4	167.4	138.1	138.1	138.1

Table AO.3 Confidence levels (in %)

Resp. prop. β (%)	Sampling rate 10%			20%			30%			40%			50%		
	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare
70	90.6	90.7	90.7	93.2	93.0	92.8	93.5	93.3	93.1	94.4	94.4	94.4	94.6	94.4	94.5
80	91.6	91.6	91.6	93.6	93.5	93.3	93.5	93.4	93.6	94.7	94.6	94.5	94.4	94.1	94.2
90	91.2	91.2	91.2	93.5	93.6	93.5	93.8	93.4	93.6	94.3	94.2	93.9	93.9	94.0	94.1
100	91.9	91.8	91.7	93.6	93.7	93.7	93.8	93.9	94.0	94.2	94.0	93.6	93.4	92.5	92.7

Tables AG

Population A, Response Mechanism G, Estimation Procedures II and I.

Table AG.1 Relative point estimator bias (in %)

Response mecha-nism	Estimation procedure	Sampling rate											
		20%			30%			40%			50%		
		Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare
Ga	II (adequate)	0.49	0.48	0.44	0.10	0.09	0.09	0.11	0.09	0.07	0.02	0.02	0.02
See (5.3)	I (non-adequate)	0.01	0.01	-0.00	-0.11	-0.10	-0.11	-0.09	-0.11	-0.12	-0.12	-0.12	-0.12
Gb	II (adequate)	0.28	0.27	0.25	0.10	0.09	0.09	0.10	0.09	0.07	0.05	0.03	0.04
See (5.4)	I (non-adequate)	0.28	0.29	0.27	0.19	0.19	0.18	0.20	0.19	0.17	0.16	0.15	0.15

Table AG.2 Estimator standard deviations and standard deviation estimation

(a) = True point estimator standard deviations by (5.8)
(b) = Average estimated standard deviation by (4.33) + (5.9)
(c) = Average estimated standard deviations by (4.35) + (5.10)
(d) = Average approximate theoretical standard deviations by (4.42) + (5.11)

Response mecha-nism	Estimation procedure		Sampling rate											
			20%			30%			40%			50%		
			Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare
Ga	II (adequate)	(a)	388.6	388.2	385.4	268.0	267.5	267.9	216.7	216.1	216.6	178.6	177.9	178.5
		(b)	324.7	325.5	324.9	263.3	262.4	262.6	218.2	217.3	217.2	181.0	180.3	181.0
		(c)	324.7	325.5	324.9	263.3	262.5	262.5	218.2	217.3	217.1	181.0	180.4	180.8
		(d)	356.9	356.0	356.2	272.5	271.6	271.2	220.0	219.2	218.9	184.0	183.3	183.2
See (5.3)	I (non-adeq.)	(a)	295.0	296.0	296.4	228.6	229.2	229.2	185.9	186.1	186.7	154.2	154.0	154.6
		(b)	267.5	268.1	268.6	214.8	215.3	215.2	178.4	178.2	178.5	150.9	150.7	150.9
		(c)	267.5	268.1	268.6	214.8	215.3	215.2	178.4	178.3	178.5	150.9	150.8	150.9
		(d)	306.0	306.0	306.0	238.0	238.1	238.1	195.3	195.3	195.3	164.7	164.7	164.6
Gb	II (adequate)	(a)	352.1	351.5	351.7	252.8	252.2	252.0	204.2	203.4	203.1	166.6	165.7	166.5
		(b)	317.1	317.0	316.8	248.9	248.3	248.2	202.5	201.5	201.6	167.4	167.0	167.4
		(c)	317.1	317.0	316.8	248.9	248.3	248.2	202.5	201.5	201.6	167.4	167.0	167.2
		(d)	335.6	335.2	335.0	252.1	251.4	251.5	203.5	202.9	202.7	169.2	168.6	168.6
See (5.4)	I (non-adeq.)	(a)	328.9	329.4	330.2	257.2	257.2	257.3	210.7	210.9	211.0	176.1	175.8	176.8
		(b)	301.0	301.5	302.2	243.3	243.8	243.8	203.9	203.8	204.2	174.5	174.4	174.7
		(c)	301.0	301.6	302.2	243.3	243.8	243.8	203.9	203.8	204.2	174.5	174.5	174.6
		(d)	314.9	314.8	314.8	245.6	245.5	245.5	202.1	202.0	202.0	171.0	170.9	170.9

Table AG.3 Empirical confidence levels (in %)

Response mecha-nism	Estimation procedure	Sampling rate											
		20%			30%			40%			50%		
		Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare
Ga	II (adequate)	85.3	85.1	85.3	91.1	91.1	91.0	93.1	93.1	92.9	93.4	93.4	93.2
See (5.3)	I (non-adequate)	89.6	89.5	89.4	91.9	91.9	91.8	92.8	92.7	92.3	92.8	92.7	92.8
Gb	II (adequate)	88.8	88.7	88.7	91.8	91.6	91.8	92.9	92.7	92.8	93.6	93.6	93.5
See (5.4)	I (non-adequate)	89.9	89.7	89.6	91.1	90.9	91.0	92.3	92.1	91.9	93.1	93.0	93.1

Tables BO

Population B, Response Mechanism O, Estimation Procedure I.

Table BO.1 Relative point estimator bias (in %)															
Resp. prop.	Sampling rate														
β	10%			20%			30%			40%			50%		
(%)	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare
70	-0.11	-0.07	-0.03	0.02	0.08	0.13	-0.04	0.04	0.05	-0.07	-0.03	-0.01	0.00	0.00	0.02
80	-0.13	-0.10	-0.06	-0.02	0.05	0.10	-0.09	-0.02	-0.01	-0.11	-0.06	-0.02	-0.02	-0.01	-0.00
90	-0.10	-0.07	-0.04	0.01	0.07	0.11	-0.04	0.03	0.05	-0.07	-0.03	0.01	-0.00	0.00	0.01
100	-0.14	-0.08	-0.06	-0.11	-0.07	-0.01	-0.09	-0.03	0.02	-0.10	-0.02	0.01	-0.02	-0.02	-0-02

Table BO.2 Estimator standard deviations and standard deviation estimation

(a) = True point estimator standard deviations by (5.8)
(b) = Average estimated standard deviation by (4.16) + (5.9)
(c) = Average estimated standard deviations by (4.20) + (5.10)
(d) = Average approximate theoretical standard deviations by (4.41) + (5.11)

Resp. prop.		Sampling rate														
β		10%			20%			30%			40%			50%		
(%)		Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare
70	(a)	3364	3351	3343	2244	2239	2249	1745	1731	1723	1445	1404	1396	1214	1157	1151
	(b)	3336	3336	3337	2239	2235	2238	1742	1733	1741	1433	1412	1434	1215	1187	1245
	(c)	3336	3336	3336	2239	2236	2236	1742	1736	1733	1433	1417	1413	1215	1184	1176
	(d)	3341	3338	3338	2239	2233	2235	1746	1734	1741	1438	1416	1436	1220	1191	1247
80	(a)	3106	3095	3088	2079	2078	2085	1607	1584	1576	1315	1274	1265	1086	1013	1002
	(b)	3053	3053	3053	2053	2047	2049	1590	1576	1579	1297	1267	1278	1088	1032	1066
	(c)	3053	3053	3053	2053	2048	2048	1590	1579	1575	1297	1273	1268	1088	1042	1030
	(d)	3067	3064	3064	2055	2048	2048	1594	1579	1581	1303	1272	1281	1094	1036	1069
90	(a)	2905	2896	2890	1927	1923	1934	1485	1468	1462	1203	1159	1149	967.8	882.6	868.8
	(b)	2829	2829	2828	1905	1898	1897	1464	1448	1446	1184	1146	1145	980.3	895.7	905.2
	(c)	2829	2830	2828	1905	1898	1897	1464	1449	1445	1184	1150	1142	980.3	908.3	893.5
	(d)	2856	2853	2853	1910	1902	1901	1469	1451	1449	1190	1150	1149	985.7	899.9	908.7
100	(a)	2676	2673	2678	1779	1774	1764	1364	1339	1329	1081	1031	1017	858.8	771.2	758.6
	(b)	2655	2656	2655	1773	1765	1765	1354	1334	1329	1084	1038	1027	884.1	768.0	750.9
	(c)	2655	2656	2655	1773	1765	1765	1354	1334	1329	1084	1038	1027	884.1	768.0	750.9
	(d)	2670	2668	2667	1780	1771	1769	1359	1338	1332	1090	1042	1030	889.4	773.2	755.6

