On Sampling with Probability Proportional to Size

Bengt Rosén

SCB

R&D Report
Statistics Sweden
Research - Methods - Development
1996:1

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 1996:1. On sampling with probability proportional to size / Bengt Rosén.
Digitaliserad av Statistiska centralbyrån (SCB) 2016.

urn:nbn:se:scb-1996-X101OP9601

On Sampling with Probability Proportional to Size

Bengt Rosén

SCB

R&D Report
Statistics Sweden
Research - Methods - Development
1996:1

Från trycket Maj 1996
Producent Statistiska centralbyrån, utvecklingsavdelningen
Ansvarig utgivare Lars Lyberg

Förfrågningar Bengt Rosén
tel 08-783 44 90
telefax 08-667 77 88

ISSN 0283-8680

Printed May, 1996
Producer Statistics Sweden
Publisher Lars Lyberg

Inquiries Bengt Rosén
telephone +46 8 783 44 90
telefax +46 8 667 77 88

ISSN 0283-8680

April 1996

On Sampling with Probability Proportional to Size

Bengt Rosén

Abstract. One of the techniques to exploit auxiliary information is to use a sampling scheme with inclusion probabilities proportional to given size measures, a πps scheme. The paper addresses the following πps problem: Exhibit a πps scheme with prescribed sample size, which leads to good estimation precision and has good variance estimation properties.

Rosén (1995) presented a new class of sampling schemes, called order sampling schemes, which are shown to provide interesting contributions to the πps problem. A notion "order sampling with fixed distribution shape" (OSFS) is introduced, and employed to construct a general class of πps schemes, OSFSπps schemes. A particular scheme, Pareto πps, is proved to be optimal among OSFSπps schemes, in the sense that it minimizes estimator variances.

Comparisons are made of three OSFSπps schemes and three other πps schemes; Sunter πps and systematic πps with frame ordered at random respectively by the sizes. The conclusions are as follows. Pareto πps is superior among πps schemes which admit objective assessment of sampling errors. Without this requirement, systematic πps with frame ordered by the sizes sometimes, but not always, leads to better estimator precision.

CONTENTS

On Sampling with Probability Proportional to Size.

1 Introduction

An efficient sample survey should utilize available auxiliary information, and a number of techniques for that purpose exist. One is to use sampling schemes with inclusion probabilities proportional to given sizes, πps schemes. The paper addresses the problem of exhibiting a πps scheme with prescribed sample size, which leads to good estimation precision and has good variance estimation properties. A novel πps scheme with favorable qualities is presented.

1.1 Notation and Basic Assumptions

$U = (1, 2, ..., N)$ denotes a finite population. A ***variable*** on it, $\mathbf{z} = (z_1, z_2.., z_N)$, associates a numerical value to each unit in U. A ***sampling frame***, in the guise of a ***list*** with records that one - to - one correspond with the units in the population, is available, and we let U refer to frame as well as population. We confine to ***without replacement*** (***wor***) probability sampling. Then the outcome of a sampling scheme is specified by the ***sample inclusion indicators*** $I_1, I_2, ..., I_N$ ($I_i = 1$ if unit i is sampled and $= 0$ otherwise). P, E, V and $\hat{V}$ denote probability, expectation, variance and variance estimator. First and second order ***inclusion probabilities*** are denoted by $\pi_i = P(I_i = 1)$ and $\pi_{ij} = P(I_i = I_j = 1)$, $i, j = 1, 2, ..., N$. The condition $\pi_i > 0$, $i = 1, 2, ..., N$, is presumed to be in force throughout. Moreover, we confine to sampling schemes with ***fixed sample size***, denoted n. Then;

$$\pi_1 + \pi_2 + ... + \pi_N = n. \tag{1.1}$$

A ***linear statistic*** is of the form stated in (1.2), where the ***weights*** $\mathbf{w} = (w_1, w_2, ..., w_N)$ are presumed to be known for all units in U while the ***study variable*** values $\mathbf{y} = (y_1, y_2, ..., y_N)$ are known only for observed units. Unless stated otherwise, it is presumed that non-response does not occur (i.e. that sampled and observed units agree).

$$L(\mathbf{y}; \mathbf{w}) = \sum_{i=1}^{N} y_i \cdot w_i \cdot I_i \, . \tag{1.2}$$

When $\pi_{ij} > 0$, $i, j = 1, 2, ..., N$, the corresponding ***Sen - Yates - Grundy variance estimator*** is;

$$\hat{V}[L(\mathbf{y}; \mathbf{w})] = \frac{1}{2} \cdot \sum_{i=1}^{N} \sum_{i=1}^{N} (y_i \cdot w_i - y_j \cdot w_j)^2 \cdot (\pi_i \cdot \pi_j / \pi_{ij} - 1) \cdot I_i \cdot I_j \, . \tag{1.3}$$

The central finite population inference task is estimation of a ***population total*** $\tau(\mathbf{y}) = \Sigma\, y_i$. The ***Horvitz - Thompson (HT) estimator*** for $\tau(\mathbf{y})$, denoted $\hat{\tau}(\mathbf{y})_{HT}$, is (1.2) with weights $w_i = 1/\pi_i$. It yields unbiased estimation of $\tau(\mathbf{y})$, and (1.3) specializes to a formula for $\hat{V}[\hat{\tau}(\mathbf{y})_{HT}]$.

Remark on notation: The printer which was used for this paper did unfortunately not allow for consequent use of bold face letters for variables. Latin letter variables are in bold face, while Greek letter variables are underlined.

1.2 The πps Problem

Let $\underline{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_N)$ be a set of desired inclusion probabilities for a size n wor sample, referred to as ***target inclusion probabilities***. By (1.1), the condition below must be met;

$$0 < \lambda_i \leq 1, \quad i = 1, 2, ..., N, \quad \text{and} \quad \lambda_1 + \lambda_2 + ... + \lambda_N = n. \tag{1.4}$$

The ***πps problem*** concerns, in its first round, exhibition of a size n wor sampling scheme with inclusion probabilities equal to, or at least close to, the target ones, i.e. such that;

$$\pi_i = (\text{or} \approx)\ \lambda_i, \ i = 1, 2, ..., N, \tag{1.5}$$

The problem is usually formulated slightly differently, which complies better with the abbreviation πps = inclusion probabilities $\underline{\pi}$ proportional to sizes. A ***size variable*** $\mathbf{s} = (s_1, s_2, ..., s_N)$,

$s_i > 0$, $i = 1, 2, ..., N$, is specified, and the task is to exhibit a sampling scheme with the π_i proportional to the s_i. This is equivalent to the above πps problem with $\lambda_i = n \cdot s_i / \sum_j s_j$, $i = 1, 2, ...,$ N. Note, though, that some size variables **s** lead to "impossible" problems, since one or more of the λ:s exceed 1. When met in practice, this obstacle is usually circumvented by introducing a "take all" stratum of units with large sizes.

The first round of the πps problem has many solutions. Therefore, in a second round the problem also concerns exhibition of a good, preferably even a best, scheme. To that end additional requirements have to be introduced, usually the following ones.

(i) The sample selection should be simple to implement.
(ii) The scheme should lead to good estimation precision.
(iii) The scheme should have good variance estimation properties. (1.6)

Requirement (i) needs no comment, (ii) only that estimation precision is, unless stated otherwise, understood to relate to the HT-estimator. Requirement (iii) is more intricate. The number one desire under (iii) is that the scheme should satisfy (1.7) below, i.e. be ***measurable*** in the terminology of Särndal et al. (1992). This condition is sufficient and essentially necessary for existence of a consistent variance estimator (for the HT-estimator), or in other words for admitting ***objective assessment of sampling errors.***

$$\pi_{ij} > 0, \quad i, j = 1, 2, ..., N. \tag{1.7}$$

For measurable sampling schemes, additional desires under (iii) are;

A variance estimation procedure should be exhibited, the simpler the better. (1.8)
Variance estimates should preferably never become negative. (1.9)

One way to meet (1.8) is to exhibit a formula/algorithm for computation of π_{ij}, $i, j = 1, 2, ..., N$. Then (1.3) (or some of its "relatives") can be employed. If also $\pi_i \cdot \pi_j - \pi_{ij} \geq 0$, $i, j = 1, 2, ..., N$, (1.9) will be satisfied for the variance estimator (1.3).

The most frequently employed πps scheme in practice, ***systematic πps***, does in general not satisfy (1.7). In fact, systematic πps is not <u>one</u> but a whole family of sampling schemes. In principle there is one scheme for each ordering of the sampling frame. Even if the HT-estimator does not depend on the frame order, its precision does. It is beneficial for estimation precision if the study variable values exhibit a "smooth trend" in frame order, and the more linear the trend is the better. Then two variance reducing forces are set to work; (a) Variation of inclusion probabilities (by πps), (b) An implicit homogeneity stratification of the population (by the frame ordering). With access to a size variable that is fairly proportional to the study variable, ***frame order by sizes*** leads to a fairly smooth trend. Other ordering principles are also used. In many survey contexts it is natural to use geographical order in the sampling frame, which leads to a desired geographical spread of the sample. Generally we speak of a sampling frame with ***fixed order***, one possibility being "ordered by size". As indicated, fixed frame order is often beneficial for estimation precision, ***but it violates* (1.7)**. Good point estimator precision is "bought" to the price that "control" over variance estimation is lost.

One way to cope with the dilemma, which is used quite widely in practice, is to order the frame at random. The sampling scheme comprises then first a totally random ordering of the frame and secondly selection of a systematic πps. This scheme, called ***systematic πps with random frame order***, is measurable. It is a bit unclear, though, how well it complies with (1.8) and (1.9). Approximate variance estimators exist, however, notably the Hartley-Rao estimator, which yields non-negative estimates. (See e.g. Section 7 in Wolter (1985).)

Sunter (1977) introduced a πps scheme, henceforth called ***Sunter πps***, with control of variance estimation. In particular, an algorithm for computation of π_{ij} was presented, and it was shown that $\pi_i \cdot \pi_j - \pi_{ij} > 0$ holds. It should be noted, though, that Sunter's scheme yields only approximate πps, in the sense that some of the original size measures have to be modified before the scheme can be set in operation, and πps refers to the modified sizes. A consequence of the modification is that a portion of the list is sampled by simple random sampling, and this portion is larger, the higher the sampling fraction is. Section 3.6.2 in Särndal et al. (1992) gives an account of the πps problem. In particular they present Sunter's scheme on pp. 93 - 96, to which the reader is referred.

1.3 Outline of the Paper

Our chief objective is to show that the ***order sampling schemes***, which are introduced and studied in Rosén (1995), can provide interesting contributions to the πps problem. To make the paper self-contained, basic concepts and results for order sampling are briefly reviewed in Section 2, where also some extensions are made. In particular the notion of ***order sampling with fixed distribution shape***, abbreviated OSFS, is introduced. In Section 3 this class of sampling schemes is employed to construct a class of (approximate) πps schemes, called OSFSπps schemes. Moreover, it is shown that a particular OSFSπps scheme, called ***Pareto πps***, is (asymptotically) optimal among the OSFSπps schemes, to the effect that it minimizes estimator variances. Section 4 contains a deepened study of Pareto πps.

The last Section 5 presents evaluations and comparisons of πps schemes, which comprise three OSFSπps schemes; ***Poisson πps, exponential πps*** and Pareto πps, and three πps schemes outside OSFSπps; Sunter πps and systematic πps, with frame ordered at random respectively by sizes. The main reason for including systematic πps is that it is widely used in practice. As regards Sunter πps, we read Särndal et al. (1992) so that they indicate Sunter πps to be a candidate for "best so far" among πps scheme with variance estimation control.

The conclusions from the evaluation are as follows. Pareto πps is superior among the considered πps schemes, relative to all three requests (i) - (iii) in (1.6). However, if (iii) is disregarded and one is prepared to loose control over variance estimation, systematic πps with frame ordered by sizes leads often, but not always, to better point estimator precision than Pareto πps. It is a matter of judgment/beliefs for the sampler to chose between the two schemes.

2 On Order Sampling

As a preparation for the subsequent approach to the πps problem we review, and also extend, some concepts and results on order sampling, presented in Rosén (1995).

2.1 Review of Basic Concepts and Results on Order Sampling

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

It is obvious that OS is wor. Moreover, since the F_i:s are assumed to be continuous, the realized Q - values contain (with probability 1) no ties. Hence, OS(n; F) has fixed sample size n.

