R&D Report 2000:5
Research – Methods – Development

Statistiska centralbyrån Statistics Sweden

Generalized Regression Estimation and Pareto πps

Bengt Rosén

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 2000:5. Generalized regression estimation and Pareto πps / Bengt Rosén.
Digitaliserad av Statistiska centralbyrån (SCB) 2016.
Vid digitalisering har rättelseblad ”Corrections..” lagts till.

urn:nbn:se:scb-2000-X101OP0005

Bengt Rosén 2000-12-29

Corrections to the Statistics Sweden R&D Report 2000:5, B. Rosén: Generalized Regression Estimation and Pareto πps

The chief reason for this correction note is that I (a bit ashamed) have to call attention to a flaw due to a programming error. Since a correction note is required, also "ordinary" printer's errors are pointed out (admitting that nowadays printer = author).

1 Correction of printer's errors (fa/fb = from above/below)

Page	Row	Reads	Should be
3	18 fa	strategy which satisfies	strategy satisfies
3	3 fb	Is yield the following	yield the following
4	14 fa	study y variable	study variable y
4	20 fb	(2.20)	(2.22)
4	17 fa	(2.20)	(2.22)
4	19 fb	(2.22) and (2.23)	(2.24) and (2.25)
4	18 fb	*(2.20) + (2.4)*	*(2.22) + (2.4)*
4	15 fb	(2.23)	(2.25)
5	3 fa	(2.21) + (2.4)	(2.22) + (2.4)
5	18 fb	(2.21)	(2.22)
5	8 fb	(2.21)	(2.22)
6	1 fb	REG	GREG
9	14 fb	(4.7)	(4.6)
9	19 fb	for belief	for the belief
9	14 fb	strategy is	strategy performance is
10	18 fa	Although	in addition to
11	8+ fa		RBPE $=(E[\hat{\tau}(y)]/\tau(y)-1)\cdot 100\%$. (6.7)
13	14 fa	har	has
13	23 fa	The optimal strategy	an optimal strategy
14			Table 6.2 should read as stated on next page.
16	1 fa	And A.6 numerical	and A.6 provide numerical
	27-32		The c - values should be read as in the subsequent Tables A.13-A.18.

2 Corrections caused by the programming error

The last sentence in the paragraph on the middle of page 16 reads as follows :

> The reason for blank columns under RESTD for spread magnitude c_2 is that we (in last minute) came to suspect a program bug, which could not be sorted out.

The suspected bug turned out to exist. Its elimination affects the RBSTD - values at the bottom of Tables A.13 - A.18. Corrected tables are presented on subsequent pages. The RBVE - values are unchanged. As a consequence of the altered figures the last paragraph on page 16 should be modified as stated below.

> **On the accuracy of approximate formulas for theoretical estimator variance :** From a practical point of view this issue is not important. The interesting aspect is if we could have dispensed of all the simulations and compared strategies by using the approximate formulas (5.3) and (4.5) instead of simulations. Tables A.13-A.18 show that the approximate variance formulas mostly work with surprisingly good accuracy, also in situations with misjudged superpopulation. However, a warning for [SRS, GREG] should be issued also here.

Corrected tables

Table 6.2. Studied strategies. True spread σ is presumed to be proportional to x^{γ}.

		Use of γ			
		In design and estimator	**In estimator only**	**In design only**	**Not at all**
Use of x	**In both design and estimator**	**Correct spread guestimate** [PAR(γ), GREG(γ)] **Overguestimated spread** [PAR($\gamma\cdot 1.2$)), GREG($\gamma\cdot 1.2$)] [PAR($\gamma\cdot 1.5$), GREG($\gamma\cdot 1.5$)] **Underguestimated spread** [PAR($\gamma\cdot 0.8$), GREG($\gamma\cdot 0.8$)] [PAR($\gamma\cdot 0.5$), GREG($\gamma\cdot 0.5$)]	[PAR(1), GREG(γ)]		
	In estimator only		[SRS, GREG(γ)]		
	In design only			**Correct spread** [PAR(γ), πps] **Overguestimated spread** [PAR($\gamma\cdot 1.2$)), πps] [PAR($\gamma\cdot 1.5$), πps] **Underguestimated spread** [PAR($\gamma\cdot 0.8$)), πps] [PAR($\gamma\cdot 0.5$), πps]	[PAR(1), πps]
	Not at all				[SRS, HT]

A.2.2 Variance estimator bias and approximation accuracy for theoretical variances

Table A.13. RBVE and RESTD (in %) for test situations of Type A. See (6.8), (6.9) and Table 6.1.

Strategy and variance estimator	n=10			n=25			n=50			n=80		
	c			c			c			c		
	0.9	1.8	3.5	0.9	1.8	3.5	0.9	1.8	3.5	0.9	1.8	3.5
					RBVE in %							
Correct spread guestimate												
[PAR(γ),GREG(γ)]/V_1	-2.9	-5.3	0.4	-2.4	5.8	-8.4	-1.3	4.0	-4.2	0.4	1.8	-8.1
[PAR(γ),GREG(γ)]/V_2	-0.4	-3.9	2.7	-1.7	6.4	-7.6	-0.9	4.2	-3.8	0.7	2.0	-7.9
[SRS,GREG(γ)]/V_1	-12.0	-14.8	-7.6	-5.1	2.9	-9.2	-1.3	2.3	-12.1	3.7	0.6	-6.3
[SRS,GREG(γ)]/V_2	-8.7	-13.8	-4.9	-3.7	3.9	-7.7	-0.8	2.7	-11.8	4.1	0.8	-6.0
[PAR(1),πps]	0.8	-3.2	-1.5	-2.0	2.2	-6.3	-0.4	3.0	-2.1	-1.2	2.5	-1.3
[PAR(γ),πps]	-0.4	-2.0	3.3	4.2	6.4	-5.6	-6.4	-1.7	-0.6	-1.2	2.5	-1.3
[SRS,HT]	2.9	2.0	4.4	2.3	1.5	-1.7	-4.2	-4.2	-4.4	-3.9	-3.6	-6.5
Overguestimated spread												
[PAR(γ·1.2)),GREG(γ·1.2)]/V_1	-2.6	-3.7	-0.6	-1.6	7.0	-9.1	-1.0	5.0	-3.9	1.9	-0.6	-7.2
[PAR(γ·1.2)),GREG(γ·1.2)]/V_2	-0.1	-2.1	1.9	-0.8	7.6	-8.2	-0.5	5.3	-3.5	2.2	-0.4	-7.0
[PAR(γ·1.5),GREG(γ·1.5)]/V_1	-0.3	-3.8	0.3	-1.8	4.7	-7.0	-2.5	3.9	-2.2	0.1	-1.5	-4.8
[PAR(γ·1.5),GREG(γ·1.5)]/V_2	2.7	-1.8	3.4	-0.7	5.6	-5.7	-1.8	4.5	-1.4	0.6	-1.1	-4.3
[PAR(γ·1.2)),πps]	-0.4	-0.5	2.7	2.6	6.9	-6.8	-5.7	0.3	-1.2	-2.1	-0.5	-5.4
[PAR(γ·1.5),πps]	0.1	-1.4	2.3	2.0	4.9	-6.6	-3.1	3.3	-0.8	-6.4	-1.0	-4.2
Underguestimated spread												
[PAR(γ·0.8),GREG(γ·0.8)]/V_1	-3.8	-7.2	-1.9	-4.0	5.8	-8.8	-1.1	4.5	-5.9	2.4	1.2	-7.2
[PAR(γ·0.8),GREG(γ·0.8)]/V_2	-1.2	-5.9	0.0	-3.2	6.4	-8.1	-0.7	4.8	-5.5	2.6	1.3	-7.1
[PAR(γ·0.5),GREG(γ·0.5)]/V_1	-6.8	-8.7	-3.7	-2.4	4.3	-7.7	-1.0	4.5	-6.9	4.8	0.0	-6.4
[PAR(γ·0.5),GREG(γ·0.5)]/V_2	-3.8	-7.2	-1.3	-1.4	5.1	-6.8	-0.6	4.8	-6.5	5.1	0.2	-6.2
[PAR(γ·0.8)),πps]	-0.6	-2.1	-0.3	2.9	4.7	0.0	-5.5	-1.7	0.0	-2.7	-3.1	0.1
[PAR(γ·0.5),πps]	0.0	-1.2	1.2	2.7	3.8	-1.4	-5.3	-2.0	-2.3	-3.2	-3.7	-2.8
					RESTD in %							
Correct spread guestimate												
[PAR(γ),GREG(γ)]	-2.0	-2.4	-0.2	-1.4	2.8	-4.2	-0.7	1.9	-2.2	0.2	0.8	-4.1
[SRS,GREG(γ)]	-3.4	-4.0	-1.1	-1.6	2.7	-3.5	-0.2	1.7	-5.8	2.1	0.7	-2.8
[PAR(1),πps]	0.1	-1.1	-0.4	-0.9	1.3	-2.5	0	1.6	-1.0	-0.2	1.4	-0.4
[PAR(γ),πps]	-0.2	-0.9	1.5	2.2	3.2	-2.6	-3.1	-0.8	-0.3	-1.9	-1.7	-3.3
[SRS,HT]	1.4	0.9	2.1	1.2	0.8	-0.7	-2.1	-2.1	-2.2	-0.6	-0.7	-0.6
Overguestimated spread												
[PAR(γ·1.2)),GREG(γ·1.2)]	-2.2	-2.1	-1.1	-1.2	3.2	-4.8	-0.6	2.3	-2.2	0.8	-0.4	-3.7
[PAR(γ·1.5),GREG(γ·1.5)]	-1.9	-2.8	-1.6	-1.6	2.0	-4.2	-1.5	1.6	-1.5	-0.1	-0.9	-2.9
[PAR(γ·1.2)),πps]	-0.1	-0.2	1.2	1.3	3.4	-3.3	-2.8	0.1	-0.6	-0.9	-0.1	-2.6
[PAR(γ·1.5),πps]	0.2	-0.7	1.4	1.1	2.4	-3.4	-1.4	1.8	-0.4	-3.0	-0.4	-2.0
Underguestimated spread												
[PAR(γ·0.8),GREG(γ·0.8)]	-1.8	-2.8	-0.8	-2.0	3.0	-4.3	-0.5	2.3	-3.0	1.2	0.6	-3.6
[PAR(γ·0.5),GREG(γ·0.5)]	-2.6	-2.6	-0.6	-0.8	2.6	-3.4	-0.3	2.4	-3.4	2.4	0.1	-3.1
[PAR(γ·0.8)),πps]	-0.2	-0.6	1.4	1.7	2.4	-1.9	-2.8	-0.9	-0.4	-1.3	-1.5	-2.1
[PAR(γ·0.5),πps]	0-0	-0.5	0.8	1.4	1.9	-0.6	-2.7	-1.0	-1.2	-1.5	-1.8	-1.4

Table A.14. RBVE and RESTD (in %) for test situations of Type B. See (6.8), (6.9) and Table 6.1.