Table BO.3 Confidence levels (in %)															
Resp. prop.	Sampling rate														
β	10%			20%			30%			40%			50%		
(%)	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare
70	87.5	87.8	87.9	92.0	92.0	92.0	93.4	93.3	93.3	93.1	93.7	94.0	93.8	94.3	95.3
80	88.3	88.5	88.6	91.8	91.8	91.6	92.6	92.7	92.8	92.9	93.4	93.8	93.2	93.7	94.7
90	89.0	89.2	89.5	92.5	92.6	92.4	93.1	92.9	93.0	93.1	93.3	93.4	93.8	93.8	94.5
100	90.4	90.4	90.3	92.3	92.3	92.5	92.8	92.8	93.2	93.1	93.3	93.7	93.7	93.5	93.4

Tables BG

Population B, Response Mechanism G, Estimation Procedures II and I

Table BG.1 Relative point estimator bias (in %)

Response mechanism	Estimation procedure	Sampling rate 20%			30%			40%			50%		
		Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare
Ga	II (adequate)	0.27	0.26	0.25	-0.02	-0.02	-0.01	-0.02	-0.02	-0.02	-0.07	-0.06	-0.05
See (5.3)	I (non-adequate)	-1.87	-1.81	-1.77	-1.94	-1.88	-1.86	-1.97	-1.93	-1.90	-1.90	-1.90	-1.89
Gb	II (adequate)	0.12	0.11	0.11	-0.03	-0.02	-0-02	-0.02	-0-02	-0.01	-0.02	-0.02	-0.02
See (5.4)	I (non-adequate)	1.98	2.05	2.10	1.94	2.02	2.04	1.91	1.96	2.00	1.97	1.99	2.00

Table BG.2 Estimator standard deviations and standard deviation estimation

(a) = True point estimator standard deviations by (5.8)
(b) = Average estimated standard deviation by (4.33) + (5.9)
(c) = Average estimated standard deviations by (4.35) + (5.10)
(d) = Average approximate theoretical standard deviations by (4.42) + (5.11)

Response mechanism	Estimation procedure		Sampling rate 20%			30%			40%			50%		
			Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare
Ga	II (adequate)	(a)	1510	1504	1485	882.3	877.6	886.2	710.1	705.5	704.7	575.4	570.6	566.5
		(b)	1057	1053	1052	854.4	850.7	851.4	701.9	699.9	701.3	585.7	582.4	583.6
		(c)	1057	1053	1052	854.4	850.8	850.9	701.9	700.0	700.5	585.7	582.3	581.8
		(d)	1147	1144	1145	874.4	871.2	869.9	703.6	700.3	699.8	586.3	582.7	583.2
See (5.3))	I (non-adeq.)	(a)	1968	1959	1970	1516	1504	1497	1238	1197	1188	1021	956.2	945.9
		(b)	867.9	868.3	869.6	693.1	693.8	693.6	572.4	572.3	572.8	482.1	481.0	482.1
		(c)	867.9	868.4	869.4	693.1	693.8	693.4	572.4	572.5	572.6	482.1	480.7	480.4
		(d)	985.4	985.3	985.3	764.4	763.9	764.0	624.8	623.8	623.9	524.6	523.1	523.8
Gb	II (adequate)	(a)	1356	1339	1339	842.7	840.9	839.9	676.3	674.3	672.8	533.8	527.0	525.0
		(b)	1042	1040	1037	807.3	804.2	803.9	651.4	648.7	648.8	538.7	538.2	539.8
		(c)	1042	1040	1037	807.3	804.2	803.6	651.4	648.6	648.2	538.7	536.5	536.0
		(d)	1084	1083	1082	812.3	809.6	809.7	653.6	650.9	650.4	541.4	539.3	540.3
See (5.4)	I (non-adeq.)	(a)	2160	2158	2168	1681	1661	1656	1379	1338	1331	1131	1058	1048
		(b)	976.3	977.2	978.0	784.4	784.9	784.5	654.1	653.6	654.3	557.5	557.4	558.6
		(c)	976.3	977.2	978.0	784.4	785.0	784.3	654.1	653.8	654.0	557.5	557.0	556.7
		(d)	1015	1014	1014	788.8	788.0	788.1	647.1	645.9	646.0	545.2	544.0	545.1

Table BG.3 Empirical confidence levels (in %)

Response mechanism	Estimation procedure	Sampling rate 20%			30%			40%			50%		
		Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare
Ga	II (adequate)	84.5	84.7	84.6	89.9	90.0	90.2	91.3	91.6	91.7	92.2	92.2	92.4
See (5.3)	I (non-adequate)	51.8	52.5	52.7	53.1	54.0	54.6	50.4	52.3	53.6	51.5	52.6	52.9
Gb	II (adequate)	88.0	88.2	88.3	91.6	91.8	91.8	92.8	92.7	92.6	93.3	93.9	94.2
See (5.4)	I (non-adequate)	57.6	57.6	57.3	58.3	57.6	57.8	57.3	57.7	57.8	55.6	58.3	58.6

Tables CO

Population C, Response Mechanism O, Estimation Procedure I.

Table CO.1 Relative point estimator bias (in %)															
Resp. prop.	Sampling rate														
β	10%			20%			30%			40%			50%		
(%)	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare
70	0.33	0.28	0.21	0.15	0.06	-0.03	0.10	0.00	-0.02	0.18	0.11	0.02	0.08	0.05	0.01
80	0.23	0.20	0.12	0.17	0.09	-0.02	0.10	0.01	-0.01	0.17	0.11	0.01	0.09	0.05	0.02
90	0.24	0.20	0.14	0.17	0.08	0.00	0.06	-0.03	-0.06	0.15	0.09	-0.00	0.05	0.02	-0.01
100	0.18	0.08	0.06	0.20	0.14	0.10	0.12	0.03	-0.02	0.14	0.06	0.04	0.06	0.06	0.07

Table CO.2 Estimator standard deviations and standard deviation estimation

(a) = True point estimator standard deviations by (5.8)
(b) = Average estimated standard deviation by (4.16) + (5.9)
(c) = Average estimated standard deviations by (4.20) + (5.10)
(d) = Average approximate theoretical standard deviations by (4.41) + (5.11)