The following Proposition 2.1, which in essence is a restatement of Approximation Result 3.2 in Rosén (1995), contains key estimation results for general order sampling schemes. The core results in Rosén (1995) are stringent limit theorems. However, from a sampling practical point of view limit results are not so interesting, since practical situations are always "finite". The main merit of limit theorems is that they lay ground for approximations which can be employed in practical, finite situations, and the approximation aspect is emphasized in the following. A reader with particular interest in precise limit theorem versions of the results to be formulated is referred to Rosén (1995) for additional information. We use the (admittedly sweeping) expression that a result holds "with good approximation under general conditions" to indicate that there is a corresponding limit theorem, which states that the claim is asymptotically true (under appropriate, but general, conditions).

We prepare the formulation of Proposition 2.1 by introducing a quantity ξ and the ***order sampling estimator*** $\hat{\tau}(\mathbf{y})_{OS}$ of $\tau(\mathbf{y})$, as stated below;

$$\xi \text{ solves the equation (in t, on } 0 \le t < \infty): \quad \sum_{i=1}^{N} F_i(t) = n, \tag{2.1}$$

$$\hat{\tau}(\mathbf{y})_{OS} = \sum_{i=1}^{N} \frac{y_i}{F_i(\xi)} \cdot I_i. \tag{2.2}$$

Remark 2.1: A sufficient condition for a solution ξ to (2.1) to be unique, is that $f_i(\xi) > 0$ for at least one i, $i=1,2,...,N$. See Remark 3.1 in Rosén (1995). #

Remark 2.2: As discussed in Section 3 in Rosén (1995) there are good grounds to believe that the following ***approximation formula for*** OS(n; **F**) ***inclusion probabilities*** π_i;

$$\pi_i \approx F_i(\xi), \quad i = 1,2,...,N, \tag{2.3}$$

holds with good approximation under general conditions. This is, however, only a conjecture, (2.3) has not been given a stringent justification in terms of a limit theorem.

With (2.3) as background, the estimator (2.2) can be viewed as a "quasi-HT-estimator", with approximate inclusion probabilities instead of the exact ones. #

The following Proposition 2.1 tells that (2.2) is a good estimator under general conditions. As already stated, Proposition 2.1 is in essence a restatement of Approximation Result 3.2 in Rosén (1995), to which the reader is referred for justifications. $N(\mu, \sigma^2)$ denotes the normal distribution with mean μ and variance σ^2.

PROPOSITION 2.1: Consider OS(n; **F**) sampling from the population U, on which the variable **y** is defined, and let notation be as in (2.1) and (2.2). Then a) - d) hold under general conditions;

a) The distribution of $\hat{\tau}(\mathbf{y})_{OS}$ is well approximated by $N(\tau(\mathbf{y}), \sigma^2)$, with (2.4)

$$\sigma^2 = \frac{N}{N-1} \cdot \sum_{i=1}^{N} \left(\frac{y_i}{F_i(\xi)} - \gamma \right)^2 \cdot F_i(\xi) \cdot (1 - F_i(\xi)), \quad \text{where} \tag{2.5}$$

$$\gamma = \sum_{i=1}^{N} \frac{y_i}{F_i(\xi)} \cdot f_i(\xi) \Big/ \sum_{i=1}^{N} f_i(\xi). \tag{2.6}$$

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

c) In particular, σ^2 yields good approximation of $V[\hat{\tau}(\mathbf{y})_{OS}]$.

d) A good variance estimator for $\hat{\tau}(\mathbf{y})_{OS}$ is provided by;

$$\hat{V}[\hat{\tau}(\mathbf{y})_{OS}] = \frac{n}{n-1} \cdot \sum_{i=1}^{N} \left(\frac{y_i}{F_i(\xi)} - \hat{\gamma} \right)^2 \cdot (1 - F_i(\xi)) \cdot I_i, \text{ where} \tag{2.7}$$

$$\hat{\gamma} = \sum_{i=1}^{N} \frac{y_i}{F_i(\xi)^2} \cdot f_i(\xi) \cdot I_i \Big/ \sum_{i=1}^{N} \frac{f_i(\xi)}{F_i(\xi)} \cdot I_i. \tag{2.8}$$

Remark 2.3: There is a slight discrepancy in formulation between the above result and Approximation Result 3.2 in Rosén (1995), to the effect that the factor N/(N-1) is inserted in

(2.5). This is an ad hoc adjustment in order to make (2.5) to the exactly correct formula in the case when all F_i are equal, i.e. when the OS scheme is nothing but simple random sampling. Note that the factor asymptotically equals 1. #

Remark 2.4: As a continuation of Remark 2.2 we stress the following. Justification of Approximation Result 3.2 in Rosén (1995) is given by Theorem 3.1 in Rosén (1995). This theorem does not rely on the approximation (2.3). As a consequence, the claims in Proposition 2.1 hold asymptotically under the general conditions in Theorem 3.1 in Rosén (1995), irrespective of whether (2.3) is a good approximation or not. #

2.2 A Uniqueness Issue for Order Sampling

Definition 2.1 states that a pair $(n;\mathbf{F})$, $\mathbf{F}=(F_1,F_2,...,F_N)$, determines an order sampling scheme. However, there is not one - to - one correspondence between OS schemes and pairs $(n;\mathbf{F})$. To realize that, consider a function φ on $[0,\infty)$ which is strictly increasing, continuous and has range $[0,\infty)$. Since an increasing function preserves order, the following holds.

Two OS schemes with ranking variables $Q_1,Q_2,...,Q_N$ and $Q^*_1,Q^*_2,...,Q^*_N$, which are related by $Q^*_i=\varphi(Q_i)$, $i=1,2,...,N$, are (probabilistically) equivalent. (2.9)

The corresponding order distributions $\mathbf{F}=(F_1,F_2,...,F_N)$ and $\mathbf{F}^*=(F_1^*,F_2^*,...,F_N^*)$ are related by $F_i^*(t)=F_i(\varphi^{-1}(t))$, $i=1,2,...,N$, where φ^{-1} denotes inverse function. Hence, (2.9) can be expressed as follows, which shows that there is *not* one - to - one correspondence between OS schemes and pairs $(n;\mathbf{F})$.

For a continuous, strictly increasing φ, the schemes $OS(n;\mathbf{F})$ and $OS(n;\mathbf{F}(\varphi^{-1})$ agree. (2.10)

2.3 Order Sampling with Fixed Order Distribution Shape

Here we introduce a sub - class of OS schemes which will play an instrumental role in the subsequent considerations of the πps problem.

DEFINITION 2.2: Let $H(t)$ be a probability distribution with density $h(t)$, $t\geq 0$, and let $\underline{\theta}=(\theta_1,\theta_2,..,\theta_N)$, $\theta_i>0$, be real numbers. An $OS(n;\mathbf{F})$ scheme, $\mathbf{F}=(F_1,F_2,...,F_N)$, is said to have ***fixed order distribution shape*** $H(t)$ and ***intensities*** $\underline{\theta}$, if either, and hence both, of the following two equivalent conditions are met.

(i) The ranking variables $Q_1\ Q_2\ ...,Q_N$ are of type $Q_i=Z_i/\theta_i$, $i=1,2,...,N$, where $Z_1,Z_2,...,Z_N$ are independent, identically distributed (iid) random variables with common distribution H. (2.11)

(ii) The order distributions are $F_i(t)=H(t\cdot\theta_i)$, $0\leq t<\infty$, $i=1,2,...,N$. (2.12)

Such a sampling scheme is referred to by the notation **OSFS**$(n;H;\underline{\theta})$.

The following particular OSFS schemes are given special attention in Rosén (1995). As regards sequential Poisson sampling, we also refer to Ohlsson (1995).

DEFINITION 2.3: a) ***Sequential Poisson sampling*** is $OSFS(n;H;\underline{\theta})$ with H = the ***standard uniform distribution***; $H(t)=\min(t,1)$, $0\leq t<\infty$, and $h(t)=1$ on $[0,1]$ and 0 outside $[0,1]$.

b) ***Successive sampling*** is $OSFS(n;H;\underline{\theta})$ with H = the ***standard exponential distribution***; $H(t)=1-e^{-t}$ and $h(t)=e^{-t}$, $0\leq t<\infty$.

3 OSFS Schemes and the πps Problem

Here we employ OSFS schemes to treat the πps problem in Section 1.2. In the first round we presume that the shape distribution is given, and we seek intensities which lead to inclusion probabilities that are equal to, or at least close to, the target inclusion probabilities. In a second round, we regard the shape distribution as optional, and consider optimal choice of it.

3.1 The πps Problem when the Shape Distribution is Given

Let $\underline{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_N)$ and H be given. We consider the following problem.

Problem: Find intensities $\underline{\theta}$ such that the inclusion probabilities $\underline{\pi} = (\pi_1, \pi_2, ..., \pi_N)$ for the scheme OSFS(n; H; $\underline{\theta}$) agree with the target probabilities $\underline{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_N)$.

We start with a heuristic treatment of the problem, in which the approximation (2.3) is regarded as an exact relation. Then, by combining (2.3) and (2.12) we obtain the relations: $\pi_i = H(\xi \cdot \theta_i)$, $i = 1, 2, ..., N$. Thus, the desired equality $\pi_i = \lambda_i$ holds if intensities θ_i and target values λ_i are related as follows: $\lambda_i = H(\xi \cdot \theta_i)$, $i = 1, 2, ..., N$. By solving for θ_i we get, where H^{-1} denotes inverse function: The intensities $\theta_i = H^{-1}(\lambda_i)/\xi$, $i = 1, 2, ..., N$, provide a solution to the above problem. Moreover, an OSFS scheme is invariant under re - scaling of the intensities (i.e. by multiplying them with the same positive number). This is seen from (2.10) with $\varphi(t) = c \cdot t$, $c > 0$. Hence, an alternative solution to the problem is $\theta_i = H^{-1}(\lambda_i)$, $i = 1, 2, ..., N$.

We use the last formula to construct ranking variables Q_i for an OSFS scheme. By (2.11), $Q_i = Z_i/\theta_i$, $i = 1, 2, ..., N$, where $Z_1, Z_2, ..., Z_N$ are iid random variables with common distribution H. Note that such Z-variables can be generated as;

$$Z_i = H^{-1}(U_i), \quad i = 1, 2, ..., N, \quad \text{where} \tag{3.1}$$

$U_1, U_2, ..., U_N$ are independent standard uniform random variables (see Def. 2.3). (3.2)

Thereby we have arrived at the following ***heuristic solution***: The OS scheme with ranking variables $Q_i = H^{-1}(U_i)/H^{-1}(\lambda_i)$, $i = 1, 2, ..., N$, (which is an OSFS scheme) has inclusion probabilities $\underline{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_N)$. Against this background we make the following formal definition.

DEFINITION 3.1: The ***OSFSπps scheme*** with sample size n, target inclusion probabilities $\underline{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_N)$ and shape distribution H, in full denoted OSFSπps(n; $\underline{\lambda}$; H), is the OS scheme which satisfies either, and hence both, of the following two equivalent conditions. In particular, the conditions imply that the scheme is an OSFS scheme.

(i) Ranking variables $Q_1, Q_2, ..., Q_N$ are;

$$Q_i = \frac{H^{-1}(U_i)}{H^{-1}(\lambda_i)}, \quad i = 1, 2, ..., N, \quad \text{with U:s as in (3.2).} \tag{3.3}$$

(ii) Order distributions $\mathbf{F} = (F_1, F_2, ..., F_N)$ are;

$$F_i(t) = H(t \cdot H^{-1}(\lambda_i)), \quad \text{with density } f_i(t) = h(t \cdot H^{-1}(\lambda_i)) \cdot H^{-1}(\lambda_i), \quad i = 1, 2, ..., N. \tag{3.4}$$

Proof of the equivalence between (i) and (ii): With ranking variables as in (3.3), we have;

$$F_i(t) = P(Q_i \leq t) = P(H^{-1}(U_i)/H^{-1}(\lambda_i) \leq t) = P(U_i \leq H(t \cdot H^{-1}(\lambda_i)) = H(t \cdot H^{-1}(\lambda_i)), \quad i = 1, 2, ..., N,$$

which is the formula for $F_i(t)$ in (3.4). The density $f_i(t)$ is obtained by differentiation. #

Since OSFSπps is a special type of OS scheme, the results in Proposition 2.1 can, and will, be applied. We start by exhibiting expressions for the ξ, $F(\xi)$ and $f(\xi)$ which enter in Proposition 2.1. By (3.4), equation (2.1) here takes the form;

$$\sum_{i=1}^{N} H(t \cdot H^{-1}(\lambda_i)) = n\,. \tag{3.5}$$

By virtue of (1.4), (3.5) is solved by $\xi = 1$. This together with (3.4) yields;

$$F_i(\xi) = F_i(1) = H(H^{-1}(\lambda_i)) = \lambda_i, \quad \text{and} \quad f_i(\xi) = f_i(1) = h(H^{-1}(\lambda_i)) \cdot H^{\prime -1}(\lambda_i), \quad i = 1,2,...,N. \tag{3.6}$$

Now, by applying Proposition 2.1 with $F_i(\xi) = \lambda_i$, and $f_i(\xi) = h(H^{-1}(\lambda_i)) \cdot H^{\prime -1}(\lambda_i)$, the following Approximation Result 3.1 is obtained. Some algebra is left to the reader.