Strategy and variance estimator	n=10			n=25			n=50			n=80		
	c			c			c			c		
	0.06	0.12	0.25	0.06	0.12	0.25	0.06	0.12	0.25	0.06	0.12	0.25
					RBVE in %							
Correct spread guestimate												
[PAR(γ),GREG(γ)]/V_1	0.2	-3.1	-1.3	-1.4	5.2	-5.4	-0.3	3.6	-3.2	-0.9	2.3	-3.9
[PAR(γ),GREG(γ)]/V_2	0.2	-3.1	-1.3	-1.4	5.2	-5.4	-0.3	3.6	-3.2	-0.9	2.3	-3.9
[SRS,GREG(γ)]/V_1	-11.8	-15.5	.7.5	-5.7	2.8	-8.0	-1.3	1.0	-14.2	3.1	-0.6	-5.5
[SRS,GREG(γ)]/V_2	-10.8	-17.0	-7.1	-4.9	3.2	-7.1	-1.1	1.2	-13.9	3.4	-0.4	-5.4
[PAR(1),πps]	0.2	-3.1	-1.3	-1.4	5.2	-5.4	-0.3	3.6	-3.2	-0.3	3.6	-3.2
[PAR(γ),πps]	0.2	-3.1	-1.3	-1.4	5.2	-5.4	-0.3	3.6	-3.2	-0.3	3.6	-3.2
[SRS,HT]	2.8	2.3	4.5	2.0	1.5	-1.4	-4.4	-4.5	-5.3	-1.8	-1.9	-1.5
Overguestimated spread												
[PAR(γ·1.2)),GREG(γ·1.2)]/V_1	-3.4	-4.4	-1.9	-3.0	3.2	-4.7	-3.0	2.8	-3.2	-1.2	2.1	-2.1
[PAR(γ·1.2)),GREG(γ·1.2)]/V_2	-2.8	-3.7	-1.4	-2.8	3.5	-4.7	-3.0	2.9	-3.3	-1.2	2.1	-2.3
[PAR(γ·1.5),GREG(γ·1.5)]/V_1	-7.4	-9.3	-8.9	-6.4	0.0	-7.7	-7.0	0.2	-5.3	-3.9	-1.1	-4.4
[PAR(γ·1.5),GREG(γ·1.5)]/V_2	-5.6	-7.0	-7.0	-5.7	1.0	-7.2	-6.4	0.6	-5.4	-3.4	-1.1	-4.7
[PAR(γ·1.2)),πps]	2.0	-1.3	-1.4	1.6	5.0	-2.1	-3.4	0.0	-4.2	3.4	2.3	-1.5
[PAR(γ·1.5),πps]	2.5	1.5	-3.2	1.5	3.5	-0.4	-3.3	-2.5	-2.8	0.7	0.1	-1.6
Underguestimated spread												
[PAR(γ·0.8),GREG(γ·0.8)]/V_1	-1.6	-3.0	0.9	-0.9	6.7	-6.8	-1.0	3.1	-3.8	-1.7	-2.1	-5.8
[PAR(γ·0.8),GREG(γ·0.8)]/V_2	-1.6	-3.4	0.8	-0.9	6.5	-6.9	-1.0	3.1	-3.7	-1.7	-2.1	-5.8
[PAR(γ·0.5),GREG(γ·0.5)]/V_1	-2.5	-5.5	0.9	-3.3	5.5	-7.8	0.0	4.7	-7.1	1.1	1.8	-8.4
[PAR(γ·0.5),GREG(γ·0.5)]/V_2	-2.2	-6.0	1.0	-3.2	5.4	-7.7	0.0	4.6	-7.1	1.2	1.8	-8.4
[PAR(γ·0.8)),πps]	-0.1	-2.4	-0.1	0.6	4.8	0.0	-2.2	2.7	0.0	-5.6	-2.2	0.0
[PAR(γ·0.5),πps]	-0.1	-1.7	4.0	4.0	6.5	-4.4	-6.7	-3.3	-2.8	-4.0	-3.9	-6.8
					RESTD in %							
Correct spread guestimate												
[PAR(γ),GREG(γ)]	0.0	-1.5	-0.6	-0.8	2.4	-2.7	0.0	1.7	-1.6	-0.4	1.1	-2.0
[SRS,GREG(γ)]	-2.1	-3.3	0.0	-1.5	3.2	-2.5	0.1	1.4	-6.7	2.0	0.3	-2.3
[PAR(1),πps]	0.0	-1.5	-0.6	-0.8	2.4	-2.7	0.0	1.7	-1.6	-0.4	1.1	-2.0
[PAR(γ),πps]	0.0	-1.5	-0.6	-0.8	2.4	-2.7	0.0	1.7	-1.6	-0.4	1.1	-2.0
[SRS,HT]	1.3	1.0	2.1	1.0	0.8	-0.5	-2.2	-2.2	-2.6	-0.9	-0.9	-0.7
Overguestimated spread												
[PAR(γ·1.2)),GREG(γ·1.2)]	-0.4	-1.3	0.3	-1.1	1.9	-1.6	-1.0	1.8	-1.1	-0.1	1.5	-0.5
[PAR(γ·1.5),GREG(γ·1.5)]	1.0	-0.9	0.4	-0.4	2.6	-0.6	-1.2	2.2	0.2	0.6	1.9	1.1
[PAR(γ·1.2)),πps]	2.2	-0.2	-0.5	1.3	2.7	-0.8	-1.4	0.2	-2.0	2.1	1.4	-0.5
[PAR(γ·1.5),πps]	2.4	1.2	-1.0	2.0	3.1	0.7	-0.7	-0.1	-0.3	0.6	0.2	-0.6
Underguestimated spread												
[PAR(γ·0.8),GREG(γ·0.8)]	-1.0	-1.4	0.2	-0.6	3.2	-3.5	-0.5	1.4	-2.0	-0.9	-1.1	-3.0
[PAR(γ·0.5),GREG(γ·0.5)]	-0.7	-1.7	0.9	-1.5	3.0	-3.6	0.1	2.4	-3.5	0.6	0.9	-4.3
[PAR(γ·0.8)),πps]	0.0	-1.2	1.5	0.6	2.5	-3.6	-1.0	1.3	-1.1	-2.6	-1.0	-3.0
[PAR(γ·0.5),πps]	-0.1	-0.8	1.8	2.1	3.3	-2.0	-3.3	-1.6	-1.4	-1.9	-1.9	-3.4

Table A.15. RBVE and RESTD (in %) for test situations of Type C. See (6.8), (6.9) and Table 6.1.

Strategy and variance estimator	n=10			n=25			n=50			n=80		
	c			C			c			c		
	3	7	13	3	7	13	3	7	13	3	7	13
					RBVE in %							
Correct spread guestimate												
[PAR(γ),GREG(γ)]/V_1	-5.9	-8.3	-4.1	-1.9	4.4	-6.7	-2.0	4.2	-4.7	5.4	0.9	-6.3
[PAR(γ),GREG(γ)]/V_2	-1.9	-5.7	-0.5	-0.6	5.6	-5.4	-1.5	4.7	-4.2	5.8	1.2	-6.0
[SRS,GREG(γ)]/V_1	-10.6	-14.3	-7.4	-5.3	2.8	-8.0	-1.5	2.7	-10.1	4.7	1.7	-6.8
[SRS,GREG(γ)]/V_2	-6.2	-11.9	-3.5	-3.7	4.1	-6.4	-0.8	3.2	-9.5	5.2	2.1	-6.5
[PAR(1),πps]	1.9	-1.1	-2.1	-3.4	0.1	-5.7	1.0	2.0	-2.2	-1.2	2.3	0.5
[PAR(γ),πps]	0.4	-1.0	0.7	2.6	3.2	-1.1	-5.9	-1.5	-1.3	-3.3	-3.6	-3.8
[SRS,HT]	3.2	2.0	4.0	1.9	0.8	-2.1	-4.8	-4.3	-4.0	-1.9	-1.9	-2.9
Overguestimated spread												
[PAR(γ·1.2)),GREG(γ·1.2)]/V_1	-4.8	-7.9	-3.8	-3.8	5.1	-7.1	-2.1	4.9	-5.2	4.6	2.3	-5.8
[PAR(γ·1.2)),GREG(γ·1.2)]/V_2	-0.9	-5.3	-0.3	-2.4	6.2	-5.8	-1.5	5.4	-4.6	5.0	2.6	-5.5
[PAR(γ·1.5),GREG(γ·1.5)]/V_1	-3.8	-7.3	-2.4	-4.2	5.8	-7.7	-2.7	4.5	-3.8	2.7	0.2	-6.8
[PAR(γ·1.5),GREG(γ·1.5)]/V_2	0.2	-4.5	1.2	-3.0	6.8	-6.4	-2.2	4.9	-3.3	3.0	0.4	-6.5
[PAR(γ·1.2)),πps]	0.8	-0.6	1.2	2.3	3.9	-2.2	-7.9	-2.5	-2.7	-2.7	-4.3	-3.3
[PAR(γ·1.5),πps]	-0.2	-1.3	1.7	2.6	4.1	-3.6	-7.1	-1.3	-0.7	-2.8	-3.6	-4.2
Underguestimated spread												
[PAR(γ·0.8),GREG(γ·0.8)]/V_1	-5.7	-9.5	-4.7	-3.6	3.8	-7.1	-1.9	3.7	-5.5	4.8	1.8	-5.9
[PAR(γ·0.8),GREG(γ·0.8)]/V_2	-1.7	-7.0	-1.2	-2.3	4.8	-5.8	-1.3	4.2	-5.0	5.2	2.1	-5.6
[PAR(γ·0.5),GREG(γ·0.5)]/V_1	-8.2	-10.5	-6.5	-3.3	3.7	-7.2	-1.8	4.4	-6.1	3.6	3.1	-5.8
[PAR(γ·0.5),GREG(γ·0.5)]/V_2	-4.0	-8.0	-2.9	-1.9	4.8	-5.9	-1.2	4.9	-5.5	4.0	3.4	-5.5
[PAR(γ·0.8)),πps]	-0.6	-1.3	-0.2	0.7	1.3	0.0	-4.8	-1.1	0.0	-3.4	-3.5	0.0
[PAR(γ·0.5),πps]	1.7	0.9	1.6	1.4	1.7	-1.5	-4.6	-1.5	-1.5	-2.5	-3.2	-3.3
					RESTD in %							
Correct spread guestimate												
[PAR(γ),GREG(γ)]	-2.9	-3.2	-1.5	-0.8	2.4	-3.0	-0.9	2.2	-2.3	2.7	0.5	-3.2
[SRS,GREG(γ)]	-3.5	-4.7	-1.7	-2.0	2.4	-3.2	-0.4	1.7	-4.8	2.5	1.1	-3.2
[PAR(1),πps]	-0.7	0.6	-0.3	-0.6	1.0	0.0	0.6	1.2	-1.4	0.0	1.3	0.7
[PAR(γ),πps]	0.1	-0.5	0.6	1.4	1.6	-0.4	-2.9	-0.7	-0.6	-1.6	-1.8	-1.9
[SRS,HT]	1.5	0.8	2.1	1.0	0.8	-0.9	-2.4	-2.1	-2.0	-0.9	-0.9	-1.4
Overguestimated spread												
[PAR(γ·1.2)),GREG(γ·1.2)]	-2.6	-3.3	-1.6	-1.8	2.6	-3.4	-1.0	2.4	-2.6	2.2	1.1	-3.0
[PAR(γ·1.5),GREG(γ·1.5)]	-2.5	-3.5	-1.6	-2.3	2.7	-3.9	-1.4	2.2	-2.0	1.3	0.0	-3.5
[PAR(γ·1.2)),πps]	0.3	-0.2	1.0	1.2	1.9	-0.9	-4.0	-1.2	-1.3	-1.3	-2.2	-1.7
[PAR(γ·1.5),πps]	-0.1	-0.5	1.1	1.4	2.1	-1.5	-3.5	-0.6	-0.3	-1.3	-1.8	-2.1
Underguestimated spread												
[PAR(γ·0.8),GREG(γ·0.8)]	-2.6	-3.6	-1.5	-1.5	2.2	-3.1	-0.8	1.9	-2.6	2.4	0.9	-2.9
[PAR(γ·0.5),GREG(γ·0.5)]	-3.2	-3.6	-2.1	-1.3	2.4	-3.1	-0.7	2.4	-2.9	1.9	1.7	-2.8
[PAR(γ·0.8)),πps]	-0.2	-0.6	0.7	0.4	0.7	-1.0	-2.3	-0.5	-0.2	-1.7	-1.8	-2.0
[PAR(γ·0.5),πps]	0.8	0.3	1.0	0.7	0.8	-0.6	-2.2	-0.7	-0.7	-1.2	-1.6	-1.6

Table A.16. RBVE and RESTD (in %) for test situations of Type D. See (6.8), (6.9) and Table 6.1.