Resp. prop.		Sampling rate														
β		10%			20%			30%			40%			50%		
(%)		Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare
70	(a)	110.1	110.1	110.9	77.8	78.3	78.8	62.9	62.7	62.3	53.3	52.9	52.7	46.3	45.4	45.5
	(b)	109.3	109.6	110.1	77.6	77.8	78.2	62.1	62.1	62.1	52.5	52.0	52.6	45.7	45.2	46.5
	(c)	109.3	109.6	110.1	77.6	77.8	78.1	62.1	62.1	61.9	52.5	52.1	52.1	45.7	45.3	45.3
	(d)	114.4	114.3	114.3	78.3	78.2	78.2	62.6	62.4	62.5	53.0	52.6	53.0	46.4	45.9	46.8
80	(a)	104.1	104.1	107.9	72.0	72.5	73.0	58.1	57.9	57.3	49.0	48.4	48.4	42.5	41.4	41.4
	(b)	101.6	101.8	102.2	71.2	71.3	71.7	57.2	57.0	57.0	48.3	47.6	48.0	41.9	41.2	41.9
	(c)	101.6	101.8	102.2	71.2	71.4	71.6	57.2	57.1	56.9	48.3	47.8	47.8	41.9	41.4	41.3
	(d)	105.3	105.3	105.3	72.4	72.3	72.3	57.8	57.6	57.6	48.9	48.4	48.5	42.7	41.8	42.3
90	(a)	98.0	98.2	98.9	67.6	68.2	68.5	54.5	54.2	53.9	45.3	44.6	44.5	39.1	37.7	37.8
	(b)	95.7	96.0	96.3	66.6	66.7	66.9	53.6	53.4	53.4	45.1	44.3	44.5	39.1	37.9	38.1
	(c)	95.7	96.0	96.3	66.6	66.7	66.9	53.6	53.4	53.3	45.1	44.4	44.4	39.1	38.1	38.0
	(d)	98.4	98.4	98.3	67.8	67.6	67.6	54.0	53.7	53.6	45.5	44.9	44.9	39.6	38.4	38.5
100	(a)	92.7	93.1	93.1	62.8	62.6	62.6	49.6	49.2	49.2	42.0	41.1	40.8	36.2	34.8	34.5
	(b)	91.3	91.6	91.3	62.9	62.8	62.7	50.0	49.7	49.6	42.1	41.3	41.1	36.5	34.7	34.4
	(c)	91.3	91.6	91.3	62.9	62.8	62.7	50.0	49.7	49.6	42.1	41.3	41.1	36.5	34.7	34.4
	(d)	92.3	92.3	92.3	63.7	63.5	63.5	50.6	50.3	50.2	42.6	41.9	41.7	37.0	35.4	35.1

Table CO.3 Confidence levels (in %)															
Resp. prop.	Sampling rate														
β	10%			20%			30%			40%			50%		
(%)	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare
70	80.4	80.4	80.5	84.4	84.2	84.3	84.7	84.9	85.1	84.5	85.1	85.8	85.3	85.5	86.5
80	81.4	81.4	81.7	84.0	83.9	83.9	84.4	84.7	84.9	84.1	84.4	85.4	84.7	85.3	85.9
90	81.5	81.6	81.9	84.3	84.3	84.3	83.9	84.0	84.3	85.4	85.7	86.3	84.8	85.3	85.9
100	82.1	82.4	82.3	84.3	84.3	84.5	85.6	86.0	86.2	85.8	86.5	86.5	86.0	86.2	85.9

Tables CG

Population C, Response Mechanism G, Estimation Procedures II and I

Table CG.1 Relative point estimator bias (in %)

Response mechanism	Estimation procedure	Sampling rate 20%			30%			40%			50%		
		Unif	Exp	Paret	Unif	Exp	Paret	Unif	Exp	Paret	Unif	Exp	Paret
Ga	II (adequate)	0.67	0.68	0.62	0.18	0.17	0.18	0.21	0.20	0.14	0.15	0.12	0.10
See (5.3)	I (non-adequate)	2.44	2.36	2.29	2.37	2.29	2.27	2.44	2.38	2.30	2.35	2.33	2.30
Gb	II (adequate)	0.43	0.41	0.38	0.18	0.15	0.15	0.18	0.17	0.11	0.15	0.10	0.08
See (5.4)	I (non-adequate)	-2.23	-2.31	-2.41	-2.31	-2.42	-2.45	-2.24	-2.30	-2.41	-2.33	-2.38	-2.41

Table CG.2 Estimator standard deviations and standard deviation estimation

(a) = True point estimator standard deviations by (5.8)
(b) = Average estimated standard deviation by (4.33) + (5.9)
(c) = Average estimated standard deviations by (4.35) + (5.10)
(d) = Average approximate theoretical standard deviations by (4.42) + (5.11)

Response mechanism	Estimation procedure		Sampling rate 20%			30%			40%			50%		
			Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare
Ga	II (adequate)	(a)	76.9	76.3	75.8	55.3	54.4	54.2	46.0	45.7	45.7	38.6	38.7	39.0
		(b)	62.6	62.3	62.1	53.1	52.9	52.8	45.0	44.7	44.9	38.4	38.5	38.7
		(c)	62.6	62.3	62.1	53.1	52.9	52.8	45.0	44.7	44.8	38.4	38.5	38.6
		(d)	71.7	71.5	71.5	56.2	56.0	55.9	46.6	46.4	46.3	40.2	40.0	40.0
See (5.3)	I (non-adeq.)	(a)	66.1	66.7	67.0	53.7	53.5	53.2	45.5	45.1	44.9	39.4	38.4	38.5
		(b)	48.3	48.6	48.8	40.8	40.8	40.7	35.1	35.0	35.1	30.9	30.9	31.0
		(c)	48.3	48.6	48.8	40.8	40.8	40.7	35.1	35.0	35.1	30.9	30.9	31.0
		(d)	59.6	59.6	59.6	47.7	47.6	47.6	40.4	40.3	40.3	35.3	35.2	35.2
Gb	II (adequate)	(a)	63.2	62.9	63.0	47.6	47.4	47.4	39.9	39.5	39.5	33.6	33.7	33.9
		(b)	57.8	57.7	57.8	47.1	46.9	46.9	39.7	39.3	39.4	33.8	33.9	34.0
		(c)	57.8	57.7	57.8	47.1	46.9	46.9	39.7	39.3	39.4	33.8	33.9	33.9
		(d)	63.1	63.0	62.9	48.7	48.4	48.4	40.5	40.3	40.2	34.9	34.6	34.6
See (5.4)	I (non-adeq.)	(a)	77.4	78.0	78.5	62.3	62.0	61.7	52.1	51.4	51.4	45.1	43.8	43.8
		(b)	58.4	58.7	59.0	48.9	49.0	48.9	42.2	42.0	42.2	37.2	37.2	37.4
		(c)	58.4	58.7	59.0	48.9	49.0	48.9	42.2	42.1	42.2	37.2	37.3	37.4
		(d)	61.2	61.1	61.1	49.0	48.9	48.9	41.5	41.4	41.4	36.4	36.3	36.3

Table CG.3 Empirical confidence levels (in %)

Response mechanism	Estimation procedure	Sampling rate 20%			30%			40%			50%		
		Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare	Unif	Exp	Pare
Ga	II (adequate)	78.9	78.9	79.0	82.3	82.6	82.6	83.2	83.5	83.6	83.7	84.0	83.7
See (5.3)	I (non-adequate)	64.4	64.4	64.8	64.3	65.0	65.1	63.8	64.3	64.9	64.9	65.3	65.2
Gb	II (adequate)	83.3	83.2	83.4	84.4	84.8	84.8	85.1	85.3	85.5	85.1	85.4	85.2
See (5.4)	I (non-adequate)	77.1	76.9	76.9	79.7	80.4	80.7	82.0	82.9	83.7	84.0	85.2	85.7

Tables DO

Population D (N=500), Response Mechanism O, Estimation Procedure I.