APPROXIMATION RESULT 3.1: Consider OSFSπps(n; $\underline{\lambda}$; H) sampling from the population U, on which the variable **y** is defined. Define the estimator $\hat{\tau}(\mathbf{y})_\lambda$ of $\tau(\mathbf{y})$ by;

$$\hat{\tau}(\mathbf{y})_\lambda = \sum_{i=1}^{N} \frac{y_i}{\lambda_i} \cdot I_i\,. \tag{3.7}$$

Then a) - d) below hold under general conditions

a) The distribution of $\hat{\tau}(\mathbf{y})_\lambda$ is well approximated by $N(\tau(\mathbf{y}), \sigma^2(\mathbf{y}; H; \underline{\lambda}))$, with (3.8)

$$\sigma^2(\mathbf{y}; \underline{\lambda}; H) = \frac{N}{N-1} \cdot \sum_{i=1}^{N} \left(\frac{y_i}{\lambda_i} - \sum_{j=1}^{N} \frac{y_j \cdot \alpha_j}{\lambda_j} \Big/ \sum_{j=1}^{N} \alpha_j \right)^2 \cdot \lambda_i \cdot (1 - \lambda_i), \quad \text{where} \tag{3.9}$$

$$\alpha_i = \alpha_i(\underline{\lambda}; H) = h(H^{-1}(\lambda_i)) \cdot H^{\prime -1}(\lambda_i), \quad i = 1,2,...,N. \tag{3.10}$$

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

c) In particular, $V[\hat{\tau}(\mathbf{y})_\lambda]$ is well approximated by $\sigma^2(\mathbf{y}; \underline{\lambda}; H)$.

d) A good variance estimator for $\hat{\tau}(\mathbf{y})_\lambda$ is provided by;

$$\hat{V}[\hat{\tau}(\mathbf{y})_\lambda] = \frac{n}{n-1} \cdot \sum_{i=1}^{N} \left(\frac{y_i}{\lambda_i} - \sum_{j=1}^{N} \frac{y_j \cdot a_j}{\lambda_j} \cdot I_j \Big/ \sum_{j=1}^{N} a_j \cdot I_j \right)^2 \cdot (1 - \lambda_i) \cdot I_i\,, \quad \text{where} \tag{3.11}$$

$$a_i = a_i(\underline{\lambda}; H) = \alpha_i(\underline{\lambda}; H)/\lambda_i = h(H^{-1}(\lambda_i)) \cdot H^{\prime -1}(\lambda_i)/\lambda_i\,, \quad i = 1,2,...,N. \tag{3.12}$$

Remark 3.1: In the heuristic reasoning which precedes Definition 3.1, the approximation (2.3) plays a crucial role. However, as stated in Remark 2.2, even if (2.3) is conjectured to be a good approximation, its goodness is an open question. We want to stress, though, (cf. Remark 2.4) that justification of the above Approximation Result 3.1 is given by Theorem 3.1 in Rosén (1995), and this theorem does not rely on the approximation (2.3). Hence, the claims in Approximation Result 3.1 hold asymptotically under the general conditions in Theorem 3.1 in Rosén (1995), whether (2.3) is a good approximation or not. #

Next we introduce terms and notation for the OSFSπps schemes which relate to the sampling procedures in Definition 2.3. Our general terminological rule is that OSFSπps(n; $\underline{\lambda}$; H) is named by its shape distribution H. We start, though, with an exception from the rule. ***Poisson πps***, denoted ***POIπps***(n; $\underline{\lambda}$) or just ***POIπps***, and ***exponential πps***, denoted ***EXPπps***(n; $\underline{\lambda}$) or just ***EXPπps***, are the OSFSπps(n; $\underline{\lambda}$; H) schemes with H = standard uniform distribution respectively H = standard exponential distribution (see Definition 2.3).

Remark 3.2: By the above terminology rule, Poisson πps ought to be named "uniform πps". A reason for diverging is that H = the standard uniform distribution (UNIF) has the special property $H(t) = H^{-1}(t)$ (on its support [0,1]). Hence, OSFSπps(n; $\underline{\lambda}$; UNIF) = OSFS(n; UNIF; $\underline{\lambda}$). Since Ohlsson (1995) calls OSFS(n; UNIF; $\underline{\lambda}$) "sequential Poisson sampling", we keep Poisson, but leave out sequential for brevity throughout this paper. We are aware, though, that confusion may arise in more general contexts, since "Poisson sampling" is already booked for another scheme, see e.g. Särndal et al. (1992), pp. 85-87. #

Verification of the following lemma is left to the reader.

LEMMA 3.1: a) For POIπps(n; $\underline{\lambda}$), $\underline{\alpha}$, **a**, Q:s and σ^2 in (3.10), (3.12), (3.3) and (3.9) are;

$$\alpha_i = \lambda_i, \quad a_i = 1, \quad Q_i = U_i/\lambda_i, \quad i = 1,2,...,N, \tag{3.13}$$

$$\sigma^2(\mathbf{y};\underline{\lambda};\mathrm{POI}) = \frac{N}{N-1} \cdot \sum_{i=1}^{N} \left(\frac{y_i}{\lambda_i} - \frac{1}{n} \cdot \sum_{j=1}^{N} y_j \right)^2 \cdot \lambda_i \cdot (1-\lambda_i). \tag{3.14}$$

b) For EXPπps(n; λ), $\underline{\alpha}$, **a**, Q:s and σ^2 in (3.10), (3.12), (3.3) and (3.9) are;

$$\alpha_i = -\ln(1-\lambda_i)\cdot(1-\lambda_i), \; a_i = \frac{-\ln(1-\lambda_i)\cdot(1-\lambda_i)}{\lambda_i}, \; Q_i = \frac{\ln(1-U_i)}{\ln(1-\lambda_i)}, \; i = 1,...,N, \tag{3.15}$$

$$\sigma^2(\mathbf{y};\underline{\lambda};\mathrm{EXP}) = \frac{N}{N-1} \cdot \sum_{i=1}^{N} \left(\frac{y_i}{\lambda_i} - \frac{\sum_{j=1}^{N} y_j \cdot \ln(1-\lambda_j)\cdot(1-\lambda_j)/\lambda_j}{\sum_{j=1}^{N} \ln(1-\lambda_j)\cdot(1-\lambda_{ji})} \right)^2 \cdot \lambda_i \cdot (1-\lambda_i). \tag{3.16}$$

Remark 3.3: The associated variance estimators are obtained by inserting the a:s in (3.11). #

3.2 Optimal Shape Distribution

The point estimator $\hat{\tau}(\mathbf{y})_\lambda$ in (3.7) is the same for all shape distributions H. The variance of $\hat{\tau}(\mathbf{y})_\lambda$ depends on H, though, as is seen from (3.9) and (3.10), and exemplified in (3.14) and (3.16). Therefore it is natural to wonder about H:s which minimize the asymptotic variance $\sigma^2(\mathbf{x};\underline{\lambda};\mathrm{H})$ in (3.9). The answer is stated below.

THEOREM 3.1: The asymptotic variance $\sigma^2(\mathbf{y};\underline{\lambda};\mathrm{H})$ in (3.9) is minimized, for any study variable **y** and any $\underline{\lambda}$, by the shape distribution;

$$H(t) = \frac{t}{1+t}, \; 0 \le t < \infty, \quad \text{with density} \quad h(t) = \frac{1}{(1+t)^2}, \; 0 \le t < \infty. \tag{3.17}$$

DEFINITION 3.2: The distribution in (3.17) is called the ***standard Pareto distribution***, and the associated OSFSπps scheme ***Pareto πps***, denoted ***PARπps***(n; $\underline{\lambda}$) or just ***PARπps***.

Remark 3.4: The fact that Pareto πps minimizes $\sigma^2(\mathbf{y};\underline{\lambda};\mathrm{H})$ for any **y** and $\underline{\lambda}$ is referred to by saying that Pareto πps is (asymptotically) ***uniformly optimal*** among OSFSπps schemes. #

Proof of Theorem 3.1: The algebraic details in the following proof are left to the reader. We shall employ the following representation of $\sigma^2(\mathbf{y};\underline{\lambda};\mathrm{H})$ in (3.9);

$$\sigma^2(\mathbf{y};\underline{\lambda};\mathrm{H}) = M(\mathbf{y};\underline{\lambda}) + R(\mathbf{y};\underline{\lambda};\mathrm{H}), \quad \text{where,} \tag{3.18}$$

$$M(\mathbf{y};\underline{\lambda}) = \frac{N}{N-1} \cdot \sum_{i=1}^{N} \left(\frac{y_i}{\lambda_i} - \sum_{j=1}^{N} y_j \cdot (1-\lambda_j) \Big/ \sum_{j=1}^{N} \lambda_j \cdot (1-\lambda_j) \right)^2 \cdot \lambda_i \cdot (1-\lambda_i), \tag{3.19}$$

$$R(\mathbf{y};\underline{\lambda};\mathrm{H}) = \frac{N}{N-1} \cdot \left(\sum_{i=1}^{N} \frac{y_i}{\lambda_i} \cdot \left[\frac{\alpha_i}{\sum_{j=1}^{N} \alpha_j} - \frac{\lambda_i \cdot (1-\lambda_i)}{\sum_{j=1}^{N} \lambda_j \cdot (1-\lambda_j)} \right] \right)^2 \cdot \sum_{i=1}^{N} \lambda_i \cdot (1-\lambda_i). \tag{3.20}$$

The representation (3.18) - (3.20) is obtained by using the following identity;

$$\sum_{i=1}^{N}(z_i-\gamma)^2\cdot\beta_i=\sum_{i=1}^{N}\left(z_i-\sum_{j=1}^{N}z_j\cdot\beta_j\Big/\sum_{j=1}^{N}\beta_j\right)^2\cdot\beta_i+\left(\gamma-\sum_{i=1}^{N}z_i\cdot\beta_i\Big/\sum_{i=1}^{N}\beta_i\right)^2\cdot\sum_{i=1}^{N}\beta_i, \quad (3.21)$$

with $z_i=y_i/\lambda_i$, $\gamma=\sum_{i=1}^{N}\frac{y_i\cdot\alpha_i}{\lambda_i}\Big/\sum_{i=1}^{N}\alpha_i$ and $\beta_i=\lambda_i\cdot(1-\lambda_i)\cdot N/(N-1)$, and comparing with (3.9).

Now, the following claims (i) - (iii) are straightforward consequences of (3.19) and (3.20); (i) $M(\mathbf{y};\underline{\lambda})$ does not depend on H, (ii) $R(\mathbf{y};\underline{\lambda};H)\geq 0$ for all H, (iii) $R(\mathbf{y};\underline{\lambda};H)=0$ if α_i is proportional to $\lambda_i\cdot(1-\lambda_i)$. By combining (3.18), (i) - (iii) and c) in Approximation Result 3.1, the following result is obtained.