Strategy and variance estimator	n=10			n=25			n=50			n=80		
	c			c			c			c		
	2.5	5	10	2.5	5	10	2.5	5	10	2.5	5	10
				RBVE in %								
Correct spread guestimate												
[PAR(γ), GREG(γ)]/V_1	-4.1	-5.9	0.0	0.3	5.4	-8.0	-5.0	4.7	-4.6	-2.4	-0.5	-8.7
[PAR(γ), GREG(γ)]/V_2	-0.7	-4.1	2.4	1.5	6.2	-7.2	-4.5	5.0	-4.2	-2.0	-0.2	-8.5
[SRS, GREG(γ)]/V_1	-15.3	-16.2	-8.0	-7.5	0.9	-8.5	-3.2	2.7	-12.9	-1.9	0.0	-7.0
[SRS, GREG(γ)]/V_2	-11.5	-14.7	-5.1	-5.6	2.1	-7.1	-2.4	3.2	-12.2	-1.5	0.3	-6.8
[PAR(1), πps]	-2.8	-3.4	-1.1	-1.2	1.9	-4.8	-4.7	5.2	-2.2	-5.0	3.2	-1.3
[PAR(γ), πps]	0.4	-1.1	3.0	4.0	5.4	-4.6	-7.6	-3.3	-2.6	-3.3	-3.2	-6.6
[SRS, HT]	2.7	2.1	3.9	2.2	1.3	-1.7	-4.8	-4.6	-5.2	-1.8	-1.7	-2.5
Overguestimated spread												
[PAR(γ·1.2)), GREG(γ·1.2)]/V_1	-3.4	-3.2	-0.5	-0.5	6.9	-8.4	-4.8	6.1	-4.3	-0.8	-0.5	-7.6
[PAR(γ·1.2)), GREG(γ·1.2)]/V_2	0.0	-1.3	2.1	0.8	7.8	-7.5	-4.1	6.5	-3.8	-0.5	-0.3	-7.3
[PAR(γ·1.5), GREG(γ·1.5)]/V_1	0.2	-2.5	1.0	0.8	5.1	-5.7	-3.1	5.9	-2.6	-4.9	-1.2	-4.7
[PAR(γ·1.5), GREG(γ·1.5)]/V_2	3.7	-0.2	4.1	2.1	6.1	-4.4	-2.3	6.5	-1.8	-4.4	-0.8	-4.3
[PAR(γ·1.2)), πps]	0.4	0.4	2.4	2.2	5.3	-5.6	-7.1	-1.4	-2.9	-1.2	0.0	-5.3
[PAR(γ·1.5), πps]	1.5	0.5	2.8	2.4	4.3	-5.0	-4.6	1.3	-2.3	-5.2	-1.4	-4.2
Underguestimated spread												
[PAR(γ·0.8), GREG(γ·0.8)]/V_1	-7.4	-8.0	-2.5	-2.2	4.7	-7.9	-3.9	5.6	-6.1	-1.4	-0.5	-7.6
[PAR(γ·0.8), GREG(γ·0.8)]/V_2	-4.0	-6.3	-0.3	-1.0	5.4	-7.1	-3.4	5.9	-5.8	-1.1	-0.3	-7.4
[PAR(γ·0.5), GREG(γ·0.5)]/V_1	-9.6	-9.5	-4.6	-4.1	3.0	-6.7	-3.9	4.9	-8.1	0.6	-1.9	-6.5
[PAR(γ·0.5), GREG(γ·0.5)]/V_2	-5.8	-7.8	-2.2	-2.7	3.9	-5.8	-3.3	5.2	-7.7	0.9	-1.6	-6.3
[PAR(γ·0.8)), πps]	0.1	-1.1	-0.4	2.6	3.8	0.1	-6.1	-2.5	0.0	-2.4	-2.6	0.1
[PAR(γ·0.5), πps]	0.7	-0.4	1.4	2.6	3.3	-0.8	-4.9	-2.2	-3.5	-3.1	-3.7	-3.4
				RESTD in %								
Correct spread guestimate												
[PAR(γ), GREG(γ)]	-3.0	-3.0	-0.5	-0.2	2.6	-3.9	-2.6	2.2	-2.3	-1.3	-0.3	-4.5
[SRS, GREG(γ)]	1.4	-4.1	-0.6	4.1	2.8	-2.3	6.1	3.0	-5.3	6.6	1.4	-2.4
[PAR(1), πps]	-1.1	-1.5	0.0	0.1	1.3	-1.8	-1.9	2.7	-1.0	-2.2	1.8	-0.4
[PAR(γ), πps]	0.1	-0.5	1.4	2.1	2.8	-2.1	-3.7	-1.6	-1.2	-1.6	-1.6	-3.3
[SRS, HT]	1.3	0.9	2.0	1.1	0.7	-0.7	-2.4	-2.3	-2.6	-0.9	-0.8	-1.2
Overguestimated spread												
[PAR(γ·1.2)), GREG(γ·1.2)]	-3.8	-2.4	-1.3	-1.7	2.9	-4.5	-3.6	2.5	-2.4	-1.6	-0.6	-4.0
[PAR(γ·1.5), GREG(γ·1.5)]	-3.3	-2.8	-1.2	-2.2	1.6	-3.7	-3.7	2.1	-1.9	-4.5	-1.2	-2.8
[PAR(γ·1.2)), πps]	0.3	0.2	1.1	1.2	2.7	-2.6	-3.6	-0.7	-1.4	-0.5	0.1	-2.5
[PAR(γ·1.5), πps]	0.8	0.1	1.7	1.3	2.2	-2.5	-2.2	0.8	-1.1	-2.4	-0.6	-2.0
Underguestimated spread												
[PAR(γ·0.8), GREG(γ·0.8)]	-2.9	-3.2	-1.0	0.1	2.7	-3.6	-0.7	3.0	-2.9	0.6	0.0	-3.7
[PAR(γ·0.5), GREG(γ·0.5)]	-1.0	-2.8	-0.8	1.6	2.6	-2.4	1.6	3.2	-3.5	4.0	-0.2	-2.8
[PAR(γ·0.8)), πps]	0.1	-0.4	1.3	1.5	2.0	-1.4	-3.0	-1.3	-1.2	-1.2	-1.3	-2.3
[PAR(γ·0.5), πps]	0.3	-0.2	0.8	1.3	1.7	-0.3	-2.5	-1.1	-1.7	-1.5	-1.6	-1.7

Table A.17. RBVE and RESTD (in %) for test situations of Type E. See (6.8), (6.9) and Table 6.1.												
Strategy and variance estimator	n=10			n=25			n=50			n=80		
	c			c			c			c		
	12	25	50	12	25	50	12	25	50	12	25	50
				RBVE in %								
Correct spread guestimate												
[PAR(γ), GREG(γ)]/V_1	-3.4	-5.6	-0.8	0.8	3.9	-7.6	-6.5	0.8	-1.9	-1.6	-1.8	-8.4
[PAR(γ), GREG(γ)]/V_2	0.3	-3.1	1.8	2.2	5.0	-6.7	-5.9	1.2	-1.5	-1.2	-1.5	-8.1
[SRS, GREG(γ)]/V_1	-19.0	-18.0	-9.7	-7.5	-2.1	-8.3	-6.0	-2.0	-10.3	-3.0	-1.0	-7.3
[SRS, GREG(γ)]/V_2	-14.4	-15.2	-6.5	-4.9	-0.4	-6.7	-4.8	-1.2	-9.7	-2.4	-0.5	-7.0
[PAR(1), πps]	-2.7	-2.7	-0.8	0.9	2.0	-3.6	-5.2	3.6	-0.7	-3.1	3.3	-1.1
[PAR(γ), πps]	0.9	-0.8	2.5	3.4	4.6	-4.2	-7.3	-3.6	-1.5	-2.4	-2.9	-6.2
[SRS, HT]	2.1	1.4	3.1	2.1	1.1	-1.7	-4.9	-4.8	-4.8	-1.1	-1.0	-2.3
Overguestimated spread												
[PAR(γ·1.2)), GREG(γ·1.2)]/V_1	-1.9	-2.2	-0.8	-0.2	4.8	-7.6	-6.5	1.3	-2.2	0.0	0.1	-7.2
[PAR(γ·1.2)), GREG(γ·1.2)]/V_2	1.6	0.3	1.9	1.1	5.9	-6.7	-5.9	1.8	-1.7	0.4	0.4	-6.9
[PAR(γ·1.5), GREG(γ·1.5)]/V_1	1.7	-0.3	1.2	2.0	4.5	-4.7	-5.1	2.6	-1.1	-4.1	-1.0	-4.7
[PAR(γ·1.5), GREG(γ·1.5)]/V_2	4.7	2.2	4.3	3.2	5.5	-3.5	-4.5	3.2	-0.4	-3.7	-0.6	-4.2
[PAR(γ·1.2)), πps]	0.9	0.6	1.9	1.6	4.1	-5.1	-7.1	-2.7	-2.0	-0.5	0.4	-4.9
[PAR(γ·1.5), πps]	1.7	1.2	2.5	2.4	3.8	-4.1	-5.6	-0.3	-1.5	-4.2	-1.3	-4.1
Underguestimated spread												
[PAR(γ·0.8), GREG(γ·0.8)]/V_1	-7.0	-7.8	-3.3	-1.4	2.6	-7.1	-5.5	2.2	-3.3	-1.2	-1.5	-7.0
[PAR(γ·0.8), GREG(γ·0.8)]/V_2	-3.1	-5.4	-0.9	0.0	3.7	-6.3	-4.8	2.7	-2.9	-0.9	-1.2	-6.9
[PAR(γ·0.5), GREG(γ·0.5)]/V_1	-10.0	-9.7	-5.9	-3.9	0.8	-6.3	-6.2	1.7	-5.6	-0.9	-3.3	-6.0
[PAR(γ·0.5), GREG(γ·0.5)]/V_2	-5.4	-7.0	-3.2	-2.0	2.1	-5.3	-5.5	2.2	-5.1	-0.5	-3.0	-5.8
[PAR(γ·0.8)), πps]	0.5	-0.8	-0.4	2.0	3.0	0.1	-5.9	-2.5	0.0	-1.7	-2.1	0.1
[PAR(γ·0.5), πps]	0.8	-0.3	1.0	2.2	2.7	-1.0	-5.8	-2.8	-2.7	-2.7	-3.4	-3.2
				RESTD in %								
Correct spread guestimate												
[PAR(γ), GREG(γ)]	-2.6	-3.1	-0.9	0.1	1.9	-3.7	-3.4	0.3	-1.1	-0.9	-1.0	-4.3
[SRS, GREG(γ)]	5.2	-2.0	0.2	10.7	5.1	-0.3	11.3	4.5	-2.1	12.9	4.9	-0.7
[PAR(1), πps]	-1.0	-1.2	0.1	0.9	1.4	-1.3	-2.4	1.9	-0.3	-1.2	1.8	-0.2
[PAR(γ), πps]	0.4	-0.4	1.2	1.7	2.4	-1.9	-3.6	-1.8	-0.7	-1.2	-1.4	-3.1
[SRS, HT]	1.0	0.6	1.6	1.0	0.6	-0.7	-2.4	-2.4	-2.4	-0.6	-0.5	-1.1
Overguestimated spread												
[PAR(γ·1.2)), GREG(γ·1.2)]	-3.7	-2.6	-1.7	-2.2	1.3	-4.4	-5.2	-0.4	-1.7	-1.9	-0.8	-4.1
[PAR(γ·1.5), GREG(γ·1.5)]	-3.5	-2.8	-1.4	-2.6	0.3	-3.7	-5.8	-0.3	-1.6	-5.2	-2.1	-3.2
[PAR(γ·1.2)), πps]	0.5	0.3	0.9	0.8	2.1	-2.4	-3.6	-1.4	-1.0	-0.2	0.3	-2.4
[PAR(γ·1.5), πps]	0.9	0.5	1.6	1.2	1.9	-2.1	-2.8	0.0	-0.7	-1.9	-0.5	-2.0
Underguestimated spread												
[PAR(γ·0.8), GREG(γ·0.8)]	-1.9	-2.8	-1.1	1.3	2.3	-2.9	-0.7	2.0	-1.1	1.6	0.2	-3.1
[PAR(γ·0.5), GREG(γ·0.5)]	1.2	-1.6	-0.7	4.4	3.3	-1.3	3.0	3.5	-1.4	6.0	0.9	-1.6
[PAR(γ·0.8)), πps]	0.3	-0.3	1.1	1.1	1.6	-1.3	-3.0	-1.3	-0.6	-0.8	-1.1	-2.1
[PAR(γ·0.5), πps]	0.4	-0.2	0.7	1.1	1.4	-0.4	-2.9	-1.4	-1.4	-1.3	-1.7	-1.6