Table DO.1 Relative point estimator bias (in %)

Resp. prop. β (%)	10% Unif	10% Exp	10% Pare	20% Unif	20% Exp	20% Pare	30% Unif	30% Exp	30% Pare	40% Unif	40% Exp	40% Pare	50% Unif	50% Exp	50% Pare
70	0.05	0.04	0.01	0.12	0.12	0.13	0.08	0.09	0.08	0.04	0.05	0.05	0.04	0.04	0.04
80	0.05	0.04	0.01	0.11	0.11	0.12	0.06	0.06	0.05	0.03	0.03	0.03	0.02	0.02	0.02
90	0.04	0.03	0.01	0.12	0.12	0.13	0.07	0.08	0.07	0.04	0.05	0.05	0.04	0.05	0.04
100	0.07	0.04	0.01	-0.07	-0.09	-0.10	-0.00	-0.02	-0.02	0.01	-0.00	-0.01	0.00	-0.00	-0.01

Table DO.2 Estimator standard deviations and standard deviation estimation

(a) = True point estimator standard deviations by (7.4)
(b) = Estimated standard deviations (4.16) + (7.5)
(c) = Estimator standard deviations (4.20) + (7.6)
(d) = Approximate theoretical standard deviations (4.41) + (7.8)

Resp. prop β (%)		10% Unif	10% Exp	10% Pare	20% Unif	20% Exp	20% Pare	30% Unif	30% Exp	30% Pare	40% Unif	40% Exp	40% Pare	50% Unif	50% Exp	50% Pare
70	(a)	486.3	485.3	487.2	340.5	341.2	340.2	274.8	271.8	271.5	233.7	229.9	228.3	203.1	197.1	195.8
	(b)	491.4	491.3	493.1	340.7	340.3	339.5	272.8	271.1	272.0	230.4	228.1	229.9	201.5	198.7	204.2
	(c)	491.4	491.3	493.1	340.7	340.4	339.3	272.8	271.4	271.4	230.4	228.6	228.1	201.5	198.9	198.4
	(d)	493.6	493.4	493.5	342.3	341.8	341.9	273.9	272.8	273.4	232.4	230.4	232.1	203.5	200.7	205.7
80	(a)	455.3	454.6	455.9	320.7	320.9	320.2	256.8	254.7	254.3	217.6	213.3	210.9	188.1	181.4	179.8
	(b)	460.2	460.2	461.7	316.3	315.8	315.0	252.5	250.7	251.0	213.0	210.1	210.8	185.8	181.0	184.1
	(c)	460.2	460.2	461.7	316.3	315.9	314.9	252.5	251.0	250.8	213.0	210.7	210.0	185.8	182.0	181.2
	(d)	460.2	460.0	460.0	318.2	317.6	317.7	254.0	252.7	252.9	214.8	212.1	212.8	187.4	182.6	185.3
90	(a)	429.3	429.0	430.5	297.7	297.6	297.5	237.8	235.8	235.1	202.4	197.6	195.7	174.0	166.7	165.3
	(b)	432.4	432.2	433.4	296.3	295.7	295.0	236.0	234.1	233.9	198.4	194.9	194.7	172.2	165.4	166.3
	(c)	432.4	432.2	433.4	296.3	295.8	294.9	236.0	234.2	233.8	198.4	195.4	194.4	172.2	166.6	165.5
	(d)	432.3	432.1	432.1	298.3	297.6	297.5	237.3	235.7	235.5	200.0	196.8	196.6	173.8	167.2	167.7
100	(a)	405.0	405.9	406.7	283.1	283.2	283.9	227.4	225.4	224.5	188.4	184.2	183.0	163.1	155.4	154.0
	(b)	403.6	403.7	404.1	282.4	281.9	282.1	222.6	220.8	220.3	186.6	182.6	181.5	161.2	152.7	151.5
	(c)	403.6	403.7	404.1	282.4	281.9	282.1	222.6	220.8	220.3	186.6	182.6	181.5	161.2	152.7	151.5
	(d)	408.7	408.5	408.4	281.2	280.4	280.2	223.0	221.2	220.7	187.2	183.5	182.4	162.0	153.7	152.1

Table DO.3 Confidence levels (in %)

Resp. prop. β (%)	10% Unif	10% Exp	10% Pare	20% Unif	20% Exp	20% Pare	30% Unif	30% Exp	30% Pare	40% Unif	40% Exp	40% Pare	50% Unif	50% Exp	50% Pare
70	90.0	89.9	90.0	90.7	90.7	90.8	91.5	91.6	91.6	91.6	91.6	91.9	92.6	92.4	93.2
80	89.9	89.9	90.0	91.0	90.9	90.9	91.3	91.4	91.6	92.0	91.8	92.2	92.5	92.5	93.0
90	90.0	90.1	90.2	91.5	91.4	91.2	92.0	91.9	91.8	92.3	91.9	91.9	92.0	92.0	92.4
100	90.9	91.0	90.9	92.0	92.2	92.2	91.7	91.9	92.1	92.3	92.1	92.3	92.6	92.2	92.0

6 Quota random retain order sampling

Here we enter the theoretical justifications of the non-response adjusted estimation procedures in Section 4. They are admittedly a bit involved, but the simplest we know of. Unfortunately the case with non-response cannot be brought back on the full response case "right away", but on the other hand situations with non-response are more complicated than full response situations. In the following an instrumental role is played by an auxiliary type of sampling scheme, called "quota random retain order sampling". Although this notion lies close to "order sampling with random non-response" it is introduced as a separate concept, an important reason being that the crucial limit result will concern this sampling scheme.

6.1 Definition

To each unit i in $U=(1,2,...,N)$ is associated a pair (Q_i, R_i) of random variables. The Q:s are called ***ranking variables*** and the R:s ***retain indicators*** with ***retain frequency*** ρ. The following assumptions are made.

$Q_1, Q_2, ..., Q_N$ are independent random variables with distributions $\mathbf{F}=(F_1, F_2, ..., F_N)$, where F_i is a probability distribution on $[0,\infty)$ with density function f_i. (6.1)

$R_1, R_2, ..., R_N$ are independent equally distributed 0-1 variables with $P(R_i=1)=\rho$. (6.2)

The random vectors $(Q_1, Q_2, ..., Q_N)$ and $(R_1, R_2, ..., R_N)$ are independent. (6.3)

The sampling process starts with realization of $(Q_1, Q_2, ..., Q_N)$ and $(R_1, R_2, ..., R_N)$. Thereafter at least the following two variations of "order sampling with random retain" can be thought of.

The Q:s select a first-stage order sample of (fixed) size n from all of U, according to Definition 2.1. The final sample consists of the retained units (those with $R=1$) in the first-stage sample. Then the size n' of the final sample is random. (6.4)

A sample size n' is prescribed. The sample is selected as the retained units (those with $R=1$) with the n' smallest Q-values. Then the sample size n' is fixed, but it may happen that the desired sample size cannot be reached. (6.5)

The (6.4) variant is nothing but order sampling with non-response. However, in the sequel we consider the (6.5) type of sampling scheme.

DEFINITION 6.1: ***Quota random retain order sampling*** with (fixed) ***sample size*** n', ***order distributions*** $\mathbf{F}$ and ***retain frequency*** ρ, denoted OSRRQ(n'; $\mathbf{F}$; ρ), is carried out as follows. Variables $Q_1, Q_2, ..., Q_N$ and $R_1, R_2, ..., R_N$ that satisfy (6.1) - (6.3) are realized. The sample consists of the units with $R=1$ which have the n' smallest Q-values.