LEMMA 3.2: For a fixed sample size n, OSFSπps(n; $\underline{\lambda}$; H) has, for any **y**, minimal asymptotic variance $\sigma^2(\mathbf{y};\underline{\lambda};H)$ for $\hat{\tau}(y_\lambda)$ among OSFSπps(n; $\underline{\lambda}$) schemes if;

$\alpha_i(\underline{\lambda};H)$ in (3.10) is proportional to $\lambda_i\cdot(1-\lambda_i)$. (3.22)

Next we search an H which makes (3.22) satisfied. Relations (3.22) and (3.10) lead to the following functional - differential equation, k being a positive constant;

$$h(H^{-1}(\lambda))\cdot H^{-1}(\lambda)=k\cdot\lambda\cdot(1-\lambda),\ 0<\lambda<1,\quad \text{where } h(t)=H'(t)\ (\text{' for derivative}). \quad (3.23)$$

By inverting (3.23) (with respect to multiplication) and using the general differentiation rule $dy^{-1}(\lambda)/d\lambda=1/y'(y^{-1}(\lambda))$, the following differential equation is obtained;

$$\frac{dH^{-1}(\lambda)/d\lambda}{H^{-1}(\lambda)}=\frac{d[\ln H^{-1}(\lambda)]}{d\lambda}=(\text{by }(3.23))=\frac{1}{k\cdot\lambda\cdot(1-\lambda)}=\frac{1}{k}\cdot\frac{d}{d\lambda}[\ln\frac{\lambda}{1-\lambda}]. \quad (3.24)$$

It is readily deduced that the probability distributions H which satisfy (3.24) are specified by;

$$H^{-1}(\lambda;k;c)=c\cdot\left(\frac{\lambda}{1-\lambda}\right)^{1/k},\ \text{implying}\ H(t;k;c)=\frac{c\cdot t^k}{1+c\cdot t^k},\ 0\leq t<\infty,\ k,c>0. \quad (3.25)$$

In the first round (3.25) seems to provides quite an extensive class of solutions to (3.22), which is not the case, though. From (3.3) and (3.25) it is seen that the associated ranking variables Q(k; c) are;

$$Q(k;c)_i=\frac{H^{-1}(U_i;k;c)}{H^{-1}(\lambda_i;k;c)}=\left(\frac{U_i\cdot(1-\lambda_i)}{\lambda_i\cdot(1-U_i)}\right)^{1/k},\quad i=1,2,\ldots,N. \quad (3.26)$$

Since the function $\varphi(t)=t^{1/k}$, $0\leq t<\infty$, $k>0$, has range $[0,\infty)$, is increasing and continuous, (3.26) and (2.9) yield that all OSFSπps(n;$\underline{\lambda}$; H) schemes specified by (3.26) are equivalent. The "standard" version with $c=k=1$ is chosen in Theorem 3.1 and Definition 3.2. #

Remark 3.5: (3.23) - (3.26) yield a constructive derivation of Pareto πps as solution to (3.22). If the reader should question the rigor in that derivation, argue instead as follows. Formula (4.1) tells that PAR(n; $\underline{\lambda}$) satisfies (3.22). Once a shape distribution H that makes (3.22) satisfied is exhibited, Lemma 3.2 "directly", and rigorously (but non - constructively), tells that no other OSFSπps scheme has smaller asymptotic variance $\sigma^2(\mathbf{y};\underline{\lambda};H)$. #

4 Pareto πps

The (asymptotic) optimality property stated in Theorem 3.1 gives Pareto πps (see Definition 3.2) particular interest. In this section we provide further information on Pareto πps.

4.1 Properties of Pareto πps

The algebra relating to (4.1) - (4.3) is left to the reader. Recall relation (1.4).

LEMMA 4.1: For PARπps(n; $\underline{\lambda}$), $\underline{\alpha}$, **a**, Q:s and σ^2 in (3.10), (3.12), (3.3) and (3.9) are;

$$\alpha_i = \lambda_i \cdot (1-\lambda_i), \quad a_i = 1-\lambda_i, \quad Q_i = \frac{U_i \cdot (1-\lambda_i)}{\lambda_i \cdot (1-U_i)}, \quad i = 1,2,\ldots,N, \tag{4.1}$$

$$\sigma^2(\mathbf{y};\underline{\lambda};\mathrm{PAR}) = \frac{N}{N-1} \cdot \sum_{i=1}^{N} \left(\frac{y_i}{\lambda_i} - \frac{\sum_{j=1}^{N} y_j \cdot (1-\lambda_j)}{\sum_{j=1}^{N} \lambda_j \cdot (1-\lambda_j)} \right)^2 \cdot \lambda_i \cdot (1-\lambda_i) =$$

$$= \frac{N}{N-1} \cdot \{ \sum_{i=1}^{N} \frac{y_i^2}{\lambda_i} \cdot (1-\lambda_i) - \frac{[\sum_{i=1}^{N} y_i \cdot (1-\lambda_i)]^2}{n - \sum_{i=1}^{N} \lambda_i^2} \}. \tag{4.2}$$

The corresponding variance estimator (3.11) is

$$\hat{V}[\hat{\tau}(\mathbf{y})_\lambda] = \frac{N}{N-1} \cdot \sum_{i=1}^{N} \left(\frac{y_i}{\lambda_i} - \frac{\sum_{j=1}^{N} y_j \cdot (1-\lambda_j)/\lambda_j}{\sum_{j=1}^{N} (1-\lambda_j)} \right)^2 \cdot (1-\lambda_i). \tag{4.3}$$

Remark 4.1: By the Lagrange equality version of Schwarz inequality, (4.2) can be re-written;

$$\sigma^2(\mathbf{y};\underline{\lambda};\mathrm{PAR}) = \frac{N}{N-1} \cdot \sum_{j=1}^{N} \sum_{j=1}^{N} \lambda_i \cdot \lambda_j \cdot (1-\lambda_i) \cdot (1-\lambda_j) \cdot \left(\frac{y_i}{\lambda_i} - \frac{y_j}{\lambda_j} \right)^2 \Big/ (n - \sum_{j=1}^{N} \lambda_i^2). \tag{4.4}$$

Remark 4.2: The uniform asymptotic optimality of PARπps (see Remark 3.3) for estimation of population totals $\tau(\mathbf{y})$ entails that PARπps is asymptotically optimal also for estimation of other population characteristics. In the discussion of this matter we use the following conventions. Arithmetic operations on variables are to be interpreted component-wise, e.g. for $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_N)$, $\mathbf{x} \cdot \mathbf{y} = (x_1 \cdot y_1, \ldots, x_N \cdot y_N)$. Moreover, optimal stands for "minimizing estimator variance".

Consider estimation of a ***ratio*** $\mu(\mathbf{y};\mathbf{x}) = \tau(\mathbf{y})/\tau(\mathbf{x})$. By standard principles, $\mu(\mathbf{y};\mathbf{x})$ is estimated by $\hat{\mu}(\mathbf{y};\mathbf{x})_\lambda = \hat{\tau}(\mathbf{y})_\lambda / \hat{\tau}(\mathbf{x})_\lambda$. Taylor linearization leads to the asymptotic variance formula $V[\hat{\mu}(\mathbf{y};\mathbf{x})_\lambda] \approx V[\hat{\tau}(\mathbf{y} - \mu(\mathbf{y};\mathbf{x}) \cdot \mathbf{x})_\lambda]/\tau(\mathbf{x})^2$. Hence, minimization of $V[\hat{\mu}(\mathbf{y};\mathbf{x})_\lambda]$ is asymptotically equivalent to minimization of $V[\hat{\tau}(\mathbf{y} - \mu(\mathbf{y};\mathbf{x}) \cdot \mathbf{x})_\lambda]$, i.e. minimization of the variance for the estimator of the total of $\mathbf{x} - \mu(\mathbf{y};\mathbf{x}) \cdot \mathbf{y}$. Since, by Theorem 3.1, Pareto πps effectuates minimization uniformly (irrespective of the study variable), it is asymptotically optimal also for estimation of ratios.

The optimality extends also to estimation of ***domain characteristics*** (totals and ratios). To illustrate, we consider estimation of a domain ratio $\mu(\mathbf{y};\mathbf{x};D) = \tau(\mathbf{y};D)/\tau(\mathbf{x};D)$, where $\tau(\mathbf{y};D)$ denotes the $\mathbf{y}$-total over domain D, which can be written $\tau(\mathbf{y};D) = \tau(\mathbf{y} \cdot \mathbf{1}_D)$, where $\mathbf{1}_D$ is the domain D indicator. The previous reasoning leads to $\hat{\mu}(\mathbf{y};\mathbf{x};D)_\lambda = \hat{\tau}(\mathbf{y};D)_\lambda / \hat{\tau}(\mathbf{x};D)_\lambda$ with asymptotic variance $V[\hat{\mu}(\mathbf{y};\mathbf{x};D)_\lambda] \approx V[\hat{\tau}([\mathbf{y} - \mu(\mathbf{x};\mathbf{y}) \cdot \mathbf{x}] \cdot \mathbf{1}_D)_\lambda]/\tau(\mathbf{x} \cdot \mathbf{1}_D)^2$. Theorem 3.1 tells again that PARπps is asymptotically optimal. #

4.2 A Limit Theorem

THEOREM 4.1: Consider, for $k=1,2,3,...$, PARπps($n_k;\underline{\lambda}_k$) sampling from the population $U_k=(1,2,...,N_k)$, on which the variable $\mathbf{y}_k=(y_{k1},y_{k2},...,y_{kN_k})$ is defined, with total $\tau(\mathbf{y}_k)$. Let $\hat{\tau}(\mathbf{y}_k)_\lambda$ be in accordance with (3.7). Then, under conditions (C1) - (C3) below, (4.5) holds for σ_k defined in accordance with (4.2).

$$[\hat{\tau}(\mathbf{y}_k)_\lambda-\tau(\mathbf{y}_k)]/\sigma_k \text{ converges in distribution to } N(0,1), \text{ as } k\to\infty. \tag{4.5}$$

(C1) $n_k\to\infty$, as $k\to\infty$.

(C2) $\overline{\lim_{k\to\infty}}\frac{1}{n_k}\cdot\sum_{i=1}^{N_k}\lambda_{ki}^2<1$.

(C3) $\max_i\left|\frac{y_{ki}}{\lambda_{ki}}-\frac{\sum_{j=1}^{N_k}y_{kj}\cdot(1-\lambda_{kj})}{\sum_{j=1}^{N_k}\lambda_{kj}\cdot(1-\lambda_{kj})}\right|\to 0$, as $k\to\infty$.

Remark 4.3: Loose interpretations of (C2) and (C3) run as follows. In order that normal approximation of the distribution of $\hat{\tau}(\mathbf{y}_k)_\lambda$ shall be good, a larger n_k is required (i) the closer to 1 some target inclusion probabilities lie, (ii) the more "dispersed" the sampling situation is, where "dispersed" relates to the of spread of $\{y_{ki}/\lambda_{ki};\ i=1,2,...,N\}$. #

Proof of Theorem 4.1: If we show that conditions (C1) - (C3) imply fulfillment of conditions (B1) - (B5) in Theorem 3.1 in Rosén (1995), the result follows from that theorem. First note that (B1) and (B2) in fact are the same as (C1) and (C3). To verify (B3), note that (3.6) yields;

$$f_{ki}(\xi_k)=\lambda_{ki}\cdot(1-\lambda_{ki}),\ i=1,2,...,N_k. \tag{4.6}$$

Condition (B3) concerns the asymptotic behavior of the quantity;

$$\frac{1}{n_k}\cdot\sum_{i=1}^{N_k}f(\xi_{ki})=\text{ (by (4.6)) }=\frac{1}{n_k}\cdot\sum_{i=1}^{N_k}\lambda_{ki}\cdot(1-\lambda_{ki})=\text{ (by (1.4)) }=1-\frac{1}{n_k}\cdot\sum_{i=1}^{N_k}\lambda_{ki}^2. \tag{4.7}$$

From (4.7) it is readily seen that (C2) implies (B3). The verifications of (B4) and (B5) which are straightforward, are left to the reader. #

4.3 Operational formulation of the Pareto πps sampling - estimation scheme

Below we list the steps in practical application of the PARπps sampling - estimation procedure. Justifications are given by the results presented so far.

The Pareto πps sampling - estimation procedure

1. A ***sampling frame*** $U=(1,2,...,N)$ with ***sizes*** $\mathbf{s}=(s_1,s_2,...,s_N)$, $s_i>0$, is at hand, and a ***sample size*** n is specified.
2. Compute the ***target inclusion probabilities*** $\lambda_i=n\cdot s_i/\sum_{j=1}^{N}s_j$, $i=1,2,...,N$.

   It is presumed that $\lambda_i<1$, $i=1,2,...,N$. If not, introduce a take all stratum or modify the size measures.
3. Realize independent standard uniform random variables $U_1,U_2,...,U_N$, thereby realizing the ***ranking variables***

   $$Q_i=\frac{U_i\cdot(1-\lambda_i)}{\lambda_i\cdot(1-U_i)},\quad i=1,2,...,N.$$

The *sample* consists of the n units with labels $(J_1, J_2, ..., J_n)$ that are determined by $Q_{J_1}, Q_{J_2}, ..., Q_{J_N}$ being the *n smallest* values among the realized $Q_1, Q_2, ..., Q_N$.

Variable values y and $\underline{\lambda}$ ***for sampled units*** are denoted (Y_ν, Λ_ν), $\nu = 1, 2, ..., n$.