Table A.18. RBVE and RESTD in % for test situations of Type F. See (6.8), (6.9) and Table 6.1.												
	n=10			n=25			n=50			n=80		
Strategy and	c			c			c			c		
variance estimator	0.2	0.4	0.7	0.2	0.4	0.7	0.2	0.4	0.7	0.2	0.4	0.7
				RBVE in %								
Correct spread guestimate												
[PAR(γ), GREG(γ)]/V_1	-0.7	-3.0	-1.2	-0.6	3.4	-4.4	-5.4	-2.8	-8.5	-0.3	1.9	-6.2
[PAR(γ), GREG(γ)]/V_2	6.3	0.9	2.1	1.8	4.8	-3.2	-4.3	-2.2	-8.0	0.4	2.3	-5.8
[SRS, GREG(γ)]/V_1	-17.7	-15.4	-10.1	-3.4	2.5	-7.0	-6.3	0.1	-13.1	5.3	1.7	-5.8
[SRS, GREG(γ)]/V_2	-10.2	-12.5	-6.6	-0.3	4.3	-5.3	-5.0	0.8	-12.4	6.0	2.1	-5.5
[PAR(1), πps]	1.8	-0.4	-3.5	2.8	3.3	-3.5	-2.7	-3.2	-4.3	2.1	-0.4	-1.5
[PAR(γ), πps]	-1.1	-4.2	3.1	4.0	7.6	-7.3	-6.5	2.2	-0.9	-4.6	-1.6	-7.9
[SRS, HT]	3.8	2.5	5.6	2.3	1.7	-2.3	-4.5	-4.0	-4.6	-2.9	-2.4	-3.1
Overguestimated spread												
[PAR(γ·1.2)), GREG(γ·1.2)]/V_1	2.2	-1.7	-1.4	0.2	3.7	-5.3	-5.5	-2.9	-8.0	1.7	-0.6	-5.5
[PAR(γ·1.2)), GREG(γ·1.2)]/V_2	9.8	2.6	2.1	2.9	5.4	-3.8	-4.2	-2.0	-7.2	2.5	-0.1	-5.0
[PAR(γ·1.5), GREG(γ·1.5)]/V_1	8.1	1.1	0.1	0.4	2.7	-3.4	-6.9	-4.5	-6.3	-0.2	-3.8	-5.1
[PAR(γ·1.5), GREG(γ·1.5)]/V_2	17.0	6.8	5.4	4.0	5.3	-1.0	-5.1	-3.2	-4.9	1.1	-2.8	-4.2
[PAR(γ·1.2)), πps]	-2.8	-3.1	1.2	0.3	8.4	-8.3	-3.6	4.8	-2.1	-2.1	-0.2	-7.2
[PAR(γ·1.5), πps]	1.6	-4.2	-0.9	-2.6	4.5	-6.1	-5.1	1.6	-3.0	2.2	-2.3	-5.3
Underguestimated spread												
[PAR(γ·0.8), GREG(γ·0.8)]/V_1	-1.5	-4.9	-3.0	-1.7	3.8	-6.1	-3.4	-1.5	-9.3	1.2	1.9	-6.3
[PAR(γ·0.8), GREG(γ·0.8)]/V_2	5.3	-1.4	-0.1	0.6	5.1	-5.0	-2.4	-0.9	-8.8	1.7	2.2	-6.1
[PAR(γ·0.5), GREG(γ·0.5)]/V_1	-6.0	-8.0	-3.6	0.1	3.6	-5.6	-4.4	-1.2	-8.9	1.1	1.3	-6.2
[PAR(γ·0.5), GREG(γ·0.5)]/V_2	1.2	-4.4	-0.4	2.6	5.1	-4.3	-3.4	-0.6	-8.4	1.7	1.6	-5.9
[PAR(γ·0.8)), πps]	-1.2	-4.1	-0.2	2.9	6.0	0.1	-6.4	1.1	0.0	-3.5	-2.1	0.0
[PAR(γ·0.5), πps]	-0.1	-2.2	2.2	3.0	4.7	-2.4	-5.6	0.1	-1.8	-3.8	-3.3	-3.9
				RESTD in %								
Correct spread guestimate												
[PAR(γ), GREG(γ)]	-2.6	-2.5	-1.5	-1.0	1.4	-2.3	-3.0	-1.6	-4.4	-0.4	0.7	-3.2
[SRS, GREG(γ)]	8.1	2.0	0.4	16.7	10.2	0.9	15.0	8.3	-3.1	21.8	9.0	0.6
[PAR(1), πps]	0.8	0.4	-1.4	1.9	2.5	-0.3	-0.9	-1.1	-2.0	1.6	0.1	-0.4
[PAR(γ), πps]	-0.6	-1.9	1.5	2.0	3.8	-3.4	-3.1	1.2	-0.4	-2.3	-0.8	-4.0
[SRS, HT]	1.8	1.1	2.7	1.2	0.9	-1.0	-2.2	-2.0	-2.3	-1.4	-1.2	-1.5
Overguestimated spread												
[PAR(γ·1.2)), GREG(γ·1.2)]	-4.0	-3.6	-2.5	-3.1	0.0	-3.5	-5.5	-3.0	-4.9	-1.7	-1.8	-3.6
[PAR(γ·1.5), GREG(γ·1.5)]	-5.0	-4.6	-4.4	-5.1	-1.8	-3.5	-8.0	-5.3	-5.0	-4.5	-4.6	-4.1
[PAR(γ·1.2)), πps]	-1.1	-1.4	0.7	0.2	4.1	-4.0	-1.7	2.4	-1.0	-0.9	-0.1	-3.6
[PAR(γ·1.5), πps]	0.5	-1.8	-0.3	-1.1	2.6	-2.8	-2.3	0.8	-1.5	1.2	-1.1	-2.6
Underguestimated spread												
[PAR(γ·0.8), GREG(γ·0.8)]	0.4	-1.4	-1.2	1.5	3.2	-2.3	0.9	0.7	-4.1	3.4	2.4	-2.6
[PAR(γ·0.5), GREG(γ·0.5)]	3.7	0.4	0.6	7.8	5.9	-0.7	5.7	3.3	-2.7	8.8	4.7	-1.4
[PAR(γ·0.8)), πps]	-0.4	-1.8	1.6	1.7	3.0	-2.6	-3.1	0.6	-0.6	-1.6	-1.0	-2.9
[PAR(γ·0.5), πps]	-0.1	-1.0	1.3	1.5	2.4	-1.1	-2.8	0.1	-0.9	-1.8	-1.7	-1.9

R&D Report **2000:5**
Research – Methods – Development

Generalized Regression Estimation and Pareto πps

Bengt Rosén

R&D Report 2000:5
Research - Methods - Development

Generalized Regression Estimation and Pareto πps

Från trycket November 2000

Producent Statistiska centralbyrån, *Statistics Sweden*, metodenheten
Box 24300, *SE*-104 51 STOCKHOLM

Förfrågningar Bengt Rosén
bengt.rosen@scb.se
telefon 08- 506 944 90
telefax 08- 667 77 88

ISSN 0283-8680

Printed in Sweden
SCB-Tryck, Örebro 2000

November 2000

Generalized Regression Estimation and Pareto πps

Bengt Rosén

Abstract

The topic is encounter between ***generalized regression estimation*** (GREG) and ***Pareto πps*** (πps for probability proportional to size), for a general and a special reason. The former is that GREG is a way to employ auxiliary information which can be used for *any* probability sample design. It is of interest to see what it leads to for a particular design as Pareto πps.

The special reason is as follows. The embryo to the GREG estimator was presented by Cassel et al. (1976), where it appeared as a proxy for the estimator part in an ***optimal sampling - estimation strategy***, strategy standing for a pair [sample design, estimator]. They showed that a strategy is optimal if the sample design belongs to a specific class of πps schemes and the estimator is what can be characterized as a "forerunner" to the GREG estimator. After Cassel et al. (1976) much effort has been devoted to the estimator part of the optimal strategy but only little to the design part, the πps scheme. A possible reason may be shortage of πps schemes with attractive properties. However, at least in the author's opinion, Pareto πps is such a πps scheme. Hence, it is of interest to revisit the optimal strategy problem by studying the performance of strategies of type [Pareto πps, GREG]. On the basis of the Cassel et al. results this strategy is conjectured to be close to optimal. The main conclusion from the findings is that they strongly support the conjecture.

CONTENTS

Generalized Regression Estimation and Pareto πps

1 Introduction and outline

The theme in the paper is encounter between ***generalized regression estimation*** (GREG), for which Särndal et al. (1992) provide a basic and excellent source, and ***Pareto πps*** (with πps for sampling with probability proportional to given sizes), presented in Rosén (1997). There is a general as well as a special reason why this encounter is of interest.

The general reason is that GREG is an approach to exploit auxiliary information, which can be applied for any probability sample design. It is of interest to see what it leads to for the particular design Pareto πps.

The special reason is as follows. The embryo to the GREG estimator was presented in Cassel et al. (1976), as a proxy for the estimator part in an ***optimal sampling-estimation strategy***. Here "strategy" stands for a pair [sample design, estimator] and optimal relates to expected variance as specified in Section 2.3. Cassel et al. showed that optimal strategies are characterized by sample design being a specific type of πps schemes and the estimator a "forerunner" to the GREG estimator.

Since Cassel et al. (1976) much effort has been devoted to the estimator part of the optimal strategy but only little to the design part, the πps scheme. A possible reason may be shortage of πps schemes with good properties. However, at least in the author's opinion, Pareto πps is a πps scheme with attractive properties. It has fixed sample size, simple sample selection, yields good estimation precision, admits objective assessment of sampling errors (i.e consistent variance estimation) and allows sample coordination by (permanent) random numbers. It is therefore of interest to revisit the optimal strategy problem, by studying the performance of strategies of type [Pareto πps, GREG]. This is the chief task in the paper.

General regression estimation belongs to the realm "inference under a superpopulation model". In that context we follow Särndal et al. (1992) and confine to *model assisted* (in contrast to *model dependent*) inference, thereby having a "safety-net" if the model is misjudged. In particular, for an estimator to be regarded as admissible it should be design unbiased (at least have negligible bias). The chief role of the superpopulation model is to (hopefully) guide to estimators with good precision.

The paper is organized as follows. To make it fairly self-contained, the next three sections give brief reviews of certain basic concepts and results; Sections 2 as regards optimal strategies, based on Cassel et al. (1976). Sections 3 as regards GREG estimation, based on Särndal et al. (1992). Sections 4 as regards Pareto πps, based on Rosén (1997). Strategies of type [Pareto πps, GREG] are introduced and studied in Sections 5 and 6. In particular, Section 6 describes a numerical study relating to the optimal strategy problem, with findings presented in the Appendix.

2 On inference from sample surveys

The frame-work for the subsequent inference considerations is as follows. We consider ***list sampling***, i.e. sampling frame and population one-to-one correspond. Probability sampling ***without replacement*** (wor for short) and with ***fixed sample size*** n, is employed. Ideal data collection conditions prevail, full response and no frame and/or measurement problems.

As stated above, notions and results in this Section 2 stem from Cassel et al. (1976).

2.1 Stochastic models for sample data

2.1.1 Sample design

U= $(1,2,\dots,N)$ denotes the finite population The distribution of the ***sample inclusion indicators*** $\mathbf{I}=(I_1,I_2,\dots,I_N)$ is referred to as the (sample) ***design***, and is denoted by P. Expectation and variance with respect to P are denoted by E and V, and ***inclusion probabilities*** by $\pi_k=E[I_k]$ and $\pi_{kj}=E[I_k\cdot I_j]$. Since fixed sample size is presumed, the following relations holds;

$$\pi_1+\pi_2+\dots+\pi_N=I_1+I_2+\dots+I_N=n. \tag{2.1}$$

For $\mathbf{s}=(s_1,s_2,\dots,s_N)$, $s_k>0$, a ***probability proportional to sizes* s *design***, is specified by (2.2) below. It requires that the sizes **s** are known at least up to a proportionality factor. Such a design is referred to as a ***πps***(*s*) design.

$$\pi_k \text{ is proportional to } s_k, \quad \text{i.e. } \pi_k=n\cdot s_k/\sum_{j=1}^{N}s_j,\quad k=1,2,\dots,N. \tag{2.2}$$

In the following is presumed that (2.2) leads to fulfillment of $\pi_k<1$, $k=1,2,\dots,N$. (If not, some adjusting step has to be taken, e.g. introduction of a "take for certain" stratum.)

2.1.2 Superpopulations

The values of the study variable $\mathbf{y}=(y_1,y_2,\dots,y_N)$ are seen as random. Their distribution, the ***superpopulation***, is denoted by $\mathcal{P}$. Expectation, variance and covariance with respect to $\mathcal{P}$ are denoted by $\mathcal{E}$, $\mathcal{V}$ and $\mathcal{C}$. In the sequel we confine to the simple superpopulation model below, where $\boldsymbol{\mu}=(\mu_1,\mu_2,\dots,\mu_N)$ and $\boldsymbol{\sigma}=(\sigma_1,\sigma_2,\dots,\sigma_N)$ are constants while $\boldsymbol{\varepsilon}=(\varepsilon_1,\varepsilon_2,\dots,\varepsilon_N)$ are random variables. The σ-values are referred to as (*superpopulation*) ***spread parameters.***

$$y_k=\mu_k+\varepsilon_k,\quad k=1,2,\dots,N, \tag{2.3}$$

$$\mathcal{E}[\varepsilon_k]=0,\quad \mathcal{V}[\varepsilon_k]=\sigma_k^2 \quad\text{and}\quad \mathcal{C}[\varepsilon_k,\varepsilon_j]=0,\quad k\neq j,\quad k,j=1,2,\dots,N. \tag{2.4}$$

2.1.3 Total survey randomness

Sample selection randomness and superpopulation randomness are assumed to be independent of each other, i. e. the distribution of $(\mathbf{y},\mathbf{I})=(y_1,y_2,\dots,y_N,I_1,I_2,\dots,I_N)$ is the product measure $\mathbf{P}=\mathcal{P}\times P$. Expectation and variance with respect to **P** are denoted by **E** and **V**.