***Remark 6.1*:** For $\rho=1$, OSRRQ(n'; $\mathbf{F}$; ρ) is simply "ordinary" OS(n'; $\mathbf{F}$). ¤

6.2 Approximate distribution of an OSRRQ sample sum

As a preparation for derivation of estimation procedures for OSRRQ we consider the distribution of an OSRRQ sample sum. For a variable $\mathbf{z}=(z_1, z_2 ..., z_N)$ and a sample we set;

The $\mathbf{z}$ ***sample sum*** $= \sum_{i \in \text{Sample}} z_i$. (6.6)

APPROXIMATION RESULT 6.1: $\mathbf{z}=(z_1, z_2, ..., z_N)$ is a variable on $U=(1,2,..,N)$, from which an OSRRQ(n'; $\mathbf{F}$; ρ) sample is drawn. S(n'; $\mathbf{z}$) denotes the corresponding $\mathbf{z}$ sample sum. Then, with μ and σ as specified below, (6.7) holds under general conditions;

The distribution of S(n'; $\mathbf{z}$) is well approximated by $N(\mu, \sigma^2)$. (6.7)

Specification of parameters

ξ solves the equation (in t): $\sum_{i=1}^{N} F_i(t) = n'/\rho, \quad 0 \leq t < \infty,$ (6.8)

$$\mu = \rho \cdot \sum_{i=1}^{N} z_i \cdot F_i(\xi), \tag{6.9}$$

$$\phi = \sum_{i=1}^{N} z_i \cdot f_i(\xi) \Big/ \sum_{i=1}^{N} f_i(\xi), \tag{6.10}$$

$$\sigma^2 = \sum_{i=1}^{N} (z_i - \phi)^2 \cdot \rho \cdot F_i(\xi) \cdot [1 - \rho \cdot F_i(\xi)]. \tag{6.11}$$

Justification of the above result is given in Section 8.

6.3 Estimation of a population total from an OSRRQ sample

6.3.1 Estimation when the retain frequency is known

ESTIMATION PROCEDURE 6.1: $\mathbf{y} = (y_1, y_2, ..., y_N)$ is a variable on $U = (1,...,N)$, from which an OSRRQ(n';F;ρ) sample is drawn and fully observed. The retain frequency ρ is known. The ***OSRRQ estimator*** of the total $\tau(\mathbf{y})$ is, where ξ is according to (6.8);

$$\hat{\tau}(\mathbf{y})_{OSRRQ} = \sum_{i \in \text{Sample}} y_i / [\rho \cdot F_i(\xi)]. \tag{6.12}$$

Then the following holds under general conditions.

a. The distribution of $\hat{\tau}(\mathbf{y})_{OSRRQ}$ is well approximated by $N(\tau(\mathbf{y}), \chi^2)$, with (6.13)

$$\chi^2 = \frac{N}{N-1} \cdot \sum_{i=1}^{N} (y_i / [\rho \cdot F_i(\xi)] - \gamma)^2 \cdot \rho \cdot F_i(\xi) \cdot (1 - \rho \cdot F_i(\xi)), \tag{6.14}$$

$$\gamma = \sum_{i=1}^{N} (y_i / [\rho \cdot F_i(\xi)]) \cdot f_i(\xi) \Big/ \sum_{i=1}^{N} f_i(\xi). \tag{6.15}$$

b. In particular, $\hat{\tau}(\mathbf{y})_{OSRRQ}$ yields consistent estimation of $\tau(\mathbf{y})$.

c. In particular, χ^2 yields an asymptotically correct value for $V[\hat{\tau}(\mathbf{y})_{OSRRQ}]$

d. A consistent estimator of $V[\hat{\tau}(\mathbf{y})_{OSRRQ}]$ is given by;

$$\hat{V}[\hat{\tau}(\mathbf{y})_{OSRRQ}] = \frac{n'}{n'-1} \cdot \sum_{i \in \text{Sample}} (y_i / [\rho \cdot F_i(\xi)] - \hat{\gamma})^2 \cdot [1 - \rho \cdot F_i(\xi)], \text{ where} \tag{6.16}$$

$$\hat{\gamma} = \frac{1}{\rho} \cdot \sum_{i \in \text{Sample}} [y_i / F_i(\xi)^2] \cdot f_i(\xi) \Big/ \sum_{i \in \text{Sample}} f_i(\xi) / F_i(\xi). \tag{6.17}$$

Start of justification: The above result can be derived from Approximation Result 6.1 by an almost verbatim repetition of the arguments for justification of Approximation Result 3.2 in Rosén (1997a). A reader interested in more details than given below is referred to that paper.

We regard Approximation Result 6.1 as justified (although this is not done until Section 8). The a-, b- and c-parts in the above procedure are obtained by choosing $z_i = y_i / [\rho \cdot F_i(\xi)]$ and noting that μ in (6.9) then becomes $\tau(\mathbf{y})$. In the step from (6.11) to (6.14) we make a somewhat ad hoc modification, to the effect that the factor $N/(N-1)$ is introduced. The reason is that with this "extra" factor, (6.14) becomes the correct formula in the case when all order distributions are equal (which entails that order sampling is nothing but simple random sampling). This (asymptotically negligible) modification yields improved approximation also in the general situation. In (6.16) we introduce the factor $n'/(n'-1)$ with the same motivation.

For the d-part, we view the sum in (6.14) as a population total. Its estimate by (6.12) is;

$$\hat{\chi}^2 = \frac{N}{N-1} \cdot \sum_{i \in \text{Sample}} (y_i/[\rho \cdot F_i(\xi)] - \gamma)^2 \cdot [1 - \rho \cdot F_i(\xi)]. \tag{6.18}$$

The estimate (6.18) has, however, the "defect" that γ is unknown. We circumvent this crux in the usual way, i.e. by inserting an estimate of γ into (6.18). The estimate is constructed as the ratio of estimators of nominator and denominator in (6.15). Estimation by (6.12) yields;

$$\sum_{i=1}^{N} (y_i/[\rho \cdot F_i(\xi)]) \cdot f_i(\xi) \quad \text{is estimated by} \quad \sum_{i \in \text{Sample}} y_i \cdot f_i(\xi)/[\rho \cdot F_i(\xi)]^2 , \tag{6.19}$$

$$\sum_{i=1}^{N} f_i(\xi) \quad \text{is estimated by} \quad \sum_{i \in \text{Sample}} f_i(\xi)/[\rho \cdot F_i(\xi)] . \tag{6.20}$$

The estimates in (6.19) and (6.20) now lead to the γ - estimate in (6.17). ¤

6.3.2 Estimation when the retain frequency is unknown

In Estimation Procedure 6.1 the retain frequency ρ is presumed to be known, and it enters in the estimators (6.12) and (6.16). The natural modification of the estimation procedure to a situation where ρ is unknown is to exchange ρ for an estimate of it.

ESTIMATION PROCEDURE 6.2: $\mathbf{y} = (y_1, y_2, ..., y_N)$ is a variable on $U = (1,2,...,N)$, from which an OSRRQ $(n'; H; \underline{\lambda}; \rho)$ sample is drawn. The retain frequency ρ is unknown, but an estimate $\hat{\rho}$ of it is available. The $\mathbf{y}$-values for all sampled units are observed.

The claims in Procedure 6.1 then hold after the modification that ρ is exchanged for $\hat{\rho}$. An important aspects of the modification is that equation (6.8) modifies to;

$$\xi \text{ solves the equation (in t):} \quad \sum_{i=1}^{N} F_i(t) = n'/\hat{\rho}, \quad 0 \le t < \infty. \tag{6.21}$$

7 Derivation of the estimation procedures in Section 4

We now return to order πps sampling with non-response. The non-response adjusted estimation procedures will be based on the following general inference principle.

Principle for conditional inference: A population characteristic is to be estimated from observations on a random sample. If a specific statistic carries no, or only little, information about the characteristic of interest (is ancillary), inference should preferably be made conditional on the ancillary statistics.

7.1 The case with over-all uniform response propensities

We start with Procedure 4.2, where assumptions are as follows. $\mathbf{y} = (y_1, y_2, ..., y_N)$ is a variable on the population $U = (1, 2, ..., N)$. The estimation interest concerns the total $\tau(\mathbf{y})$. An OS$(n; \mathbf{F})$ sample is drawn from U with the aim to observe $\mathbf{y}$ for all selected units. However, only n' ($\le$ n) units respond. The task is then to estimate $\tau(\mathbf{y})$ from the y-values for the responding units.

The statistic n' = "number of responding units" is regarded as ancillary under estimation of $\tau(\mathbf{y})$. Hence, the above conditioning principle says that we shall condition on n', and regard it as fixed. The following conditioning result should be evident upon some thought.