4. The ***population total*** $\tau(y) = y_1 + y_2 + ... + y_N$, is ***estimated*** by $\hat{\tau}(y)_\lambda = \sum_{\nu=1}^{n} \frac{Y_\nu}{\Lambda_\nu}$. (4.8)

5. The ***variance*** of $\hat{\tau}(y)_\lambda$ is ***estimated*** by;

$$\hat{V}[\hat{\tau}(y)_\lambda] = \frac{n}{n-1} \cdot \sum_{\nu=1}^{n} \left(\frac{Y_\nu}{\Lambda_\nu} - \frac{\sum_{\rho=1}^{n} Y_\rho \cdot (1-\Lambda_\rho)/\Lambda_\rho}{\sum_{\rho=1}^{n} (1-\Lambda_\rho)} \right)^2 \cdot (1 - \Lambda_\nu). \quad (4.9)$$

6. An approximately 95% ***confidence interval*** for $\tau(y)$ is;

$$\hat{\tau}(y)_\lambda \pm 1.96 \cdot \sqrt{\hat{V}[\hat{\tau}(y)_\lambda]}. \quad (4.10)$$

Remark 4.4: Below we state a version of (4.9), which has computational merits;

$$\hat{V}[\hat{\tau}(y)_\lambda] = \frac{n}{n-1} \cdot (A - B^2/C), \quad (4.11)$$

where

$$A = \sum_{\nu=1}^{n} \left(\frac{Y_\nu}{\Lambda_\nu} \right)^2 \cdot (1 - \Lambda_\nu), \quad B = \sum_{\nu=1}^{n} \frac{Y_\nu}{\Lambda_\nu} \cdot (1 - \Lambda_\nu), \quad C = \sum_{\nu=1}^{n} (1 - \Lambda_\nu). \quad (4.12)$$

Remark 4.5: As indicated in Remark 4.2, point and variance estimators for other population characteristics than totals can be obtained by appropriate modifications of Steps 4 - 6. #

Remark 4.6: The general structure in the above formulation of the Pareto πps sampling - estimation procedure applies to all OSFSπps schemes. Differences between specific schemes concern only the definition of the ranking variables Q (in Step 3), and the **a** - values in (3.12), which in turn affect the variance estimator (3.11), which for Pareto πps takes the form (4.9). #

Remark 4.7: From (4.9) it is seen that variance estimates for Pareto πps always are non-negative. More generally, it is seen from (3.11) that this holds for any OSFSπps procedure. #

Remark 4.8: Ohlsson (1995) emphasizes the following attractive feature of Poisson πps. If the frame contains ***over - coverage*** (out - of - scope units), a "quasi Poisson πps" sample from the (imaginary) list of in - scope units can be selected by the following procedure. Units which turn out to be out - of - scope are excluded from the sample and replaced by units with smallest ranking variable values among so far "un - used" units, until a size n sample of in - scope units is obtained. In this process one looses the full control of the inclusion probabilities, though, since the size sum over the in - scope list is not known. However, if the aim is to estimate a ratio $\mu(y;x) = \tau(y)/\tau(x)$, this does not matter, since the unknown size sum in nominator and denominator will cancel. (If the aim is to estimate a total over the in - scope units, the size sum should be estimated.) This type of "over-coverage adjustment" can be employed for any OSFSπps scheme, i.a. for Pareto πps. #

Remark 4.9: Another attractive aspect of Poisson πps which also is emphasizes in Ohlsson (1995) as follows. If the U_i:s in the ranking variables U_i/λ_i are ***permanent random numbers***, a new Poisson πps sample from an updated version of the frame is positively coordinated (= has big overlap) with the previous sample. Negative coordination (also for simultaneous samples), is obtained if U_i is exchanged for $1 - U_i$. From (3.3) is seen that any OSFSπps scheme has these coordination properties and, hence, also Pareto πps. #

Remark 4.10: The above Pareto sampling - estimation procedure is formulated for the ideal case when all sampled units respond. However, in virtually all practical surveys ***non-response*** occurs. How to cope with non-response situations depends, of course, on what non-response model is judged to be realistic in the particular situation. Below we suggest an adjustment technique for Pareto πps, under the non-response model that the population can be partitioned into disjoint groups $G_1, ..., G_g, ...$, which are (fairly) homogeneous with respects to λ-values as well as response propensities. Then, use the following adjustment procedure (which will be better motivated elsewhere).

> Within each group G_g, re-calculate λ-values by $\lambda_i' = \lambda_i \cdot n_g'/n_g$, where n_g and n'_g denote the numbers of *sampled* respectively *responding* units in group G_g. Then carry out Steps 4 and 5 within each group G_g, using the λ′:s and $n = n_g'$, which leads to estimates of group totals with variance estimates. Finally, add group total estimates and variance estimates for group totals, to obtain an estimate of the population total with a variance estimate. Note that it suffices to re-calculate λ:s for sampled units. (4.13)

5 Evaluation of πps schemes

5.1 Introduction

Even if Pareto πps has the attractive property of being asymptotically uniformly optimal among OSFSπps schemes, it cannot be recommended right away. Below we list some additional issues that should be addressed.

- Asymptotic optimality among OSFSπps schemes does not say anything about how well Pareto πps compares with πps schemes in general. There may be better schemes outside the OSFSπps class. For a global evaluation (i.e. over all πps schemes), one would like to compare Pareto πps with the best πps scheme outside OSFSπps. This is a somewhat intricate wish, though, because there is no πps scheme which is unanimously regarded as "best so far". However, as already stated, we read Särndal et. al. (1992) so that they suggest that Sunter πps is a candidate for the title. The most widely used πps scheme in practice, systematic πps, is of course also of interest.
- The asymptotic optimality of Pareto πps was justified by using Approximation Result 3.1. As the name states, this result involves approximations; regarding unbiasedness of the point and variance estimators, the theoretical estimator variance formula, normal distribution of estimators. Although the approximation errors are asymptotically negligible, adequate practical application of an OSFSπps scheme requires that the approximations are "good enough" in the particular finite situation. Unfortunately, it is unfeasible to exhibit theoretical error bounds, from which it can be read off if the approximations are good enough in a particular situation. What can be done, though, to get an idea of the approximation errors is to carry out simulation studies.
- In addition to knowing that Pareto πps is (asymptotically) optimal among OSFSπps schemes, one wants a quantitative idea of its superiority over other OSFSπps schemes.
- The optimality property of Pareto πps relates only to (ii) in (1.6). The scheme should also be judged relative to (i) and (iii).

5.2 Theoretical considerations for OSFSπps schemes

Being (asymptotically) uniformly optimal, Pareto πps plays a distinguished role. Therefore we use the performance measure ***variance increase (VI) relative to Pareto πps.*** The parameters n, $\underline{\lambda}$ and **y** are presumed to be the same for compared schemes.

$$VI(\text{scheme S}) = \frac{\text{Estimator variance under } \pi\text{ps scheme S}}{\text{Estimator variance under Pareto } \pi\text{ps}} - 1. \quad (5.1)$$

Computation of exact VI - values is unfeasible, and we will resort to approximate ones;

$$AVI(\text{scheme S}) = \frac{\text{Approximate estimator variance under } \pi\text{ps scheme S}}{\text{Approximate estimator variance under Pareto } \pi\text{ps}} - 1. \quad (5.2)$$

In the first round we confine the comparisons to OSFSπps schemes, using (5.2) with approximate estimator variances by formula (3.9). Then the asymptotic optimality of Pareto πps entails that AVI - values are non - negative. By (3.18) and (4.2), AVI for OSFSπps(n; $\underline{\lambda}$; H) is, with $M(\mathbf{x};\underline{\lambda})$ and $R(\mathbf{x};\underline{\lambda};H)$ according to (3.19) and (3.20);

$$\text{AVI}(\mathbf{y};\underline{\lambda};H) = \frac{R(\mathbf{y};\underline{\lambda};H)}{M(\mathbf{y};\underline{\lambda})}. \quad (5.3)$$

Even if (5.3) gives an explicit expressions for AVI it is difficult to look through it, to see how it depends on **y**, $\underline{\lambda}$ and H. To make things a bit more transparent, we specialize to AVI for Poisson πps. By combining (3.14) and (4.2) we get the following expression for Poisson πps AVI, after some straightforward algebra which is left to the reader;

$$\text{AVI}(\mathbf{y};\underline{\lambda};\text{POI}) = \frac{\left(\sum_{i=1}^{N} y_i \cdot (\lambda_i - \frac{1}{n} \cdot \sum_{j=1}^{N} \lambda_j^2)\right)^2}{(n - \sum_{i=1}^{N} \lambda_i^2) \cdot \sum_{i=1}^{N} \frac{y_i^2}{\lambda_i} \cdot (1-\lambda_i) - (\sum_{i=1}^{N} y_i \cdot (1-\lambda_i))^2}. \quad (5.4)$$

Formula (5.4) is still quite difficult to look through, though, and we specialize it further by considering a particular type of sampling situation.

For $a \geq 0$, let the study variable values be $y_i = i^a$, $i = 1, 2, \ldots, N$, and let the target inclusion probabilities be proportional to i, $i = 1, 2, \ldots, N$. Then, after straightforward (but cumbersome) calculations one arrives at the following asymptotic (as $N \to \infty$) expression;

$$\text{AVI(POI)} = \left(\frac{n}{N}\right)^2 \cdot \frac{16a}{9(a+2)^2} \Bigg/ \left(1 - \left(\frac{n}{N}\right) \cdot \frac{4(5a+2)(a+1)}{3(2a+1)(a+2)} + \left(\frac{n}{N}\right)^2 \cdot \frac{16a(a+1)^2}{3(2a+1)(a+2)^2}\right). \quad (5.5)$$

Two special cases of (5.5) are listed below;

$$\text{For } \frac{n}{N} = 0.5, \quad \text{AVI(POI)} = \frac{4a \cdot (2a+1)}{3(a^2 + 8a + 4)}, \quad (5.6)$$

$$\text{For } a = 2, \quad \text{AVI(POI)} = \frac{2}{9} \cdot \left(\frac{n}{N}\right)^2 \Bigg/ \left(1 - 2.4 \cdot \left(\frac{n}{N}\right) + 1.2 \cdot \left(\frac{n}{N}\right)^2\right). \quad (5.7)$$

Remark 5.1: We do not know the maximal AVI for Poisson πps (or for any other OSFSπps scheme). By letting $a \to \infty$ in (5.6), it is seen that maximal AVI(POI) is at least $8/3 = 267\%$. #

Even if (5.6) and (5.7) relate to situations which are a bit extreme, we will use them to draw tentative conclusions on how AVI depends on the sampling situation.

The ideal πps sampling situation is when **y** - and **s** -values are exactly proportional. Then the population **y** - total is estimated without error. In this ideal situation, the values of (**y**, **s**) lie along a line through the origin. In practice this situation never occurs, though, the (**y**, **s**) - values ***scatter*** more or less around a ***trend***. The trend may be ***proportional*** (= linear through the origin) or ***non - proportional***. Type of trend as well as degree of scatter may vary considerably between practical sampling situations. If "powerful" auxiliary information is available, one may be in a situation with proportional trend and little scatter. In cases where only "weak" auxiliary information is at hand, non-proportionality as well as degree of scatter may be high.

A fact which reinforces variation in scatter and trend is that most surveys are multi-purpose (i.e. involve many study variables, and commonly also many domains). However, only one size measure can be used when selecting the sample. Therefore, in practical πps contexts one meets situations with good πps properties (as regards trend and scatter) for some of the study variables, but considerably less good for others. The usual way to cope with this dilemma is to chose the size measure with regard to the variable that is judged to be most important.

Against this background we look at formula (5.6). For $a = 1$ we are in a perfect πps situation with proportional trend and no scatter. The corresponding AVI - value in (5.6) is a bit difficult to interpret, though, because it is a degenerate value of type 0/0. For $a < 1$ we are in situations with ***increasing concave trend*** (**y** grows relatively slower than **s**) and for $a > 1$ in situations with ***increasing convex trend*** (**y** grows relatively faster than **s**). In situations with a ***flat* or *decreasing*** trend, πps sampling is in fact non - favorable compared with simple random sampling. As is readily checked, AVI in (5.6) increases with a. A tentative conjecture is therefore that Pareto πps is more superior, the more convex the trend is. However, as will be seen later on, this picture is disturbed when scatter comes in, and we formulate the following looser version of the conjecture;

The degree of superiority for Pareto πps depends on the form of the (**y**,**s**) - trend. (5.8)

Next we turn to (5.7), which shows that AVI depends on the ***sampling fraction***. It is readily checked that AVI in (5.7) increases from 0 to 55% when the sampling fraction increases from 0 to 1/2, which provides background for the following conjecture;

Pareto πps is more superior, the higher the sampling fraction is. (5.9)

A requirement on a good πps scheme is that it works efficiently also in situations with high sampling rates, for the following reason. Even if high over - all sampling rates are unusual, high "partial" sampling rates occur, in particular when πps sampling is used in conjunction with the following type of stratification. The population is stratified by study domains and πps samples are selected from the strata. Some strata may be comparatively small, which in combination with the fact that the precision of a stratum estimate depends mainly on stratum sample size (not on stratum sampling rate) may lead to high sampling rates in some strata.