As a special case of the well-known formula $\mathbf{V}(Z)=\mathbf{E}[\mathbf{V}^{\mathcal{B}}(Z))]+\mathbf{V}[\mathbf{E}^{\mathcal{B}}(Z)]$, where $\mathcal{B}$ stands for some information σ-algebra, we have the following relation which will be useful later on;

$$\mathbf{V}(Z)=\mathcal{E}[V(Z\mid\mathbf{y})]+\mathcal{V}[E(Z\mid\mathbf{y})]=E[\mathcal{V}(Z\mid\mathbf{I})]+V[\mathcal{E}(Z\mid\mathbf{I})]. \tag{2.5}$$

2.2 Estimation

We confine to the "basic" estimation problem, estimation of a ***population total*** $\tau(\mathbf{y})=y_1+y_2+\dots+y_N$. An estimator $\hat{\tau}(\mathbf{y})$ is a function of **I** and the **y**-values for sampled units. It is P-***unbiased*** if $E[\hat{\tau}(\mathbf{y})]=\tau(\mathbf{y})$ holds for all conceivable **y**. It is ***linear*** if it is of the form;

$$\hat{\tau}(\mathbf{y})=W_0+\sum_{k\in\text{Sample}}y_k\cdot W_k=W_0+\sum_{k\in U}y_k\cdot W_k\cdot I_k, \tag{2.6}$$

where W_k may depend on the sample outcome, i.e. $W_k=w_k(\mathbf{I})$, $k=0,1,2,\dots,N$. The ***class of linear P-unbiased estimators*** is denoted by $\mathcal{L}_u(P)$. It is readily checked that $\hat{\tau}(\mathbf{y})$ in (2.6) belongs to $\mathcal{L}_u(P)$ if, and only if;

$$E(W_0)=0,\quad\text{and}\quad E(W_k\cdot I_k)=1,\quad k=1,2,\dots,N. \tag{2.7}$$

Generalized ***difference estimators***, denoted $\hat{\tau}(\mathbf{y};\mathbf{e})_D$, constitute a subclass of $\mathcal{L}_u(P)$. They are defined in (2.8) below, where $\mathbf{e}=(e_1,e_2,\dots,e_N)$ stand for arbitrary, known constants.

$$\hat{\tau}(\mathbf{y};\mathbf{e})_D=\tau(\mathbf{e})+\sum_{k\in\text{Sample}}(y_k-e_k)/\pi_k. \tag{2.8}$$

For $\mathbf{e} = \mathbf{0}$, $\hat{\tau}(y;\mathbf{e})_D$ is the ***Horvitz-Thompson (HT) estimator***;

$$\hat{\tau}(y)_{HT} = \sum_{k \in \text{Sample}} y_k/\pi_k . \tag{2.9}$$

From (2.8) and (2.9) is seen that $\hat{\tau}(y;\mathbf{e})_D$ also can be written;

$$\hat{\tau}(y;\mathbf{e})_D = \hat{\tau}(y)_{HT} + [\tau(\mathbf{e}) - \hat{\tau}(\mathbf{e})_{HT}]. \tag{2.10}$$

2.3 Sampling-estimation strategies

2.3.1 Some generalities

A ***sampling-estimation strategy*** is a pair $[P, \hat{\tau}(y)]$ of a sample design and an estimator.

A strategy $[P, \hat{\tau}(y)]$ is ***admissible*** if $\hat{\tau}(y)$ belongs to $\mathcal{L}_u(P)$. (2.11)

Search for good strategies is confined to admissible ones. Following Cassel et al. (1976) we consider the following performance criterion.

The smaller $\mathcal{E}(V[\hat{\tau}(y)])$ is, the ***better*** the strategy is. (2.12)

2.3.2 Optimal strategies

Theorem 2.1 below is a slight modification of a result in Cassel et al. (1976), see also Th. 4.1 in Cassel et al. (1977). For completeness we give a proof (due to Cassel et al.) afterwards.

THEOREM 2.1: Assumptions and notation are as hitherto in Section 2. In particular, $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ are as in Section 2.1.2. The minimal possible value of $\mathcal{E}(V[\hat{\tau}(y)])$ over the class of admissible strategies $[P, \hat{\tau}(y)]$ with fixed sample size is attained if, and only if, the strategy which satisfies conditions (i) and (ii) below;

(i) P is a $\pi ps(\sigma)$ scheme, i.e. π_k is proportional to σ_k, $k = 1,2,\dots,N$. (2.13)

(ii) $\hat{\tau}(y) = \hat{\tau}(y;\boldsymbol{\mu},\boldsymbol{\sigma})_D := \tau(\boldsymbol{\mu}) + \left(\sum_{k \in U} \sigma_k\right) \cdot \sum_{k \in \text{Sample}} (y_k - \mu_k)/(n \cdot \sigma_k)$. (2.14)

A strategy which satisfies (i) and (ii) is called an ***optimal strategy***.

By (2.10), the estimator in (2.14) can, under (2.13), be written;

$$\hat{\tau}(y;\boldsymbol{\mu},\boldsymbol{\sigma})_D = \hat{\tau}(y)_{HT} + [\tau(\boldsymbol{\mu}) - \hat{\tau}(\boldsymbol{\mu})_{HT}]. \tag{2.15}$$

Proof of Theorem 2.1 : In the following is presumed that the strategy is admissible, i.e. $\hat{\tau}(y)$ has form (2.6) with (2.7) in force. Application of the identity (2.5) with $Z = \hat{\tau}(y)$ yields;

$$\mathcal{E}[V(\hat{\tau}(y) \mid y)] = E[\mathcal{V}(\hat{\tau}(y) \mid \mathbf{I})] + V[\mathcal{E}(\hat{\tau}(y) \mid \mathbf{I})] - \mathcal{V}[E(\hat{\tau}(y) \mid y)]. \tag{2.16}$$

To exploit (2.16) we start with the following observations.

When $\hat{\tau}(y)$ is P-unbiased, $E(\hat{\tau}(y) \mid y) = \tau(y)$ which yields $\mathcal{V}[E(\hat{\tau}(y) \mid y)] = \Sigma_U \sigma_k^2$. (2.17)

$$E[\mathcal{V}(\hat{\tau}(y) \mid \mathbf{I})] = E[\mathcal{V}(W_0 + \textstyle\sum_1^N y_k \cdot W_k \cdot I_k \mid \mathbf{I})] = E\left(\textstyle\sum_1^N \sigma_k^2 \cdot W_k^2 \cdot I_k\right). \tag{2.18}$$

In the right hand sum in (2.18) we first apply Schwarz' inequality $(\Sigma a_k^2) \geq (\Sigma a_k \cdot b_k)^2/(\Sigma b_k^2)$ with $a_k = \sigma_k \cdot W_k \cdot \sqrt{I_k}$ and $b_k = \sqrt{I_k}$, then the fixed sample size assumption $\Sigma I_k = n$, next Jensen's inequality $E(Z^2) \geq [E(Z)]^2$, and finally (2.7), which yields;

$$E[\mathcal{V}(\hat{\tau}(y) \mid \mathbf{I})] \geq E\left[\left(\textstyle\sum_1^N \sigma_k \cdot W_k \cdot I_k\right)^2 / \textstyle\sum_1^N I_k\right] \geq \left(\textstyle\sum_1^N \sigma_k \cdot E[W_k \cdot I_k]\right)^2/n = \left(\textstyle\sum_1^N \sigma_k\right)^2/n . \tag{2.19}$$

(2.16), (2.17), (2.19) and non-negativity of the variance $V[\mathcal{E}(\hat{\tau}(y) \mid \mathbf{I})]$ is yield the following

For any admissible strategy with fixed sample size holds;

$$\mathcal{E}[V(\hat{\tau}(y) \mid y)] \geq \left(\textstyle\sum_1^N \sigma_k\right)^2/n - \textstyle\sum_1^N \sigma_k^2 . \tag{2.20}$$

The lower bound in (2.20) is attained for a strategy that satisfies (2.13) and (2.14). To realize this, use (2.16) in combination with (2.17) and the following relations, which are readily checked, $E[\mathcal{V}(\hat{t}(y;\mu,\sigma)_D|I)] = (\sum_1^N \sigma_k)^2/n$ and $V[\mathcal{E}(\hat{t}(y;\mu,\sigma)_D|I)] = 0$. Hence;

$$\mathcal{E}[V(\hat{t}(y;\mu,\sigma)_D \mid y)] = \text{right hand side in (2.20).} \tag{2.21}$$

Thereby the if-part of the theorem is shown. The "only if" part is realized as follows. For a strategy with fixed sample size to attain the minimal value in (2.20) at least the following conditions must be satisfied. (i) There is equality in Schwartz' inequality with P-probability 1. This occurs if and only if $\sigma_k \cdot W_k \cdot \sqrt{I_k}$ is proportional to $\sqrt{I_k}$ with P-probability 1, which occurs if and only if W_k is proportional to $1/\sigma_k$. (ii) $V[\mathcal{E}(\hat{t}(y) \mid I)] = 0$, which requires $\mathcal{E}(\hat{t}(y) \mid I)$ to be non-random. Upon some thought is realized that this can only occur under (2.13) and (2.14). This concludes the proof of the theorem.

2.3.3 Situations with auxiliary information

Here is presumed that values of R auxiliary variables are available for each unit in the frame, denoted by $\mathbf{x}_r = (x_{r1}, x_{r2}, \ldots, x_{rN})$, $r = 1, 2, \ldots, R$,. The study $\mathbf{y}$ variable and the auxiliary variables are assumed to be related according to the linear model (2.22) below, with (2.4) in force;

$$y_k = \sum_{r=1}^{R} \beta_r \cdot x_{rk} + \varepsilon_k, \quad k = 1, 2, \ldots, N. \tag{2.22}$$

In matrix/vector notation (2.20) can be written $\mathbf{y} = \sum_{r=1}^{R} \beta_r \cdot \mathbf{x}_r$, and even more compactly with $\times$ for matrix multiplication and ' for matrix transposition;

$$\mathbf{y} = \boldsymbol{\beta} \times \mathbf{X} + \boldsymbol{\varepsilon}, \qquad \text{with } \mathbf{X} = (\mathbf{x}_1', \mathbf{x}_2', \ldots, \mathbf{x}_R') \text{ and } \boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_R). \tag{2.23}$$

Model (2.20) is a special case of (2.3), with $\mu_k = \sum_r \beta_r \cdot x_{rk}$. Hence, Theorem 2.1 and (2.15) yield that conditions (2.22) and (2.23) below are necessary and sufficient ***for a strategy to be optimal under (2.20) + (2.4).***

$$\text{P is a } \pi ps(\sigma) \text{ scheme,} \tag{2.24}$$

$$\hat{t}(y) = \hat{t}(y)_{GREG1} := \hat{t}(y)_{HT} + \sum_{r=1}^{R} \beta_r \cdot [t(\mathbf{x}_r) - \hat{t}(\mathbf{x}_r)_{HT}]. \tag{2.25}$$

In (2.23) subscript GREG1 is used to indicate that the estimator is a first step towards the generalized regression estimator, which will be denoted $\hat{t}(\mathbf{y})_{GREG}$.

3 Generalized regression estimation

Here a brief review of generalized regression estimation (GREG) is given, based on Särndal et al. (1992). First a notation convention which is used throughout the paper. Algebra operations, $\otimes$ (= +, -, $\cdot$, /, etc.), on variables (scalar, vector or matrix valued) stand for ***component-wise operations***. For $\mathbf{y} = (y_1, y_2, \ldots, y_N)$ and $\mathbf{z} = (z_1, z_2, \ldots, z_N)$, $\mathbf{y} \otimes \mathbf{z} = (y_1 \otimes z_1, y_2 \otimes z_2, \ldots, y_N \otimes z_N)$.

3.1 Some basics

Even if Cassel et al. (1976) derive the estimator $\hat{t}(\mathbf{y})_{GREG1}$ in (2.25) in tandem with a πps design, this estimator is well-defined for any design. In fact it is P-unbiased for any design P and any values $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_R)$, be they the values in the model (2.22) or some other. The precision of the estimator $\hat{t}(\mathbf{y})_{GREG1}$, though, depends on the employed $\boldsymbol{\beta}$, the better it complies with the model $\boldsymbol{\beta}$, the better estimation precision.

However, in practice $\boldsymbol{\beta}$ is normally unknown. The GREG approach to circumvent this is to estimate $\boldsymbol{\beta}$ from the sample data. Then, in the first round one presumes that for *all* population units (**y**, **X**) as well as the spread $\boldsymbol{\sigma}$ are known. General regression analysis results then tell the following. (Recall that τ stands for population total).