LEMMA 7.1: An OSπps$(n; H; \underline{\lambda})$ sample is selected from U. During data collection, non-responses occur by Response Mechanism 3.1 with over-all uniform response propensity β. Conditional on the number n' of responding units, the $\mathfrak{R}$sample is probabilistically equivalent to an OSRRQ$(n'; H; \underline{\lambda}; \beta)$ sample from U.

The $\mathfrak{R}$sample is completely observed for **y**. Hence, Lemma 7.1 entails that Procedures 6.1 and 6.2 can be applied to the $\mathfrak{R}$sample. Since we presume that the response propensity (= retain frequency) is unknown, Procedure 6.2 with $\rho = \beta$ is the relevant result. The first task is to exhibit a β - estimate. We use the following "natural" β - estimate under Response Mechanism 3.1 with over - all uniform response propensities;

$$\hat{\beta} = n'/n. \tag{7.1}$$

With this β - estimate and order distributions according to (2.3), equation (6.21) becomes;

$$\xi \text{ solves the equation (in t): } \sum_{i=1}^{N} H(t \cdot H^{-1}(\lambda_i)) = n'/\hat{\beta} = \text{[by (6.1)]} = n, \quad 0 \leq t < \infty. \tag{7.2}$$

From (2.2) is readily seen that (7.2) is solved by $\xi = 1$. With this ξ we have;

$$F_i(\xi) = F_i(1) = H(H^{-1}(\lambda_i)) = \lambda_i \quad \text{and} \quad f_i(\xi) = f_i(1) = h(H^{-1}(\lambda_i)) \cdot H^{-1}(\lambda_i), \quad i = 1, 2, \ldots, N. \tag{7.3}$$

Now, insertion as stated in Procedure 6.2 into the formulas in Procedure 6.1 leads, after some straightforward algebra which is left to the reader, to the claims in Procedure 4.2. In the insertion step note the following consequences of the formulas in (7.3): (i) $\beta \cdot F_i(\xi)$ shall be exchanged for $\hat{\beta} \cdot F_i(\xi) = (n'/n) \cdot \lambda_i =$ by (4.13) $= \lambda_i'$, (ii) $f(\xi) = \alpha_i = \lambda_i \cdot a_i$, with a_i and α_i according to (4.4) and (4.30). Moreover, insertion into (6.14) and (6.15) yields (4.34).

7.2 The case with group - wise uniform response propensities

The case with group - wise uniform response propensities is treated along the same lines as above. Lemma 7.1 is modified as follows, and again the result should be clear upon thought.

> **LEMMA 7.2:** An OSπps(n; H; $\underline{\lambda}$) sample is selected from U. During data collection, non-responses occur by Response Mechanism 3.1 with group - wise uniform response propensities, with response uniformity groups $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_G$. Conditional on the numbers $(n'_1, n'_2, \ldots, n'_G)$ of responding units from the different groups, the $\mathfrak{R}_g$sample is probabilistically equivalent with an OSRRQ(n'_g; H; $\underline{\lambda}$; $\beta^{(g)}$) sample from $\mathcal{G}_g$, $g = 1, 2, \ldots, G$.

Let;

$$\tau(\mathbf{y})_g = \text{the } \mathbf{y}\text{-total over group } \mathcal{G}_g, \quad g = 1, 2, \ldots, G. \tag{7.4}$$

Then we have;

$$\tau(\mathbf{y}) = \tau(\mathbf{y})_1 + \tau(\mathbf{y})_2 + \ldots + \tau(\mathbf{y})_G, \tag{7.5}$$

which leads to the following estimation formula;

$$\hat{\tau}(\mathbf{y}) = \hat{\tau}(\mathbf{y})_1 + \hat{\tau}(\mathbf{y})_2 + \ldots + \hat{\tau}(\mathbf{y})_G. \tag{7.6}$$

Lemma 7.2 tells that the estimator (4.14) can be used group - wise in (7.6), which leads to (4.25). As regards the variance of the estimator (7.6), we make the following observation.

Conditional on $(n'_1, n'_2, \ldots, n'_G)$ the $\mathfrak{R}_g$samples are independent, $g = 1, 2, \ldots, G$. (7.7)

By (7.6) and (7.7) we arrive at;

$$V[\hat{\tau}(\mathbf{y}) \mid n_1', n_2', \ldots, n_G'] = \sum_{g=1}^{G} V[\hat{\tau}(\mathbf{y})_g \mid n_g'], \tag{7.8}$$

which in turn leads to the variance estimation formula;

$$\hat{V}[\hat{\tau}(\mathbf{y}) \mid n_1', n_2', \ldots, n_G'] = \sum_{g=1}^{G} \hat{V}[\hat{\tau}(\mathbf{y})_g \mid n_g']. \tag{7.9}$$

By applying the variance estimator (4.16) in (7.9), the variance estimator (4.27) is obtained. Analogously, application of (4.34) in (7.8) yields (4.35).

8 Justification of Approximation Result 6.1

The chief task in this section is to justify Approximation Result 6.1. Thereby the main step will be derivation of a limit result for the distribution of OSRRQ sample sums, which is formulated in Theorem 8.1.

8.1 A limit theorem for quota retain order sampling

THEOREM 8.1: For $k=1,2,\ldots$, an OSRRQ($n'_k;\mathbf{F}_k;\rho_k$) sample is drawn from the population $U_k=(1,2,\ldots,N_k)$. The density for F_{ki} is denoted f_{ki}. $\mathbf{z}_k=(z_{k1},z_{k2},\ldots,z_{kN_k})$ is a variable on U_k and $S_k(n'_k;\mathbf{z}_k)$ is the corresponding $\mathbf{z}_k$ sample sum. Let ξ_k, μ_k, ϕ_k and σ_k be in accordance with (7.8)-(7.11). Then, with $\Rightarrow$ denoting convergence in law;

$$\text{Law}[\,(S_k(n_k;\mathbf{z}_k)-\mu_k)/\sigma_k] \Rightarrow N(0,1),\ \text{ as } k\to\infty, \tag{8.1}$$

provided that conditions (B1) - (B6) below are satisfied.

(B1) $n'_k \to \infty$ as $k\to\infty$.

(B2) $\max_i |z_{ki}-\phi_k|/\sigma_k \to 0$, as $k\to\infty$.

(B3) $0<\underline{\lim}_{k\to\infty} \frac{\rho_k\cdot\xi_k}{n'_k}\cdot\sum_{i=1}^{N_k} f_{ki}(\xi_k) \le \overline{\lim}_{k\to\infty} \frac{\rho_k\cdot\xi_k}{n'_k}\cdot\sum_{i=1}^{N_k} f_{ki}(\xi_k)<\infty$.

(B4) For some $\delta>0$, some $C<\infty$ and some function $w(\Delta)$ which tends to 0 as $\Delta\to 0$, the following inequalities hold for $1-\delta\le t,s\le 1+\delta$;

$$|f_{ki}(t\cdot\xi_k)-f_{ki}(s\cdot\xi_k)| \le C\cdot w(|t-s|)\cdot f_{ki}(\xi_k),\ i=1,2,\ldots,N_k,\ k=1,2,3,\ldots . \tag{8.2}$$

(B5) For some $\eta>0$, we have;

$$F_{ki}(\xi_k)\cdot(1-\rho_k\cdot F_{ki}(\xi_k)) \ge \eta\cdot\xi_k\cdot f_{ki}(\xi_k),\ i=1,2,\ldots,N_k,\ k=1,2,3,\ldots . \tag{8.3}$$

(B.6) $\underline{\lim}_{k\to\infty} \frac{\rho_k\cdot N_k}{n'_k}>1$.