5.3 Numerical comparisons

5.3.1 Generation of sampling situations

Here we employ a more concrete way to indicate how AVI depends on the sampling situation, by presenting numerical AVI - values for specific situations. As always in this type of context, one meets the problem that there is an "ocean" of situations of potential interest, while the space for the paper is limited. We have to restrict, and then we let (5.8) and (5.9) be guiding.

We shall consider situations generated by the following ***sampling situation model*** for the relation between the study variable **y** and the size measure **s**;

$$s_i = i, \quad y_i = c\cdot(s_i^a + \sigma\cdot Z_i\cdot\sqrt{s_i^a}), \quad c>0, \quad \sigma\geq 0, \quad \text{the } Z_i \text{ being iid } N(0,1), \quad i=1,2,...,N. \qquad (5.10)$$

The value of the parameter c does not affect the comparisons, we include it just to indicate the scope of the model. The vital parameters are N, σ and a, and they will be varied. Variation of σ and a can be interpreted as follows. Increase of σ leads to increased scatter. When a moves away from 1, the trend moves away from proportionality, it is concave for $a < 1$ and convex for $a > 1$. To make the generated situations reproducible, we inform that the Z:s were generated by the SAS6.10 function NORMAL(seed) with first seed (for $i = 1$) = 555.

One may think that model (5.10) with $a \neq 1$ is uninteresting, since the situation can then be brought closer to the ideal πps situation by letting s^a play the role of size measure. However, we mean that $a \neq 1$ is of interest, for at least two reasons. (i) Often, the sampler has only vague knowledge of the (**y**, **s**) - trend. He / she may judge it wrongly, and believe in proportionality while "truth" is non - proportional ($a \neq 1$). (ii) The fit between **y** and **s** was a secondary concern in the sample selection, **s** was chosen with regard to a more important study variable than **y**.

Numerical AVI - values have been derived by two approaches; (a) Using the asymptotic formulas for OSFSπps estimator variances, (b) Monte Carlo simulations. Approach (a) leads to easy computations, but does not tell how well the asymptotic formulas approximate the true estimator variances. Approach (b) leads to considerably greater computation efforts but it gives, at least some, insight into the approximation goodness problem. Comparisons with Sunter πps are made by both approaches. An algorithm for computation of second order Sunter inclusion probabilities is available. Having those, standard formulas yield estimator variances and variance estimates. Since Sunter πps lies outside OSFSπps, its AVI - values may be negative, implying that it performs better than Pareto πps. The performance of Sunter confidence intervals is studied by simulations. For systematic πps we resort to simulations.

Before presenting numerical results, we formulate the tentative conclusions we shall draw from them. We do not lay claim to give a complete picture of how AVI - values vary with the sampling situation. (To the best of our understanding, this is a very complex problem.)

Conclusions concerning the measurable πps schemes; OSFSπps, Sunter πps and systematic πps with random frame order (***rfo***).

When πps is better than SRS, i.e. when the (**y**,**s**) - trend is fairly strongly increasing.

The OSFSπps schemes and systematic πps(rfo) all perform better than Sunter πps, and the more better the higher the sampling rate is. The advantage over Sunter πps is greater, the more beneficial πps sampling is relative to simple random sampling. (5.11)

Pareto, exponential and Poisson πps compare as follows.

Under proportional (y, s) - trend: All three schemes have very similar performances. There is an edge for Pareto πps, but it is of little to no practical importance. (5.12)

Under non - proportional (y, s) - trend: Pareto πps performs better than exponential πps, which in turn performs better than Poisson πps. The differences are small when sampling rates are small, but may be pronounced for high sampling rates. (5.13)

Pareto πps compares with systematic πps(rfo) in very much the same way as with the OSFSπps schemes. When the (**y**,**s**) - trend is fairly proportional, the two schemes have similar performances, with a slight edge for Pareto πps. When the (**y**,**s**) - trend is non - proportional, Pareto πps performs better than systematic πps(rfo). Differences are small, though, for small sampling rates, but may be pronounced for high sampling rates. (5.14)

When πps is worse than SRS, i.e. when the (**y**,**s**) - trend is flat or decreasing.

The bad performance of πps sampling is more accentuated for the OSFSπps schemes and systematic πps(rfo) than for Sunter πps, which performs better than the other. There is an edge for Pareto πps among "non - Sunter" schemes. (5.15)

Conclusions concerning systematic πps with frame ordered by size (sfo).

Systematic πps(sfo) has considerably better point estimation precision than the other πps schemes when the (**y**,**s**) - trend is distinctly non - proportional. However, under fairly proportional (**y**, **s**) - trend, systematic πps(sfo) and Pareto πps compare in an erratic way, they take turns to be better than the other. (5.16)

5.3.2 Comparisons of OSFSπps and Sunter πps, based on asymptotic formulas

In the following we present numerical AVI - values for exponential, Poisson and Sunter πps and also for simple random sampling, SRS. SRS is not included in the study as a "competitor", though, only to show the relative efficiency of πps - sampling versus SRS. For the OSFSπps schemes the asymptotic formulas were used to compute (approximate) estimator variances, while the exact standard formulas were used for Sunter πps and SRS. We now turn to the numerical results which provide background for the claims in (5.11)-(5.16).

We start by illustrating (5.11) and (5.12), by considering situations with a = 1 in (5.10), i.e. situations with proportional trend.

Table 1	AVI for situation (5.10) with N=100, σ=2, a=1				
Sampling fraction f	**f=0.1**	**f=0.2**	**f=0.3**	**f=0.4**	**f=0.5**
Poisson πps	$4 \cdot 10^{-4}$ %	$2 \cdot 10^{-3}$ %	$6 \cdot 10^{-3}$ %	0.02 %	0.04 %
Exponential πps	$1 \cdot 10^{-4}$ %	$6 \cdot 10^{-4}$ %	$2 \cdot 10^{-3}$ %	$7 \cdot 10^{-3}$ %	0.03 %
Sunter πps	2.6 %	58 %	181 %	280 %	401 %
SRS	450 %	446 %	440 %	432 %	421 %

The AVI - values in Table 1 comply with (5.11) and (5.12). One may wonder, though, if the fairly small population size (N = 100) affects the AVI - values in some particular direction. A "magnified" situation with roughly the same degree of scatter (at least for large y - values) is obtained by increasing N by a factor 4 and σ by a factor 2. The corresponding AVI - values are shown in Table 2, which again comply with (5.11) and (5.12). They also indicate that population size is not a crucial factor. Therefore, in the rest of this section we confine to N=100.

Table 2	AVI for situation (5.10) with N=400, σ=4, a=1				
Sampling fraction f	**f=0.1**	**f=0.2**	**f=0.3**	**f=0.4**	**f=0.5**
Poisson πps	$4 \cdot 10^{-4}$ %	$2 \cdot 10^{-3}$ %	$7 \cdot 10^{-3}$ %	0.02 %	0.05 %
Exponential πps	$1 \cdot 10^{-4}$ %	$5 \cdot 10^{-4}$ %	$1 \cdot 10^{-3}$ %	$3 \cdot 10^{-3}$ %	$5 \cdot 10^{-3}$ %
Sunter πps	13 %	79 %	192 %	344 %	469 %
SRS	423 %	430 %	438 %	450 %	468 %

Next we vary the degree of scatter. In Table 3 the scatter is increased (compared with Table 1) to σ = 4, and in Table 4 it is decreased to σ = 0.5. The figures in these tables again illustrate (5.11) and (5.12), in particular the latter part of (5.11).

Table 3	AVI for situation (5.10) with N=100, σ=4, a=1				
Sampling fraction f	**f=0.1**	**f=0.2**	**f=0.3**	**f=0.4**	**f=0.5**
Poisson πps	$4 \cdot 10^{-4}$ %	$2 \cdot 10^{-3}$ %	$6 \cdot 10^{-3}$ %	0.02 %	0.04 %
Exponential πps	$1 \cdot 10^{-4}$ %	$6 \cdot 10^{-4}$ %	$2 \cdot 10^{-3}$ %	$7 \cdot 10^{-3}$ %	0.03 %
Sunter πps	9.8 %	8.1 %	47 %	60 %	87 %
SRS	105 %	103 %	101 %	98 %	94 %

Table 4	AVI for situation (5.10) with N=100, σ=0.5, a=1				
Sampling fraction f	**f=0.1**	**f=0.2**	**f=0.3**	**f=0.4**	**f=0.5**
Poisson πps	$4 \cdot 10^{-4}$ %	$2 \cdot 10^{-3}$ %	$6 \cdot 10^{-3}$ %	0.02 %	0.04 %
Exponential πps	$1 \cdot 10^{-4}$ %	$6 \cdot 10^{-4}$ %	$2 \cdot 10^{-3}$ %	$7 \cdot 10^{-3}$ %	0.03 %
Sunter πps	309 %	1187 %	2933 %	4950 %	6820 %
SRS	7412 %	7347 %	7266 %	7161 %	7019 %

Next we consider variation of the **(y, s)** - trend. First we look at growingly convex situations by considering a = 1.2, 1.5 and 2. It should be noted that a = 2 yields a rather extreme situation, though. In the first round we use $\sigma = 2$. The AVI - values are shown in Tables 5 - 7. They illustrate the claim in (5.13). Also (5.11) is illustrated. Even if a negative Sunter AVI - value turns up at one place (for a = 2), we still mean that (5.11) gives the over - all picture.

Table 5	**AVI for situation (5.10) with N = 100, σ=2, a = 1.2**				
Sampling fraction f	**f = 0.1**	**f = 0.2**	**f = 0.3**	**f = 0.4**	**f = 0.5**
Poisson πps	0.06 %	0.3 %	1.0 %	2.8 %	7.8 %
Exponential πps	0.01 %	0.1 %	0.2 %	0.5 %	0.9 %
Sunter πps	33 %	128 %	418 %	790 %	1321 %
SRS	1213 %	1226 %	1249 %	1290 %	1387 %

Table 6	**AVI for situation (5.10) with N = 100, σ=2, a = 1.5**				
Sampling fraction f	**f = 0.1**	**f = 0.2**	**f = 0.3**	**f = 0.4**	**f = 0.5**
Poisson πps	0.2 %	1.3 %	4.2 %	12 %	39 %
Exponential πps	0.1 %	0.3 %	0.9 %	2.3 %	4.7 %
Sunter πps	5.8 %	62 %	262 %	653 %	1457 %
SRS	1083 %	1095 %	1130 %	1222 %	1538 %

Table 7	**AVI for situation (5.10) with N = 100, σ=2, a = 2**				
Sampling fraction f	**f = 0.1**	**f = 0.2**	**f = 0.3**	**f = 0.4**	**f = 0.5**
Poisson πps	0.3 %	1.5 %	5.1 %	15 %	54 %
Exponential πps	0.1 %	0.4 %	1.1 %	2.8 %	6.2 %
Sunter πps	-4.1 %	3.3 %	66 %	269 %	877 %
SRS	536 %	550 %	582 %	660 %	941 %

In Tables 5 - 7, σ is kept fixed when a is increased, which leads to relatively decreased **(y, s)** - scatter for units with large variable values. One may wonder if the effects that can be observed in Tables 5 - 7 are to be ascribed to change of trend or to lowered relative scatter. We believe that the trend factor is most important. To illustrate this, Table 8 presents AVI - values for the situation in Table 6 with doubled σ, i.e. with $\sigma = 4$.

Table 8	**AVI for situation (5.10) with N = 100, σ=4, a = 1.5**				
Sampling fraction f	**f = 0.1**	**f = 0.2**	**f = 0.3**	**f = 0.4**	**f = 0.5**
Poisson πps	0.2 %	0.9 %	3.0 %	8.5 %	27 %
Exponential πps	0.04 %	0.2 %	0.6 %	1.6 %	3.1 %
Sunter πps	6.7 %	44 %	198 %	466 %	1018 %
SRS	775 %	821 %	821 %	889 %	1084 %

Next we consider concave situations, i.e. situations with a < 1. In order that the relative scatter should not increase, and also to avoid negative **y** - values, we decrease σ when a is decreased. The findings in Table 9 illustrate (5.13) as well as (5.11).

Table 9	**AVI for situation (5.10) with N = 100, σ = 0.5, a = 0.7**				
Sampling fraction f	**f = 0.1**	**f = 0.2**	**f = 0.3**	**f = 0.4**	**f = 0.5**
Poisson πps	0.1 %	0.7 %	1.9 %	4.7 %	12 %
Exponential πps	0.03 %	0.2 %	0.5 %	1.7 %	2.0 %
Sunter πps	25 %	43 %	103 %	141 %	180 %
SRS	246 %	229%	213 %	197 %	186 %

In Table 9, a = 0.7. If a is further decreased, we enter the region where πps sampling no longer is advantageous over SRS. For a = 0.5, the (y, s) - trend is quite flat, and for a = 0 it is entirely flat. Corresponding AVI - values, which illustrate (5.15), are shown in Tables 10 and 11.