The BLUE estimate of $\boldsymbol{\beta}$ under (2.21) + (2.4) is $\mathbf{B} = (B_1, B_2, ..., B_R)' = \mathbf{T}^{-1} \times \mathbf{t}$, (3.1)

where

$$\mathbf{T} = \tau(\mathbf{X}' \times \mathbf{X})/\boldsymbol{\sigma}^2) = \sum_{k \in U} (\mathbf{X}_k' \times \mathbf{X}_k)/\sigma_k^2, \quad \mathbf{t} = \tau(\mathbf{y} \cdot \mathbf{X}/\boldsymbol{\sigma}^2) = \sum_{k \in U} y_k \cdot \mathbf{X}_k / \sigma_k^2. \tag{3.2}$$

To take a second step towards $\hat{\tau}(\mathbf{y})_{GREG}$, exchange β_r in (2.25) for B_r in (3.1) + (3.2);

$$\hat{\tau}(\mathbf{y})_{GREG2} = \hat{\tau}(\mathbf{y})_{HT} + \sum_{r=1}^{R} B_r \cdot [\tau(\mathbf{x}_r) - \hat{\tau}(\mathbf{x}_r)_{HT}]. \tag{3.3}$$

The estimator (3.3) is not practicable, though, since the B_r are unknown population quantities. This dilemma is circumvented in the "usual" way, by exchanging **B** for a sample estimate $\hat{\mathbf{B}}$ of it. Hereby HT - estimators are used to estimate population totals. With (3.1), (3.2) and (3.3) as background this leads to;

$$\hat{\mathbf{B}} = (\hat{B}_1, \hat{B}_2, ..., \hat{B}_R)' = \hat{\mathbf{T}}^{-1} \times \hat{\mathbf{t}}, \tag{3.4}$$

where

$$\hat{\mathbf{T}} = \hat{\tau}(\mathbf{X}_k \times \mathbf{X}_k / \boldsymbol{\sigma}^2)_{HT} = \sum_{k \in \text{Sample}} \mathbf{X}_k \times \mathbf{X}_k / (\sigma_k^2 \cdot \pi_k), \tag{3.5}$$

$$\hat{\mathbf{t}} = \hat{\tau}(\mathbf{y} \cdot \mathbf{x}/\boldsymbol{\sigma}^2)_{HT} = \sum_{k \in \text{Sample}} y_k \cdot \mathbf{X}_k / (\sigma_k^2 \cdot \pi_k). \tag{3.6}$$

Insertion into (3.3) now yields the ***generalized regression estimator***, the ***GREG-estimator***;

$$\hat{\tau}(\mathbf{y})_{GREG} = \hat{\tau}(\mathbf{y})_{HT} + \sum_{r=1}^{R} \hat{B}_r \cdot [\tau(\mathbf{x}_r) - \hat{\tau}(\mathbf{x}_r)_{HT}]. \tag{3.7}$$

For the sake of simplicity, in the sequel we confine to the case with ***one-dimensional auxiliary data*** **x**. Then (2.21), and (3.4) - (3.7) takes the following forms;

$$y_k = \beta \cdot x_k + \varepsilon_k, \quad k = 1, 2, ..., N, \tag{3.8}$$

$$B = \tau(\mathbf{y} \times \mathbf{x}/\boldsymbol{\sigma}^2)/\tau(\mathbf{x}^2/\boldsymbol{\sigma}^2) = \sum_{k \in U} (y_k \cdot x_k)/\sigma_k^2 \Big/ \sum_{k \in U} x_k^2 / \sigma_k^2, \tag{3.9}$$

$$\hat{B}(\boldsymbol{\sigma}) = \hat{\tau}(\mathbf{y} \times \mathbf{x}/\boldsymbol{\sigma}^2)_{HT} / \hat{\tau}(\mathbf{x}^2/\boldsymbol{\sigma}^2)_{HT} = \sum_{k \in \text{Sample}} (y_k \cdot x_k)/(\sigma_k^2 \cdot \pi_k) / \sum_{k \in \text{Sample}} x_k^2 / (\sigma_k^2 \cdot x_k). \tag{3.10}$$

$$\hat{\tau}(\mathbf{y})_{GREG} = \hat{\tau}(\mathbf{y})_{HT} + \hat{B}(\boldsymbol{\sigma}) \cdot [\tau(\mathbf{x}) - \hat{\tau}(\mathbf{x})_{HT}]. \tag{3.11}$$

At this junction we point at the fact that a superpopulation model plays two roles, firstly it ***describes the study variable variation*** over the population, and secondly it is an instrument for choosing the ***sampling-estimation*** strategy. As always in model contexts there is a dichotomy between ***true*** and ***believed*** model. The former describes how "nature" generated the study variable values, while the latter specifies how the statistician believes they were generated. For true as well as believed model, the *linear* model (2.21) is a possible option, but any type of model can in principle be used.

The distinction between true and believed model concerns in particular the spread parameter $\boldsymbol{\sigma}$. So far $\boldsymbol{\sigma}$ has been viewed as known, at least up to a proportionality factor. Via (3.10) $\boldsymbol{\sigma}$ also plays a role in the estimation process. It may also affect the choice of sample design.The statistician must use some value for $\boldsymbol{\sigma}$, whether he/she knows it or not, a believed ("guestimated" is another possible term) value. The true $\boldsymbol{\sigma}$ is of course preferred, but "truth" and "belief" may deviate. To cope with this possibility, we introduce the following terminology and assumption.

For ***believed models***, also called ***sampling-estimation models***, we confine to (3.8) + (2.4) with the spread parameter changed to $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_N)$. (3.12)

Formula (3.13) states how (3.10) is modified to comply with (3.12).

$$\hat{B}(\boldsymbol{\delta}) = \hat{t}(\mathbf{y} \times \mathbf{x}/\boldsymbol{\delta}^2)_{HT} / \hat{t}(\mathbf{x}^2/\boldsymbol{\delta}^2)_{HT} = \sum_{k \in \text{Sample}} (y_k \cdot x_k)/(\delta_k^2 \cdot \pi_k) \Big/ \sum_{k \in \text{Sample}} x_k^2/(\delta_k^2 \cdot \pi_k). \tag{3.13}$$

In this context we do not bother about the other model parameter, $\boldsymbol{\beta}$, though, since it enters neither in the estimation process nor in the sample design.

3.2 Estimator variance and variance estimation

The GREG estimator is a non-linear function of sampled $\mathbf{y}$-values, which makes the issues "estimator variance" and "variance estimation" a bit complex. One has to rely on approximations. Section 6.6 in Särndal et al. (1992) provides the full story, from which some excerpts are given below. As stated, we confine to the case with one-dimensional auxiliary data.

Regarding $\hat{B}$ as an error free estimate of B leads to the approximation;

$$\hat{t}(\mathbf{y})_{GREG} \approx B \cdot \tau(\mathbf{x}) + \hat{t}(\mathbf{y})_{HT} - B \cdot \hat{t}(\mathbf{x})_{HT} = B \cdot \tau(\mathbf{x}) + \hat{t}(\mathbf{E})_{HT}, \tag{3.14}$$

with

$$\mathbf{E} = (E_1, E_2, \dots, E_N), \quad E_k = y_k - B \cdot x_k, \quad k = 1, 2, \dots, N, \tag{3.15}$$

Relation (3.14) leads to the following approximate variance formula;

$$V[\hat{t}(\mathbf{y})_{GREG}] \approx V[\sum_{k \in \text{Sample}} E_k/\pi_k]. \tag{3.16}$$

We presume that a procedure $\hat{V}$ for estimation of the variance of a sample sum is available for the used sample design. Application of $\hat{V}$ to the sum to the right in (3.16) would yield a variance estimator $\hat{V}[\hat{t}(\mathbf{y})_{GREG}]$, but there is an obstacle. The E_k are not known even for sampled units, since they depend on B, which in turn depends on all $\mathbf{y}$-values in the population. To circumvent this obstacle, the following proxies e_k for the E_k are introduced;

$$e_k = y_k - \hat{B} \cdot x_k, \quad k \in \text{Sample}. \tag{3.17}$$

Next, by exchanging E_k in (3.16) for e_k and treating the e_k as constants (although they depend on the sample), the following variance estimator is obtained;

$$\hat{V}[\hat{t}(\mathbf{y})_{GREG}]_1 = \hat{V}[\sum_{k \in \text{Sample}} e_k/\pi_k], \quad \text{with } e_k \text{ treated as being non-random}. \tag{3.18}$$

A refinement of (3.18) is achieved by the so called ***g-method***. The point of departure is then the following version of $\hat{t}(\mathbf{y})_{GREG}$, see e.g. (6.5.18) in Särndal et al. (1992);

$$\hat{t}(\mathbf{y})_{GREG} = B \cdot \tau(\mathbf{x}) + \sum_{k \in \text{Sample}} \frac{E_k \cdot g_k}{\pi_k}, \quad g_k = 1 + \frac{x_k \cdot [\tau(\mathbf{x}) - \hat{t}(\mathbf{x})_{HT}]}{\delta_k^2 \cdot \hat{t}(\mathbf{x}^2/\boldsymbol{\delta}^2)_{HT}}. \tag{3.19}$$

By (3.19) the following holds: $V[\hat{t}(\mathbf{y})_{GREG}] = V[\sum_{k \in \text{Sample}} E_k \cdot g_k/\pi_k]$. The last term is then estimated as follows. Use e_k as a proxy for E_k, and treat e_k as well as g_k as non-random (although they depend on the sample), leading to the following alternative variance estimator;

$$\hat{V}[\hat{t}(\mathbf{y})_{REG}]_2 = \hat{V}[\sum_{k \in \text{Sample}} e_k \cdot g_k/\pi_k], \quad \text{with } e_k \text{ and } g_k \text{ treated as non-random}. \tag{3.20}$$

4 On πps sampling

4.1 Extension of the πps notion

We start by slightly modifying the notion of πps as formulated in Section 2.1.1. Firstly, the quantities to the right in (2.2) are re-named. For a size measure $\mathbf{s} = (s_1, s_2, \ldots, s_N)$ and a sample size n, the $\boldsymbol{\lambda} = (\lambda_1, \lambda_2.., \lambda_N)$ in (4.1) below are called the ***desired inclusion probabilities***;

$$\lambda_k = n \cdot s_k / \sum_{j=1}^{N} s_j, \quad k = 1, 2, \ldots, N. \tag{4.1}$$

As before is presumed that $\lambda_k < 1$, $k = 1, 2, \ldots, N$. (If not, some appropriate adjustment should be made.) Secondly, the notion of πps is made a bit wider than before. Definition (2.2) requires that a πps design satisfies $\pi_k = \lambda_k$. From now on we accept a sampling scheme as a ***πps design*** if, with $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_N)$ for the factual inclusion probabilities;

$$\pi_k \approx \lambda_k \text{ holds } \textit{with good approximation} \text{ for } k = 1, 2, \ldots, N. \tag{4.2}$$

For a "perfect" πps scheme (i.e. a scheme with $\pi_k = \lambda_k$) the HT-estimator is;

$$\hat{t}(y)_{\pi ps} = \sum_{k \in \text{Sample}} y_k / \lambda_k \,. \tag{4.3}$$

In particular, for a perfect πps scheme the estimator in (4.3) unbiased. Under (4.2) it is only an "approximate HT - estimator", which may have some bias. However, with no further auxiliary information available (4.3) is the "natural" estimator under a πps design (in (4.2) sense).

4.2 Pareto πps

Pareto πps as defined below was introduced in Rosén (1997).

DEFINITION 4.1: A ***Pareto πps*** sample with ***size measure*** s and ***sample size*** n is selected as follows. First compute desired inclusion probabilities $\lambda_1, \lambda_2, \ldots, \lambda_N$ by (4.1), as usual presuming that $\lambda_k < 1$, $k = 1, 2, \ldots, N$. Then realize independent random variables $U_1, U_2, \ldots, U_N$ with uniform distribution on $[0, 1]$, and compute;

$$Q_k = \frac{U_k \cdot (1 - \lambda_k)}{\lambda_k \cdot (1 - U_k)}, \quad k = 1, 2, \ldots, N. \tag{4.4}$$

The ***sample*** consists of the units with the ***n smallest Q-values.***

In spite of its name, it is not obvious that Pareto πps is a πps design in the (4.2) sense. However, this is shown in Rosén (2000) and Aires & Rosén (2000), where also is shown that in almost all practical situations $\hat{t}(y)_{\pi ps}$ has negligible bias under Pareto πps.

The results in (4.5) and (4.6) below are justified in Rosén (1997).