***Remark 8.1*:** The concrete contents of condition (B.6) is that the expected number of retained units, $\rho_k\cdot N_k$, must exceed the desired sample size n'_k. ¤

***Remark 8.2*:** Even if the conditions in Theorem 8.1 look quite complicated they are in fact satisfied under very general conditions in many, not to say most, types of sampling situations of practical interest. Illustrations are given in Rosén (1997a & b). ¤

Before proving Theorem 8.1 we discuss how Approximation Result 7.1 follows from it.

8.2 Justification of Approximation Result 6.1 from Theorem 8.1

Approximation result 6.1 is obtained from Theorem 8.1 by employing the limit distribution already in the finite situation, i.e. by the approximation Law[$(S(n;\mathbf{z})-\mu)/\sigma] \approx N(0,1)$, or equivalently Law[$S(n;\mathbf{z})] \approx N(\mu,\sigma^2)$. The "algebraic step" from Theorem 8.1 to Approximation Result 7.1 is straightforward, and left to the reader. The practical follow - up problem is then : When is a situation sufficiently "close to infinity" for the approximations to be good enough for practical purposes? This problem was discussed Section 5.

8.2 Proof of Theorem 8.1

Theorem 8.1 is very similar to Theorem 3.1 in Rosén (1997a), henceforth referred to by RTh3.1. To the best of our understanding, though, it is not similar enough to be a corollary to RTh3.1, it must be given its own proof. This, however, will be very parallel to that of RTh3.1. Therefore we leave out a great deal of proof details, and refer the interested reader to the proof of RTh3.1 for more details. The chief proof tool for Theorem 8.1 will be Theorem 5.1 in Rosén (1997a), for simplicity referred to by RTh5.1 in the sequel. Analogously, formula (x.y) in Rosén (1997a) is referred to as (R.x.y). We start with some preparations for application of RTh5.1. To begin with we disregard the sequence context, and omit subscripts k.

$Q_1, Q_2, ..., Q_N$ and $R_1, R_2, ..., R_N$ stand for the ranking variables and retain indicators in Definition 6.1. We introduce the following stochastic processes, where $\mathbf{1}_A$ is the indicator of the set A, $\mathbf{z}=(z_1, z_2 ..., z_N)$ are given real numbers and $\xi > 0$ is an arbitrary (but fixed) number;

$$H_i(t) = \mathbf{1}_{\{Q_i \le t\}}, \quad 0 \le t < \infty, \quad i=1,2,...,N, \tag{8.4}$$

$$J(t) = \sum_{i=1}^{N} H_i(t \cdot \xi) \cdot R_i, \quad 0 \le t < \infty, \text{ and } \quad L(t;\mathbf{z}) = \sum_{i=1}^{N} z_i \cdot H_i(t \cdot \xi) \cdot R_i, \quad 0 \le t < \infty. \tag{8.5}$$

The contents of these processes are as follows. $H_i(t)$ "signalizes" (by jumping from 0 to 1) the value of the ranking variable Q_i. J(t) tells the number retained units with ranking variable values $\le$ t, and $L(t;\mathbf{z})$ is the $\mathbf{z}$-sum for these units.

Let T* denote the ***passage time*** when J(t) reaches the level n', and let $A(n';\mathbf{z})$ denote the corresponding ***passage z-sum***;

$$T^* = \inf(t: J(t) = n'), \tag{8.6}$$

$$A(n';\mathbf{z}) = L(T^*;\mathbf{z}) = \sum_{i=1}^{N} z_i \cdot H_i(T^* \cdot \xi) \cdot R_i. \tag{8.7}$$

The claim in (8.8) below should be evident upon some thought. It provides the crucial link between an OSRRQ sample sum and passage variables, which enables application of RTh5.1.

An OSRRQ sample sum $S(n';\mathbf{z})$ has the same distribution as $A(n';\mathbf{z})$ in (8.7). (8.8)

We now turn to the limit considerations, and let index *k* signify that a quantity relates to situation *k*. Let $J_k(t)$, $L_k(t;\mathbf{z}_k)$, T_k^* and $A_k(n'_k;\mathbf{z}_k)$ be in accordance with (8.4) - (8.7), and let ξ_k, μ_k, ϕ_k and σ_k be in accordance with (6.8) - (6.11). As a consequence of (8.8) we have;

$$\text{Law}[(S_k(n'_k;\mathbf{z}_k) - \mu_k)/\sigma_k] = \text{Law}[(A_k(n'_k;\mathbf{z}_k) - \mu_k)/\sigma_k], \quad k=1,2,3,... . \tag{8.9}$$

We employ RTh5.1 to show that $(A_k(n'_k;\mathbf{z}_k) - \mu_k)/\sigma_k$ is asymptotically N(0,1) distributed under the conditions in Theorem 8.1. Having done that, (8.9) tells that also $(S_k(n'_k;\mathbf{z}_k) - \mu_k)/\sigma_k$ is asymptotically N(0,1) distributed, and Theorem 8.1 is proved.

For future use we list the following formulas, which are straightforward consequences of the fact that $H_{ki}(t) \cdot R_{ki}$ is a 0-1 random variable with expected value $F_i(t) \cdot \rho_k$;

$$E[H_{ki}(t) \cdot R_{ki}] = \rho_k \cdot F_{ki}(t), \quad V[H_{ki}(t) \cdot R_{ki}] = \rho_k \cdot F_{ki}(t) \cdot [1 - \rho_k \cdot F_{ki}(t)], \quad 0 \le t < \infty. \tag{8.10}$$

The entities Y_k, X_k, x_k, y_k, V_k, U_k, ε_k and τ_k in RTh5.1 are chosen as stated in (8.11) - (8.17), and the interval under consideration is $[t_0, t_1] = [1-\delta, 1+\delta]$, where δ comes from condition (B4). Moreover, ^{c} denotes centering at expectation. When checking (8.11) - (8.15), recall the expectation formula in (8.10).

$$Y_k(t) = \frac{1}{n'_k}\sum_{i=1}^{N} H_{ki}(t\cdot\xi_k)\cdot R_i = y_k(t) + \frac{1}{\sqrt{n'_k}}\cdot V_k(t),\ t_0 \le t \le t_1, \tag{8.11}$$

$$y_k(t) = \frac{\rho_k}{n'_k}\sum_{i=1}^{N_k} F_{ki}(t\cdot\xi_k), \qquad V_k(t) = \frac{1}{\sqrt{n'_k}}\sum_{i=1}^{N_k}[H_{ki}(t\cdot\xi_k)\cdot R_i]^c,\ t_0 \le t \le t_1, \tag{8.12}$$

$$X_k(t) = \frac{1}{\sqrt{n'_k}}\sum_{i=1}^{N_k}\frac{z_{ki}-\phi_k}{\sigma_k}\cdot H_{ki}(t\cdot\xi_k)\cdot R_{ki} = x_k(t) + \frac{1}{\sqrt{n'_k}}\cdot U_k(t),\ t_0 \le t \le t_1, \tag{8.13}$$

$$x_k(t) = \frac{\rho_k}{\sqrt{n'_k}}\sum_{i=1}^{N_k}\frac{z_{ki}-\phi_k}{\sigma_k}\cdot F_{ki}(t\cdot\xi_k),\ t_0 \le t \le t_1, \tag{8.14}$$

$$U_k(t) = \sum_{i=1}^{N_k}\frac{z_{ki}-\phi_k}{\sigma_k}\cdot[H_{ki}(t\cdot\xi_k)\cdot R_{ki}]^c,\ t_0 \le t \le t_1, \tag{8.15}$$

$$\varepsilon_k = 1/\sqrt{n'_k},\ \ k=1,2,3,\dots, \tag{8.16}$$

$$\tau_k = 1,\ \ k=1,2,3,\dots. \tag{8.17}$$

(B.6), (8.12) and (6.8) yield that t^*_k in (R.5.6), i.e. the solution to $y_k(t) = \tau_k = 1$, takes the value $t^*_k = 1$ for all k. Condition (B.6) is an "extra" condition in comparison with RTh3.1. Its concrete meaning is discussed in Remark 8.2. Technically it guarantees that the range of $y_k(t)$ contains 1, so that $y_k(t) = 1$ in fact has a solution which also is an interior point in $[t_0, t_1]$.