Table 10	**AVI for situation (5.10) with N = 100, σ = 0.5, a = 0.5**				
Sampling fraction f	**f = 0.1**	**f = 0.2**	**f = 0.3**	**f = 0.4**	**f = 0.5**
Poisson πps	0.1 %	0.6 %	1.8 %	4.4 %	11 %
Exponential πps	0.03 %	0.2 %	0.4 %	1.0 %	1.8 %
Sunter πps	-5.3 %	-30 %	-35 %	-46 %	-48 %
SRS	-31 %	-35 %	-40 %	-44%	-47 %

Table 11	**AVI for situation (5.10) with N = 100, σ = 0.2, a = 0**				
Sampling fraction f	**f = 0.1**	**f = 0.2**	**f = 0.3**	**f = 0.4**	**f = 0.5**
Poisson πps	0.1 %	0.4 %	1.0 %	2.4 %	5.6 %
Exponential πps	0.02 %	0.1 %	0.3 %	0.6 %	1.0 %
Sunter πps	-50 %	-76 %	-89 %	-96 %	-99 %
SRS	-98 %	-98 %	-98 %	-98 %	-98 %

5.3.3 Findings from simulations

This section reports on results from simulations with two aims: (i) To give some insight into the goodness of the approximations that are set to work in practical application of OSFSπps schemes. (ii) Comparisons with systematic πps.

We employed a standard simulation approach. A sampling situation was generated by the model (5.10), yielding a known τ(y). Repeated independent πps samples with prescribed sample sizes were drawn by Pareto, exponential, Poisson Sunter and systematic πps. 3 000 repetitions were used throughout. For each sample we computed the total - estimate $\hat{\tau}$, by (3.7) for the OSFS schemes and by the HT - estimator for the other schemes, variance estimate $\hat{V}(\hat{\tau})$, by (3.11) by for the OSFS schemes and by (1.3) for Sunter πps. Moreover, approximate 95% confidence intervals $\hat{\tau} \pm 1.96 \cdot \sqrt{\hat{V}(\hat{\tau})}$ for τ(y) were computed for OSFS and Sunter πps, and it was checked if they covered τ(y) or not. The following summary statistics were derived; the ***empirical mean***, $m(\hat{\tau})$, and ***empirical variance***, $S^2(\hat{\tau})$, for the estimates $\hat{\tau}$, the empirical mean of the variance estimates, $m(\hat{V}(\hat{\tau}))$ and ***empirical cover rates*** for confidence intervals.

On approximation goodness

The crucial questions are: When is the bias in the estimator (3.7) reasonably small? When is the bias in the variance estimator (3.11) reasonably small? When is a normal approximation confidence interval good enough? In addition we have a question of more theoretical interest: How well do asymptotic variances by (3.9) approximate the true ones?

Results for the OSFSπps and Sunter schemes are shown in Tables 12 - 14, which treat situations with quite different trend forms. Since relative discrepancies are simpler to grasp than absolute ones, we present the results in terms of ***relative bias for the point estimator***, $m(\hat{\tau})/\tau(x) - 1$, and ***relative bias for the variance estimator***, $m[\hat{V}(\hat{\tau})]/S^2(\hat{\tau}) - 1$. Sunter's scheme is not included in the bias study, since its point and variance estimator are unbiased.

Table 12. Performance of πps procedures in situation N = 100, σ = 2, a = 1											
Sampling	Relative bias for point estimator			Relative bias for variance estimator			Empirical coverage rates for confidence intervals				
rate	PAR	EXP	POI	PAR	EXP	POI	PAR	EXP	POI	SUNT
0.1	-0.1%	-0.1%	-0.1%	-4.5%	-4.1%	-4.2%	92.1%	92.1%	92.2%	92.5%
0.2	0%	0%	0%	-3.7%	-2.7%	-2.4%	93.2%	93.4%	93.6%	92.6%
0.3	0%	0%	0%	-2.9%	-2.9%	-3.6%	93.7%	93.8%	93.8%	93.2%
0.4	0%	0%	0%	0.3%	0.2%	0.2%	93.9%	94.0%	93.9%	94.0%
0.5	0%	0%	0%	1.7%	1.8%	1.7%	94.4%	94.3%	94.4%	95.0%

Table 13. Performance of πps procedures in situation N = 100, σ = 2, a = 1.5										
Sampling	Relative bias for point estimator			Relative bias for variance estimator			Empirical coverage rates for confidence intervals			
rate	PAR	EXP	POI	PAR	EXP	POI	PAR	EXP	POI	SUNT
0.1	0.1%	0.2%	0.2%	-1.4%	-1.1%	-1.3%	90.5%	90.4%	90.3%	89.5%
0.2	0.1%	0.2%	0.2%	2.0%	2.0%	0.9%	93.7%	93.7%	93.3%	91.0%
0.3	0.1%	0.2%	0.2%	0.7%	1.0%	0.2%	93.1%	92.7%	92.4%	93.7%
0.4	0%	0.1%	0.1%	-1.7%	-1.7%	-0.6%	93.0%	92.8%	93.1%	94.0%
0.5	0%	0%	0%	-4.4%	-3.5%	1.8%	93.0%	92.9%	93.3%	95.0%

Table 14. Performance of πps procedures in situation N = 100, σ = 0.5, a = 0.7										
Sampling	Relative bias for point estimator			Relative bias for variance estimator			Empirical coverage rates for confidence intervals			
rate	PAR	EXP	POI	PAR	EXP	POI	PAR	EXP	POI	SUNT
0.1	0%	0%	0%	-7.9%	-7.7%	-7.2%	85.6%	85.3%	85.4%	85.5%
0.2	-0.1%	-0.1%	-0.1%	-2.7%	-3.4%	-3.7%	88.3%	88.1%	88.1%	87.3%
0.3	0.1%	0%	-0.1%	-5.8%	-4.8%	-5.4%	88.4%	88.2%	87.7%	93.5%
0.4	0%	0%	-0.1%	-3.0%	-3.5%	-3.0%	88.5%	88.4%	88.1%	94.2%
0.5	0%	0%	0%	-2.5%	-4.3%	-2.5%	87.7%	87.8%	87.8%	94.8%

From Tables 12 - 14, and also from Tables 15 and 16, is seen that all three OSFS πps procedures have exceedingly small bias for the point estimator. Hence, that part of the approximation issue seems to cause no problem.

The picture is not equally perfect as regards variance estimators and confidence levels. For the latter the approximation goodness depends jointly on the goodness of the normal distribution approximation and the approximations behind the variance estimators. Since sample sizes are fairly small, the sampling situations in Tables 13 and 14 are quite dispersed (in the sense of Remark 4.3) and a number of target inclusion probabilities lie fairly close to 1 when $f > 0.3$, the magnitudes of the approximation errors in Tables 13 and 14 are not too surprising. Larger sample sizes are required for really good approximation.

Tables 15 and 16 treat "enlarged" versions (N is increased) of the situations in Tables 13 and 14. They show that the approximations improve when sample sizes are increased. A lesson from Tables 12 - 16 is, as in most normal approximation contexts: There is no simple rule of type "$n > 20$" which guarantees good normal approximation, a proper rule must be more complicated. The results indicate that dispersion (in the sense of Remark 4.3) is a crucial factor. When it is small (as in Table 12), confidence intervals have decent levels already for sample sizes 10-20, while larger sample sizes are required in more dispersed situations.

Another conclusion from Tables 12 - 16 runs as follows. Even if Sunter's procedure, as regards approximations, employs only the normal distribution approximation, while the OSFS procedures are based on considerably more sophisticated approximation arguments, OSFS confidence intervals seem to perform as well as Sunter confidence intervals. This holds at least as long as Sunter πps is "genuine" πps. In the considered situations Sunter πps is very close to SRS for f=0.4 and f=0.5 (which also is reflected by the AVI - values in Tables 1,6 and 9.

Table 15. Performance of πps procedures in situation N = 200, σ = 2, a = 1.5

Sampling	Relative bias for point estimator			Relative bias for variance estimator			Empirical coverage rates for confidence intervals			
rate	PAR	EXP	POI	PAR	EXP	POI	PAR	EXP	POI	SUNT
0.1	0.1%	0.1%	0.1%	1.3%	1.2%	1.7%	92.5%	92.6%	92.5%	92.7%
0.2	0.1%	0.1%	0.1%	1.8%	1.8%	1.3%	93.6%	93.6%	93.4%	92.4%
0.3	0%	0%	0%	2.3%	3.0%	3.5%	93.7%	93.7%	93.6%	93.2%
0.4	0%	0%	0%	-2.4%	-2.0%	-0.7%	94.7%	94.6%	95.0%	94.6%
0.5	0%	0%	0%	-0.5%	-0.3%	3.0%	93.5%	93.2%	94.1%	95.6%

Table 16. Performance of πps procedures in situation N = 300, σ = 0.5, a = 0.7

Sampling	Relative bias for point estimator			Relative bias for Variance estimator			Empirical coverage rates for confidence intervals			
rate	PAR	EXP	POI	PAR	EXP	POI	PAR	EXP	POI	SUNT
0.1	0.1%	0%	0%	-1.6%	-1.8%	-1.8%	89.7%	89.6%	89.6%	88.9%
0.2	0%	0%	0%	4.7%	3.9%	3.4%	91.2%	92.0%	91.7%	92.7%
0.3	0%	0%	0%	-1.9%	-2.8%	-3.1%	91.5%	91.4%	91.1%	94.2%
0.4	0%	0%	0%	-0.9%	-0.4%	-0.5%	92.3%	92.3%	92.2%	94.4%
0.5	0%	0%	0%	-1.9%	-0.1%	2.6%	91.6%	92.0%	92.2%	95.2%

It is also of interest to see how well the asymptotic variances in (3.9) comply with the empirical ones. Tables 17 - 19 present ***relative biases in the asymptotic variances*** in (3.9), defined as: [variance by asymptotic formula (3.9)] / [empirical variance] - 1.

Our conclusion from Tables 17 - 19 is that the asymptotic variance formula (3.9) work with surprisingly good approximation also for fairly small n - and N - values. Hence, the AVI - values in Tables 1-11, in fact lie quite close to the corresponding VI - values in (5.1).

Table 17	**Relative bias in asymptotic variance (3.9) for N = 100, σ = 2, a = 1**				
Sampling fraction	**f = 0.1**	**f = 0.2**	**f = 0.3**	**f = 0.4**	**f = 0.5**
Pareto πps	-5.2 %	-2.8 %	-2.3 %	1.5 %	3.5 %
Exponential πps	-4.5 %	-1.6 %	-2.1 %	1.7 %	3.8 %
Poisson πps	-4.5 %	-1.2 %	-2.5 %	1.4 %	3.0 %

Table 18	**Relative bias in asymptotic variance (3.9) for N = 200, σ = 2, a = 1.5**				
Sampling fraction	**f = 0.1**	**f = 0.2**	**f = 0.3**	**f = 0.4**	**f = 0.5**
Pareto πps	0.6%	3.5%	2.5%	-2.0 %	1.2 %
Exponential πps	0.7 %	3.7 %	3.4 %	-1.6 %	1.3 %
Poisson πps	0.8 %	3.7 %	3.9 %	-0.2 %	4.2 %

Table 19	Relative bias in asymptotic variance (3.9) for N = 300, σ = 0.5, a = 0.7				
Sampling fraction	f = 0.1	f = 0.2	f = 0.3	f = 0.4	f = 0.5
Pareto πps	-0.2 %	4.7 %	-3.5 %	-0.6 %	0.1 %
Exponential πps	-0.1 %	4.1 %	-4.8 %	-0.1 %	2.1 %
Poisson πps	-0.3 %	4.2 %	-4.6 %	-0.6 %	4.4 %

Comparisons with systematic πps

Comparisons are made between Pareto, Sunter and systematic πps, with random frame order (*rfo*) as well as size ordered frame (*sfo*) by ***empirical AVI - values*** = ($S^2(\hat{t})$ for "alternative" scheme)/(asymptotic variance for Pareto πps) - 1. Results are presented in Tables 20 - 26, which are ordered by decreasing a - values. For systematic πps(rfo) the picture is fairly clear, namely as formulated in (5.14) and (5.15).

For systematic πps(sfo) the findings, which are summarized in (5.16), are a bit surprising. When the (y, s) - trend deviates pronuoncedly from proportional (a = 2, 1.5, 0.7, 0.5 and 0), systematic πps(sfo) has dramatically better estimation precision than the measurable πps schemes. However, when the (y, s) - trend is fairly proportional (a = 1.2, and 1), which is the kind of trend one is aiming for in practice, the picture becomes fuzzy, big positive and negative AVI-values mix in an irregular way.