Asymptotically correct approximation of the estimator variance is given by;.

$$V[\hat{t}(y)_{\pi ps}] \approx \frac{N}{N-1} \cdot \sum_{k=1}^{N} \left(\frac{y_k}{\lambda_k} - \sum_{j=1}^{N} y_j \cdot (1 - \lambda_j) / \sum_{j=1}^{N} \lambda_j \cdot (1 - \lambda_j) \right)^2 \cdot \lambda_k \cdot (1 - \lambda_k). \tag{4.5}$$

Consistent estimation of $V[\hat{t}(y)_{\pi ps}]$ is given by;

$$\hat{V}[\hat{t}(y)_{\pi ps}] = \frac{n}{n-1} \cdot \sum_{k \in \text{Sample}} \left(\frac{y_k}{\lambda_k} - \sum_{j \in \text{Sample}} \frac{y_j \cdot (1 - \lambda_j)}{\lambda_j} \Big/ \sum_{j \in \text{Sample}} (1 - \lambda_j) \right)^2 \cdot (1 - \lambda_k). \tag{4.6}$$

Remark 4.1: Computation of (4.5) and (4.6) is facilitated by (4.7) and (4.8) below, where W, R and S are as stated in (4.9) and (4.10);

$$\text{Right hand side in (4.5)} = (W - R^2/S) \cdot N/(N-1), \tag{4.7}$$

$$\text{Right hand side in (4.6)} = (\hat{W} - \hat{R}^2/\hat{S}) \cdot n/(n-1), \tag{4.8}$$

$$W = \sum_{k=1}^{N} \frac{y_k^2 \cdot (1-\lambda_k)}{\lambda_k}, \quad R = \sum_{k=1}^{N} y_k \cdot (1-\lambda_k), \quad S = \sum_{k=1}^{N} \lambda_k \cdot (1-\lambda_k)\,, \tag{4.9}$$

$$\hat{W} = \sum_{k \in \text{Sample}} \left(\frac{y_k}{\lambda_k}\right)^2 \cdot (1-\lambda_k), \; \hat{R} = \sum_{k \in \text{Sample}} \frac{y_k \cdot (1-\lambda_k)}{\lambda_k}, \quad \hat{S} = \sum_{k \in \text{Sample}} (1-\lambda_k). \tag{4.10}$$

5 Regression estimation under πps

5.1 A slightly modified GREG estimator

Here we consider situations where the study variable $\mathbf{y}$ is observed for a size n πps($\mathbf{s}$) sample (in (4.2) sense) and an auxiliary variable $\mathbf{x}$ is available. For simplicity $\mathbf{x}$ is presumed one - dimensional. The task is to estimate the population total $\tau(\mathbf{y})$.

For a (wide sense) πps design the GREG estimator can often not be derived precisely as stated in Section 3, since one does not know the exact inclusion probabilities π_k, which are needed for HT - estimation. However, the "quasi HT - estimator" (4.3) is available. The definition of GREG - estimator is therefore modified by letting $\hat{\tau}(\cdot)_{\pi ps}$ play the role of $\hat{\tau}(\cdot)_{HT}$. The ***modified GREG estimator*** is as follows;

$$\hat{\tau}(\mathbf{y};\boldsymbol{\delta})_{\text{GREG}}^{\pi ps} = \hat{\tau}(\mathbf{y})_{\pi ps} + \hat{B}(\boldsymbol{\delta})_{\pi ps} \cdot [\tau(\mathbf{x}) - \hat{\tau}(\mathbf{x})_{\pi ps}], \tag{5.1}$$

where

$$\hat{B}(\boldsymbol{\delta})_{\pi ps} = \hat{\tau}(\mathbf{y} \cdot \mathbf{x}/\boldsymbol{\delta}^2)_{\pi ps} / \hat{\tau}(\mathbf{x}^2/\boldsymbol{\delta}^2)_{\pi ps} = \sum_{k \in \text{Sample}} (y_k \cdot x_k)/(\delta_k^2 \cdot \lambda_k) \Big/ \sum_{k \in \text{Sample}} x_k^2/(\delta_k^2 \cdot \lambda_k). \tag{5.2}$$

5.2 GREG estimation under Pareto πps design

5.2.1 General results

The estimator given by (5.1) + (5.2) works for any πps($\mathbf{s}$) design, in particular for Pareto πps. Formulas for estimator variance and variance estimation, though, differ for πps schemes. For Pareto πps, combination of (3.16) and (4.5) yields the following approximate variance formula for the GREG estimator, with E_k according to (3.15);

$$V[\hat{\tau}(\mathbf{y})_{\text{GREG}}^{\pi ps}] \approx \frac{N}{N-1} \cdot \sum_{k=1}^{N} \left(\frac{E_k}{\lambda_k} - \sum_{j=1}^{N} E_j \cdot (1-\lambda_j) \Big/ \sum_{j=1}^{N} \lambda_j \cdot (1-\lambda_j) \right)^2 \cdot \lambda_k \cdot (1-\lambda_k). \tag{5.3}$$

Combination of (3.18) and (4.6) yields the following variance estimator, with e_k as in (3.17);

$$\hat{V}[\hat{\tau}(\mathbf{y})_{\text{GREG}}^{\pi ps}]_1 = \frac{n}{n-1} \cdot \sum_{k \in \text{Sample}} \left(\frac{e_k}{\lambda_k} - \sum_{j \in \text{Sample}} \frac{e_j \cdot (1-\lambda_j)}{\lambda_j} \Big/ \sum_{j \in \text{Sample}} (1-\lambda_j) \right)^2 \cdot (1-\lambda_j). \tag{5.4}$$

For the g - method the definition of the ***g - coefficients*** are modified as follows;

$$g_k = 1 + \frac{x_k \cdot [\tau(\mathbf{x}) - \hat{\tau}(\mathbf{x})_{\pi ps}]}{\delta_k^2 \cdot \hat{\tau}(\mathbf{x}^2/\boldsymbol{\delta}^2)_{\pi ps}}, \quad k = 1,2,\ldots,N. \tag{5.5}$$

Combination of (3.20) and (4.6) yields the alternative variance estimator;

$$\hat{V}[\hat{\tau}(\mathbf{y})_{\text{GREG}}^{\pi ps}]_2 = \frac{n}{n-1} \cdot \sum_{k \in \text{Sample}} \left(\frac{e_k \cdot g_k}{\lambda_k} - \sum_{j \in \text{Sample}} \frac{e_j \cdot g_j \cdot (1-\lambda_j)}{\lambda_j} \Big/ \sum_{j \in \text{Sample}} (1-\lambda_j) \right)^2 \cdot (1-\lambda_k). \tag{5.6}$$

Note that (5.4) and (5.6) can be computed by employing Remark 4.1, after using the transformations $y_k \rightarrow e_k$ respectively $y_k \rightarrow e_k \cdot g_k$.

5.2.2 GREG when the auxiliary variable is used as size measure

Assume that model (3.8) is in force and that β is positive, which is the typical case in practice. Then $\mathbf{y}$ and $\mathbf{x}$ are positively correlated, often even fairly proportional to each other. The "traditional" πps approach in this type of situation is to use $\mathbf{x}$ as size measure, i.e. to use a $\pi ps(\mathbf{x})$ design, accompanied by $\hat{\tau}(\mathbf{x})_{\pi ps}$ in (4.3). A natural question is therefore: Given that the sample is selected with $\pi ps(\mathbf{x})$, can GREG estimation lead to improvements over $\hat{\tau}(\mathbf{x})_{\pi ps}$? The answer is *no*, as is seen from the following result.

Under a $\pi ps(\mathbf{x})$ design the following holds for any $\boldsymbol{\delta} \geq 0$: $\hat{\tau}(\mathbf{y};\boldsymbol{\delta})^{\pi ps}_{\text{GREG}} = \hat{\tau}(\mathbf{y})_{\pi ps}$. (5.7)

To realize (5.7), note that for any $\pi ps(\mathbf{x})$ design holds $\hat{\tau}(\mathbf{x})_{\pi ps} = \tau(\mathbf{x})$. Having this, the claim in (5.7) follows readily from (5.1). With (5.7) as background, the following results about variances should "reasonably" hold for Pareto $\pi ps(\mathbf{x})$, and they do hold.

Under Pareto $\pi ps(\mathbf{x})$: $V[\hat{\tau}(\mathbf{y})^{\pi ps}_{\text{GREG}}]$ in (5.3) = $V[\hat{\tau}(\mathbf{y})_{\pi ps}]$ in (4.5), (5.8)

Under Pareto $\pi ps(\mathbf{x})$: The two versions of $\hat{V}[\hat{\tau}(\mathbf{y})^{\pi ps}_{\text{GREG}}]$ in (5.4) and (5.6) agree, and they are both equal to $\hat{V}[\hat{\tau}(\mathbf{y})_{\pi ps}]$ in (4.7). (5.9)

It is quite straightforward to checks that (5.8) and (5.9) hold not only "reasonably" but also algebraically, and the details are left to the reader. When checking (5.9) note that for $\mathbf{s} = \mathbf{x}$ the g-coefficients are $g_k = 1$, $k=1,2,\dots,N$, which implies that (5.4) and (5.6) coincide.

6 The optimal strategy problem revisited

6.1 Introduction

In this section we pursue the optimal strategy issue. As in Section 2, the framework is confined to estimation of a population total $\tau(\mathbf{y})$ from observations of $\mathbf{y}$ on a wor probability sample with fixed sample size, when auxiliary information $\mathbf{x}$, for simplicity one-dimensional, is available for each population unit.

Theorem 2.1 provides background for belief that strategies of type $[\pi ps(\boldsymbol{\sigma}), \hat{\tau}(\mathbf{y})^{\pi ps}_{\text{GREG}}]$ are close to optimal under the model (3.8) + (2.4). Moreover, since we regard Pareto πps as a particularly attractive πps design, the strategy [Pareto $\pi ps(\boldsymbol{\sigma})$, $\hat{\tau}(\mathbf{y})^{\pi ps}_{\text{GREG}}$] will be of special interest, leading to the conjecture in (6.1) below. There, and throughout the paper, is presumed that compared strategies have the same sample size, and also that estimators have negligible bias. Hence, strategy is measured by estimator variance.

Is the following **conjecture** true? Under (3.8) + (2.4), the performance of the strategy [Pareto $\pi ps(\boldsymbol{\sigma})$, $\hat{\tau}(\mathbf{y};\boldsymbol{\sigma})^{\pi ps}_{\text{GREG}}$] is superior, at least never notably inferior, to that of any other admissible strategy. (6.1)

An aspect of (6.1) concerns quantification of the no-answer for the strategies considered in Section 5.2.2, to use the auxiliary $\mathbf{x}$ as size measure in a πps design accompanied by estimation according to (4.3). This leads to the following question.

How much inferior to [Pareto $\pi ps(\boldsymbol{\sigma})$, $\hat{\tau}(\mathbf{y})^{\pi ps}_{\text{GREG}}$] is [Pareto $\pi ps(\mathbf{x})$, $\hat{\tau}(\mathbf{y})_{\pi ps}$] ? (6.2)

Model based and model assisted procedures of course suffer in some respect if the model is misjudged. In a model based approach this commonly leads to point estimation bias. In a model assisted approach, as GREG, misjudgment does not lead to bias but affects estimation precision adversely. It is of interest to obtain quantitative information about the precision loss by model misspecification. We shall considerer two types of misspecification. In the simplest one is assumed that true and believed superpopulation models both are linear, while the true spread σ is misjudged to be δ. Striving to be close to optimal it is natural to employ the strategy [Pareto $\pi ps(\delta)$, $\hat{t}(y;\delta)^{\pi ps}_{GREG}$] in which δ affects the sample design as well as the estimator. This raises the following question.

If (3.8) + (2.4) is the true model, but (3.12) is used as sampling-estimation model, how inferior to [Pareto $\pi ps(\sigma)$, $\hat{t}(y;\sigma)^{\pi ps}_{GREG}$] is [Pareto $\pi ps(\delta)$, $\hat{t}(y;\delta)^{\pi ps}_{GREG}$]? (6.3)

A more complex misspecification possibility is that the true model is nonlinear, while the linear model (3.12) is used as sampling-estimation model, with correct or misjudged σ.

How does [Pareto $\pi ps(\delta)$, $\hat{t}(y;\delta)^{\pi ps}_{GREG}$] perform when the true model is non-linear? (6.4)

6.2 Performance measures for strategies

To study the questions (6.1) - (6.4), some performance measure for strategies must be used. Although optimality in Theorem 2.1 relates to the criterion $\mathcal{E}(V[\hat{t}(y)])$, its design analogue $V[\hat{t}(y)]$ will be employed, in spite of its drawback to depend on the specific study variable y. It is used for the following main reasons. (i) When true and believed superpopulation models differ, $V[\hat{t}(y)]$ is the more natural measure. (ii) When the models agree and the sample size is not "too small", $\mathcal{E}(V[\hat{t}(y)])$ and $V[\hat{t}(y)]$ lie close to each other. In particular, the notion "close to optimal" is fairly much the same for the two criteria. The measure $V[\hat{t}(y)]$ is not used as it stands, though, but in transformed versions. The quantities RME and RVI specified below are preferred since they, although being essentially equivalent to $V[\hat{t}(y)]$, have more concrete interpretations.