The proof is divided into two parts, stated in (8.18) and (8.19) below. Together they yield that $(A_k(n'_k;\mathbf{z}_k) - \mu_k)/\sigma_k$ is asymptotically N(0,1) distributed.

$$\text{Under conditions (B1) - (B6):}\quad \frac{A_k(n'_k;\mathbf{z}_k)-\mu_k}{\sigma_k} - U_k(1) \xrightarrow{P} 0,\ \text{ as } k\to\infty. \tag{8.18}$$

$$\text{Under conditions (B1) and (B2):}\quad \mathrm{Law}[U_k(1)] \Rightarrow N(0,1),\ \text{ as } k\to\infty. \tag{8.19}$$

We start by proving the simplest part, (8.19).

Proof of (8.19) **:** Since $Q_{k1},\dots,Q_{kN}$ are independent, so are $H_{ki}(t)$, $i = 1,\dots,N_k$. Hence, cf. (8.15),

$$U_k(1) = \sum_{i=1}^{N_k}\frac{z_{ki}-\phi_k}{\sigma_k}\cdot[H_{ki}(\xi_k)\cdot R_i]^c \tag{8.20}$$

is a sum of independent random variables with means 0. By (6.11) and the variance formula in (8.10) it is readily checked that the variance of $U_k(1)$ is 1. Liapunov's condition in 4:th moment version yields the asymptotic normality in (8.19). See the proof of RTh3.1 for details.

Proof of (8.18) **:** First we derive expressions for the two terms in (R.5.8). By (8.13), (8.5) and (8.7) is readily checked that;

$$X_k(T^*_k) = \frac{L_k(T^*_k;\mathbf{z}_k) - \phi_k\cdot J_k(T^*_k)}{\sigma_k\cdot\sqrt{n'_k}} = \frac{A_k(n'_k;\mathbf{z}_k)-\phi_k\cdot n'_k}{\sigma_k\cdot\sqrt{n'_k}}. \tag{8.21}$$

By (8.14), $t^*_k = 1$, (6.9) and (6.8) we have;

$$x_k(t^*_k) = x_k(1) = \frac{\mu_k - \phi_k\cdot n'_k}{\sigma_k\cdot\sqrt{n'_k}}. \tag{8.22}$$

From (8.21), (8.22) and (8.16) we get;

$$\frac{X_k(T^*_k) - x_k(t^*_k)}{\varepsilon_k} = \frac{A_k(n'_k;\mathbf{z}_k)-\mu_k}{\sigma_k}. \tag{8.23}$$

The formulas for y_k and x_k in (8.12) and (8.14), and the assumption that the F_{ki} have densities imply that $y_k(t)$ and $x_k(t)$ are differentiable on $[t_0, t_1]$, with the following derivatives;

$$y'_k(t) = \frac{\rho_k \cdot \xi_k}{n'_k} \cdot \sum_{i=1}^{N_k} f_{ki}(t \cdot \xi_k), \quad t_0 \leq t \leq t_1, \tag{8.24}$$

$$x'_k(t) = \frac{\rho_k \cdot \xi_k}{\sqrt{n_k}} \cdot \sum_{i=1}^{N_k} \frac{z_{ki} - \phi_k}{\sigma_k} \cdot f_{ki}(t \cdot \xi_k), \quad t_0 \leq t \leq t_1. \tag{8.25}$$

From (8.25), $t_k^* = 1$ and (6.10) follows that $x'_k(t_k^*) = x'_k(1) = 0$. Hence, the term to the right in (R.5.8) simplifies to $U_k(1)$. This together with (8.23) yields that *if* RTh5.1 applies in the present situation, it leads to (8.18). This in conjunction with (8.19) yields the algebraic part of Theorem 8.1. It remains, though, to show that RTh5.1 in fact does apply, to the effect that conditions (A.1) - (A.8) in it in fact are satisfied.

At this junction we confine to saying the following. With conditions (B.1) - (B.5) in RTh3.1 modified to those in Theorem 8.1, verifications of (A.1) - (A.8) in RTh5.1 are obtained by paralleling the verifications in the proof of RTh3.1 making the "obvious" modifications at appropriate places. ¤

References

Rosén, B. (1997a). Asymptotic Theory for Order Sampling. *J. Statist. Planning and Inf.*, **62** 135 - 158.

Rosén, B. (1997b). On Sampling with Probability Proportional to Size. *J Stat. Plan. Inf.*, **62** 159 - 191.

Rosén, B. (1998). On Inclusion Probabilities for Order Sampling. Stat. Sweden R&D Report 1988:2.

Förteckning över utkomna R&D Reports

R&D Reports är en för IT-enheten och Metodenheten gemensam publikationsserie, som 1988-01-01 ersatte de tidigare "gula" och "gröna" serierna. I serien ingick fram till årsskiftet 1992-93 även **Abstracts** (sammanfattning av metodrapporter från SCB).

Reports published during 1996 and onwards:

1996:1 (grön)	On Sampling with Probability Proportional to Size *(Bengt Rosén)*
1996:2 (grön)	Bortfallsbarometern nr 11 *(Antti Ahtiainen, Stefan Berg, Margareta Eriksson, Åsa Greijer, Dan Hedlin, Monica Rennermalm och Anita Ullberg)*
1996:3 (grön)	Regression Estimators in Theory and in Practice *(Tomas Garås)*
1996:4 (gul)	Quality Aspects of a Modern Database Service *(Pat Dean and Bo Sundgren)*
1996:5 (gul)	Metadata: A Quality Element in Official Statistics - the Swedish Approach *(Bo Sundgren and Pat Dean)*
1996:6 (gul)	Our Legacy to Future Generation - Using Databases for Better Availability and Documentation *(Gösta Guteland and Erik Malmborg)*
1997:1 (grön)	Bortfallsbarometern nr 12 *(Anti Ahtiainen, Stefan Berg, Mats Bergdahl, Fredrik Granström, Dan Hedlin, Lena Otterskog och Monica Rennermalm)*
1997:2 (grön)	Quality Concept for Official Statistics - Entry in the forthcoming update of the Encyclopedia of Statistical Sciences, Wiley & Sons *(Eva Elvers and Bengt Rosén)*
1998:1 (grön)	Preliminär statistik: Nybyggnadskostnader - en simuleringstudie *(Catarina Elffors)*
1998:2 (grön)	On Inclusion Probabilities for Order Sampling *(Bengt Rosén)*
1998:3 (grön)	On the Stratification of Highly Skewed Populations *(Dan Hedlin)*
1998:4 (grön)	Bortfallsbarometer nr 13 *(Per Nilsson, Antti Ahtiainen, Stefan Berg, Mats Bergdahl, Monica Rennermalm och Marcus Vingren)*
1998:5 (grön)	Estimation from Order πps Samples with Non - Response *(Bengt Rosén and Pär Lundqvist)*

ISSN 0283-8680

Tidigare utgivna **R&D Reports** kan beställas genom Katarina Klingberg, SCB, MET, Box 24 300, 104 51 STOCKHOLM (telefon 08-783 42 82, fax 0[illegible]

R&D Reports from 1988-1995 can - in case they ar[illegible]
attn. Katarina Klingberg, MET, Box 24 300, SE-104[illegible]
fax +46 8 783 45 99, e-mail katarina.klingberg@scb[illegible]

090693-002101- 1

CATARINA ELFFORS

MR/BY-S

SCB-Tryck, Örebro 1998.11