It is difficult to draw clear - cut practical conclusions, perhaps they should run as follows. If one believes that the (y, s) - trend lies well away from proportional, systematic πps(sfo) is to be preferred to OSFSπps and systematic πps(rfo). However, even more preferable is probably to do as follows, if possible. If a is distinctly greater than 1, transform the size measure to achieve a more proportional (y, s) - trend. If a is considerably less than 1, avoid any kind of πps scheme since SRS is better, or transform the size measure to achieve a more proportional (y, s) - trend. When the (y, s) - trend is fairly proportional, systematic πps(sfo) may lead to substantially better estimation precision than Pareto πps, but the opposite may also occur. It seems difficult to tell which alternative is at hand in a specific situation.

Table 20	Empirical AVI for (5.10) with N = 100, σ = 2.0, a = 2.0				
Sampling fraction	f = 0.1	f = 0.2	f = 0.3	f = 0.4	f = 0.5
Systematic πps, random frame order	1.9%	0.6%	1.3%	26%	35%
Systematic πps, frame order by size	-86%	-89%	-93%	-94%	-91%
Sunter πps	-4.1 %	3.3 %	66 %	269 %	877 %

Table 21	Empirical AVI for (5.10) with N = 100, σ = 2, a = 1.5				
Sampling fraction	f = 0.1	f = 0.2	f = 0.3	f = 0.4	f = 0.5
Systematic πps, random frame order	2.2%	1.0%	0.7%	21%	29%
Systematic πps, frame order by size	-74%	-75%	-73%	-86%	-73%
Sunter πps	5.8 %	62 %	262 %	653 %	1457 %

Table 22	Empirical AVI for (5.10) with N = 100, σ = 2.0, a = 1.2				
Sampling fraction	f = 0.1	f = 0.2	f = 0.3	f = 0.4	f = 0.5
Systematic πps, random frame order	2.9%	-0.7%	3.5%	3.0%	3.3%
Systematic πps, frame order by size	-18%	-27%	+46%	-41%	+1.2%
Sunter πps	33 %	128 %	418 %	790 %	1321 %

Table 23	Empirical AVI for (5.10) with N=100, σ=2, a=1				
Sampling fraction	f=0.1	f=0.2	f=0.3	f=0.4	f=0.5
Systematic πps, random frame order	2.4%	-3.5%	0.6%	-2.2%	-3.3%
Systematic πps, frame order by size	+3.6%	-16%	+88%	-21%	+12%
Sunter πps	2.6 %	58 %	181 %	280 %	401 %

Table 24	Empirical AVI for (5.10) with N=100, σ=0.5, a=0.7				
Sampling fraction	f=0.1	f=0.2	f=0.3	f=0.4	f=0.5
Systematic πps, random frame order	1.9%	-2.2%	0.3%	8.7%	13%
Systematic πps, frame order by size	-34%	-52%	-9.2%	-52%	-47%
Sunter πps	25 %	43 %	103 %	141 %	180 %

Table 25	Empirical AVI for (5.10) with N=100, σ=0.5, a=0.5				
Sampling fraction	f=0.1	f=0.2	f=0.3	f=0.4	f=0.5
Systematic πps, random frame order	3.4%	-0.4%	-0.6%	7.3%	13%
Systematic πps, frame order by size	-34%	-49%	-17%	-50%	-50%
Sunter πps	-5.3 %	-30 %	-35 %	-46 %	-48 %

Table 26	Empirical AVI for (5.10) with N=100, σ=0.2, a=0				
Sampling fraction	f=0.1	f=0.2	f=0.3	f=0.4	f=0.5
Systematic πps, random frame order	13%	11%	-6.8%	-0.4%	14%
Systematic πps, frame order by size	-26%	-33%	-26%	-40%	-43%
Sunter πps	-50 %	-76 %	-89 %	-96 %	-99 %

5.4 General conclusions

When considering computation aspects in the sequel, we pay regard to programming but not to computer time, the cost of which is negligible nowadays (at least for sample surveys). The weight to be assigned to programming work depends, of course, on the number of times the survey will be conveyed, the fewer the more important to keep programming down.

A. Comparison of OSFSπps and Sunter πps

We compare the OSFSπps schemes and Sunter πps, and we follow points (i)-(iii) in (1.6).

Simplicity of sample selection: As stated in Remark 4.6, the description of Pareto πps in Section 4.3 holds in its general structure for all OSFSπps schemes. As is seen from Steps 1 - 3 in that description, programming of sample selection for an OSFSπps scheme is very simple. Programming for Sunter's scheme is also simple, a bit less than for OSFSπps, though, because of the need to modify the size measures.

Estimation precision: Sunter's point estimator is unbiased. Tables 12 - 16 indicate that the OSFSπps schemes all have very small point estimator bias. Hence, point estimator bias can be disregarded, the crucial aspect is estimator variances.

As stated in (5.11), the OSFS schemes perform better, often considerably better, than Sunter πps in situation where πps sampling is more efficient than SRS. Also (5.15) has relevance in the comparison. Tables 10 and 11 tell that Sunter πps performs better than OSFSπps in situations where πps sampling is disadvantageous to SRS. The conclusion we draw from that is, however, not that OSFSπps should be avoided in favor of Sunter πps in such situations, but that any kind of πps sampling should be avoided (or, if possible, that the size measure should be transformed so as to obtain a more proportional (**y**, **s**)-trend.

Variance estimation properties: Pertinent aspects are listed in a) - c) below.

a) Sunter πps yields unbiased variance estimates, which are non - negative for the variance estimator (1.3). The OSFSπps schemes have consistent, non - negative variance estimates which are very mildly biased.

b) Even if the OSFSπps procedures are based on more elaborate approximation arguments than Sunter πps, their normal approximation confidence intervals seem to have as good level properties as for Sunter πps, at least when Sunter πps is not close to SRS (thereby being fairly uninteresting).

c) Variance estimation for the OSFSπps schemes is very simple to program, it has "complexity order" = computation of an ordinary sample variance (see (4.9) and Remark 4.4). The only data from the sample selection needed at the estimation stage are the normed size - values. Variance estimation for Sunter πps requires more involved programming. Second order inclusion probabilities should be derived at the sample selection stage, by a somewhat intricate algorithm, and be saved to the estimation phase. The variance estimation formulas themselves are also a bit intricate.

Summing up, we mean that the OSFSπps schemes lead to fully acceptable variance estimates, which are considerably simpler to compute than those for Sunter πps.

Comparison relating to Remarks 4.8 and 4.9: There is no Sunter πps analogue to the "over - coverage adjustment" indicated in Remark 4.8. Due to the "list-sequential" ingredients in the sample selection for Sunter πps as well as for OSFSπps, all these schemes admit sample coordination as indicated in Remark 4.9.

Combined conclusions: The OSFSπps schemes are on no vital point inferior to Sunter πps and in many respects better, notably as regards the most crucial aspect, estimation precision. Hence OSFSπps schemes should be preferred to Sunter πps.

B. Comparison of OSFSπps schemes

Here we consider Pareto, exponential and Poisson πps. As regards simplicity of sample selection, simplicity of variance estimation and level properties for confidence intervals, they are regarded as essentially equivalent. Hence, the crucial aspect is estimation precision. In many situations the schemes perform very similarly, but there are situations where estimation precisions do differ substantially (see Tables 6 - 11). Then Pareto πps, being the optimal one, performs better than the other two, and it should therefore be preferred among OSFSπps schemes. Use of Pareto πps can be seen as an "insurance without premium", it never performs worse than other OSFSπps schemes, and in some situations it performs considerably better.

C. Comparisons of Pareto πps and systematic πps

Comparisons with systematic πps generally: Sample selection is simple for systematic πps, under random as well as size frame order, (rfo) resp. (sfo). The same holds for Pareto πps.

Over - coverage adjustment and sample coordination as indicated in Remarks 4.8 and 4.9 can be made for Pareto πps but, to our knowledge, no analogues exist for systematic πps.

Comparison with systematic πps(rfo): For systematic πps(rfo), the Hartley - Rao variance estimator mostly works with good approximation, even if not entirely asymptotically correct. If the Hartley - Rao estimator is regarded to yield acceptable variance estimates, Pareto πps and systematic πps(rfo) are fairly equivalent with respect to (iii) in (1.6). Programming of variance estimation is a bit simpler, though, for Pareto πps.

With regard to (ii), the main conclusion is formulated in (5.14). It states that Pareto πps never performs (noteworthy) worse than systematic πps(rfo), and sometimes considerably better. Hence, Pareto πps should be preferred.

Comparison with systematic πps(sfo): Systematic πps(sfo) has its specific weakness with respect to (iii) in (1.6), it does not admit objective assessment of sampling errors.

As regards (ii), point estimator precision, the picture is a bit confusing. In situations that deviate pronouncedly (maybe one should say extremely) from proportional $(\mathbf{y}, \mathbf{s})$ - trend, systematic πps(sfo) leads to dramatically better estimation precision than Pareto πps, and also than the other πps schemes. It is somewhat unclear what to recommend, though; Use of systematic πps(sfo) or transformation of the sizes. In situations with fairly proportional $(\mathbf{y}, \mathbf{s})$ - trend, the ranking between systematic πps(sfo) and Pareto πps seems to be erratic, the schemes take turns to be better than the other. We leave the problem there, by concluding that the choice between Pareto πps and systematic πps(sfo) becomes a matter of judgment/beliefs on behalf of the sampler.

D. Over - all conclusion

With the A, B and C as background our tentative overall conclusions run as follows;

- Pareto πps should be preferred among πps schemes that admit objective assessment of estimation precision.
- Systematic πps with frame ordered by the sizes sometimes leads to distinctly better point estimation precision than Pareto πps, but the opposite also occurs. It is difficult to tell if a particular sampling situation is advantageous for systematic πps(sfo) or not. In any case, when πps(sfo) is used, control of variance estimation is lost. Choice between Pareto πps and systematic πps(sfo) becomes a matter of judgment/beliefs for the sampler.

References

OHLSSON, E. (1995). Sequential Poisson Sampling. Research Report from Institute of Actuarial Mathematics and Mathematical Statistics at Stockholm University.

ROSÉN, B. (1995). Asymptotic Theory for Order Sampling. R&D Report 1995:1, Statistics Sweden.

SÄRNDAL, C-E, SWENSSON, B. & WRETMAN, J. (1992). *Model Assisted Survey Sampling.* Springer-Verlag. New York.

SUNTER A. B. (1977). List Sequential Sampling with Equal or Unequal Probabilities without Replacement. *Applied Statistics* **26**, 261 - 268.

WOLTER, K. M. (1985). *Introduction to Variance Estimation.* Springer-Verlag. New York.

R & D Reports är en för U/ADB och U/STM gemensam publikationsserie, som fr o m 1988-01-01 ersätter de tidigare "gula" och "gröna" serierna.

R & D Reports Statistics Sweden are published by the Department of Research & Development within Statistics Sweden. Reports dealing with statistical methods have green (grön) covers. Reports dealing with EDP methods have yellow (gul) covers.

Reports published 1995:

1995:1 (grön)	Asymptotic Theory for Order Sampling *(Bengt Rosén)*
1995:2 (gul)	PC-programvaror för estimation och statistisk analys *(Dan Hedlin)*
1995:3 (grön)	Vad användarna tycker om SCB - Rapport från två användarstudier med QSP *(Jan Eklöf, Mats Bergdahl)*
1995:4 *(grön)*	Bortfallsbarometern nr 10 *(Mats Bergdahl, Pär Brundell, Margareta Eriksson, Dan Hedlin, Monica Rennermalm, Per Sandgren)*
1995:5 (gul)	Statistical metadata - a tutorial *(Bo Sundgren)*
1995:6 (gul)	Making statistical data more available *(Bo Sundgren)*
1995:7 (gul)	Integration of statistical information systems - theory and practice theory and practice *(Bo Sundgren and Erik Malmborg)*
1995:8 (gul)	Statistical information systems in a modern society: roles, functions, and system designs *(Bo Sundgren)*

Tidigare utgivna R & D Reports kan beställas genom Ingvar Andersson, SCB, U/LEDN, 115 81 STOCKHOLM (telefon 08-783 41 47, telefax 08-783 45 99, E-mail ingvar.andersson@scb.se)

R & D Reports listed above as well as issues from 1988-94 can - in case they are still in stock - be ordered from Statistics Sweden, att. Ingvar Andersson U/LEDN, S-115 81 STOCKHOLM SWEDEN (telephone +46 8 783 41 47, telefax +46 8 783 45 99, E-mail ingvar.andersson@scb.se)