The ***relative margin of error*** (in %), abbreviated RME, for the strategy [P, $\hat{t}(y)$] is;

$$\textbf{RME} = 100 \cdot 1.96 \cdot \sqrt{V[\hat{t}(y)] \text{ under design P}} \,/\, \tau(y) \ \%. \tag{6.5}$$

Conjecture (6.1) provides background for name and choice of denominator in the next notion.

The ***relative variance increase*** (in %), abbreviated RVI, for strategy [P, $\hat{t}(y)$] is;

$$\textbf{RVI} = \left(\frac{V[\hat{t}(y)] \text{ under the design P}}{V[\hat{t}(y)] \text{ for the strategy [Pareto } \pi ps(\sigma),\ \hat{t}(y;\sigma)^{\pi ps}_{GREG}]} - 1 \right) \cdot 100\,\%. \tag{6.6}$$

RME and RVI are not independent measures, though. The following holds. RVI [P, $\hat{t}(y)$] = [(RME[P, $\hat{t}(y)$]/RME[Pareto $\pi ps(\sigma)$, $\hat{t}(y;\delta)^{\pi ps}_{GREG}$])^2 - 1] $\cdot$ 100 %. However, we think it facilitates for the reader if both RME and RVI are presented

One would like to be able to carry out strategy comparisons relating to questions (6.1)-(6.4) by employing nice analytical formulas for RME and RVI. To the best of our understanding, though, it is in vain to hope for such formulas. The feasible approach is to carry out a numerical study for a selection of strategies and test situations, and this approach was used.

The simplest would have been to take for granted that (4.5) and (5.3) work with "good enough" accuracy, and to use them to derive RME and RVI. This would have required fairly modest numerical efforts. However, GREG as well as Pareto πps are large sample procedures, and it is not obvious how accurately their formulas work for finite samples. To gain information also on that question, a more elaborate numerical approach was used. Numerical results were derived in a Monte Carlo study, with repeated independent samples.

6.3 Approximation accuracy

The considered accuracy questions are listed below.

Point estimator bias

Särndal et al (1992) show that $\hat{\tau}(y)_{GREG}$ is consistent under general conditions. Rosén (2000) and Aires & Rosén (2000) show that $\hat{\tau}(y)_{\pi ps}$ under general conditions has negligible bias for Pareto πps. Of special interest was to see if the same holds for the combined estimator $\hat{\tau}(y\,;\delta)^{\pi ps}_{GREG}$. The performance measure is ***relative bias for point estimator*** (RBPE);

Variance estimator bias

The variance estimators (3.18), (3.20) and (4.6) are based on large - sample considerations. Even if not unbiased they are consistent under general conditions. Another task of special interest was to find out to what extent this holds for the combined variance estimators (5.4) and (5.6).The performance measure is ***relative bias for variance estimator*** (RBVE);

$$\mathbf{RBVE} = (\,E(\hat{V}[\hat{\tau}(y)])/V[\hat{\tau}(y)]-1)\cdot 100\,\%. \tag{6.8}$$

Accuracy in approximate formulas for the theoretical estimator variance

The chief interest in this context relates to the large-sample formula (5.3). When evaluating its accuracy we use *standard deviation*, abbreviated D, as basic quantity (instead of variance). The performance measure is ***relative error for theoretical standard deviation*** (RESTD);

$$\mathbf{RETSD} = (D[\hat{\tau}(y)]_{appr}/D[\hat{\tau}(y)]-1)\cdot 100\,\%,\;\; D[\hat{\tau}]_{appr} = \sqrt{V[\hat{\tau}]_{appr}},\; D[\hat{\tau}] = \sqrt{V[\hat{\tau}]}\,. \tag{6.9}$$

The performance measures (6.5) - (6.9) involve the theoretical quantities $\tau(y)$, $E[\hat{\tau}(y)]$, $V[\hat{\tau}(y)]$ and $E(\hat{V}[\hat{\tau}(y)])$. Even though all population values are known, only $\tau(y)$ can be computed exactly, in lack of manageable expressions for the others. Numerical values for them were derived as means based on 3000 independent samples, see (6.10). Since as many as 3 000 runs were made, the means are regarded as true values, even if "empirical" is a more adequate term.

$$E[\hat{\tau}(y)] = \frac{1}{3000}\cdot\sum_{u=1}^{3000}\hat{\tau}(y)_u,\;\; V[\hat{\tau}(y)] = \frac{1}{3000-1}\cdot\sum_{u=1}^{3000}(\hat{\tau}(y)_u-\overline{\hat{\tau}(y)})^2,\;\; E[\hat{V}[\hat{\tau}(y)]] = \frac{1}{3000}\cdot\sum_{u=1}^{3000}\hat{V}[\hat{\tau}(y)_u]\,. \tag{6.10}$$

6.4 The numerical study

6.4.1 Test situations

A ***test situation*** is specified by values for a study variable **y** and an auxiliary variable **x** for each unit in a population. The following general framework was used, judged to embrace a versatile family of test situations with parsimonious parameterization.

The values of the ***auxiliary variable*** **x** were set to;

$$x_k = k,\quad k = 1,2,\ldots,N. \tag{6.11}$$

The ***study variable*** values y were derived by first specifying values for (non - negative) parameters α, β, c and γ, and then generate y - values by relation (6.12) below, where $Z_1, Z_2, \ldots, Z_N$ stand for independent standard normal random variables;

$$y_k = \beta\cdot x_k^{\alpha} + c\cdot x_k^{\gamma}\cdot Z_k,\quad k = 1,2,\ldots,N. \tag{6.12}$$

In the notation used in Section 2.1.2 this means;

$$\mathcal{E}[\varepsilon_k] = 0,\;\; \mathcal{V}[\varepsilon_k] = \sigma_k^2 = c^2\cdot x_k^{2\gamma},\;\; \mathcal{C}[\varepsilon_k, \varepsilon_l] = 0,\;\; k\neq l,\; k,l = 1,2,\ldots,N. \tag{6.13}$$

In conjunction with model (6.12) we use the following terminology from Rosén (1997). The plot $\{(x_k, \beta\cdot x_k^{\alpha}),\, k = 1,2,\ldots,N\}$ is called the **y - x** - ***trend***. Its ***shape*** is determined by the parameter α, it is ***linear*** when α lies close to 1, ***convex*** when $\alpha > 1$ and ***concave*** when $\alpha < 1$. The parameters c and γ determine how much the y-values ***spread/scatter*** around the trend, γ is called ***spread shape*** and c ***spread magnitude***.

There is an abundance of potentially interesting test situations, but many reasons call for temperance, not least the space required for presentation of numerical findings. The study was confined to six types of test situations, labeled A - F, which are specified in Table 6.1. We believe, or at least hope, that these situations allow for fairly general conclusions. Two parameters were held fixed, the population size N and the regression coefficient β;

$$\text{Only population size } N = 200 \text{ was considered, } \beta \text{ was set to } 1. \tag{6.14}$$

The reason for setting $\beta = 1$ is that the performance measures depend on the parameters β and c only through their ratio β/c. Hence one of β and c can be normalized.

Table 6.1 Parameter values in (6.12) and (6.13) for the considered test situations

Label	α	γ	Resulting σ	c-values		
				c_1	c_2	c_3
A	1	0.5	$\sigma_k \propto \sqrt{x_k}$	0.9	1.8	3.5
B	1	1	$\sigma_k \propto x_k$	0.06	0.12	0.25
C	1	0.25	$\sigma_k \propto \sqrt[4]{x_k}$	3	7	13
D	1.2	0.5	$\sigma_k \propto \sqrt{x_k}$	2.5	5	10
E	1.5	0.5	$\sigma_k \propto \sqrt{x_k}$	12	25	50
F	0.7	0.5	$\sigma_k \propto \sqrt{x_k}$	0.2	0.4	0.7

Below we give some comments on the parameter choices.

(i) The superpopulation model is linear in situations A, B and C, while non-linear in D, E and F, "mildly" convex in D ($\alpha = 1.2$), "strongly" convex in E ($\alpha = 1.5$), concave in F ($\alpha = 0.7$).

(ii) The y-scatter was (hopefully) held at practically realistic levels by the following considerations. For the unit with the largest **x**-value, i.e. unit N = 200, y_N should not deviate from its trend value ($= x_N^\alpha$, when $\beta = 1$) by more than (roughly) half the trend value. With 2 as a practical upper bound for $|Z|$, the following restriction was laid on c;

$$c \cdot x_k^\gamma \cdot 2 \le x_k^\alpha / 2, \quad \text{which implies } c \le x_N^{\alpha-\gamma}/4. \tag{6.15}$$

For each combination of α and γ three spread magnitudes c were used. The largest, denoted c_3, was given by the right hand side in (6.15) for $x_{200} = 200$. The other two, c_2 and c_1, were set to (roughly) $c_2 = c_3/2$ and $c_1 = c_2/2$.

(iii) The normal variates $Z_1, Z_2, \ldots, Z_N$ were generated by the SAS - function NORMAL with seed = 555. This seed value was used to achieve comparability with findings in Rosén (1997).

6.4.2 Sampling-estimation model

Sampling-estimation models (or "believed" model) were chosen in agreement with (3.12), i.e. (3.8)+(2.4) with spread parameter $\boldsymbol{\delta}$. ***Correct spread guestimate*** is said to be at hand if $\boldsymbol{\delta} = \boldsymbol{\sigma}$, ***overgestimated spread*** if $\boldsymbol{\delta} > \boldsymbol{\sigma}$ and ***undergestimated spread*** if $\boldsymbol{\delta} < \boldsymbol{\sigma}$.

6.4.3 Sample designs and estimation procedures

Sample designs were confined to Pareto πps. Note that for $\mathbf{s} = \mathbf{1}$ Pareto πps is simple random sampling. Two estimation modes were considered, "straight" πps estimation by (4.3) (which is HT - estimation for simple random sampling), and GREG estimation by (5.1) and (5.2). The following shorthand for designs and estimation procedures is used.

$$\textbf{PAR}(\rho) \text{ stands for the Pareto } \pi\text{ps}(\mathbf{x}^{\rho}) \text{ scheme,} \tag{6.16}$$

$$\textbf{SRS} \text{ stands for simple random sampling } [= \text{PAR}(0)], \tag{6.17}$$

$$\boldsymbol{\pi}\textbf{ps} \text{ stands for the estimator } \hat{\tau}(\mathbf{y})_{\pi ps} \text{ in (4.3), } \textbf{HT} \text{ for Horvitz-Thompson estimation,} \tag{6.18}$$

$$\textbf{GREG}(\rho) \text{ stands for the estimator } \hat{\tau}(\mathbf{y};\mathbf{x}^{\rho})_{\text{GREG}}^{\pi ps} \text{ according to (5.1) and (5.2).} \tag{6.19}$$

6.4.4 Sample sizes

When sample sizes were decided on, attention was paid to *sampling rate* as well as *sample size* per se, against the following background. Performance for πps procedures commonly depends quite pronouncedly on the sampling rate, which therefore was wanted to range from "small" to "large". Approximation accuracy usually har sample size as the most vital aspect. The following sample sizes were used in the study, n = 10, 25, 50 and 80. Since population size was set to 200 [see (6.14)], sampling rates range from 5% to 40%.

6.4.5 Considered strategies

Here we adhere to the notation in (6.12). In particular, spread, which so far has been referred to by $\boldsymbol{\sigma}$, will in the sequel mostly be specified by the γ in (6.12), $\boldsymbol{\sigma}$ and γ corresponding by the relation $\boldsymbol{\sigma} = (c \cdot x_k^{\gamma} : k = 1, 2, \dots, N)$, i.e. ***spread is proportional to*** $\mathbf{x}^{\gamma}$. In combination with (6.16) - (6.19) the strategy in interest focus, [Pareto $\pi\text{ps}(\boldsymbol{\sigma}), \hat{\tau}(\mathbf{y};\boldsymbol{\sigma})_{\text{GREG}}^{\pi ps}$], is denoted;

$$[\text{PAR}(\gamma), \text{GREG}(\gamma)]. \tag{6.20}$$

Conjecture (6.1) says that (6.20) is the optimal strategy under (6.12) + (6.13), at least close to being so. When appraising the conjecture, any other admissible strategy with fixed sample size is a "challenger". The conjecture cannot be proved by a numerical study, though, which can encompass only a finite number of strategies and test situations, while it could be disproved. The latter would happen if other admissible strategies exhibit negative RVI [see (6.6)] of non-negligible magnitude. However, the conjecture is supported if no considered strategy has (substantially) negative RVI-values, the more supported the more diverse the family of alternative strategies is.

One way of classifying strategies is by mode for exploiting auxiliary information. In the present context two kinds of auxiliary information is at hand, the "y-prognostic" variable $\mathbf{x}$ and the "uncertainty measure" $\boldsymbol{\sigma}$. Both may enter into a strategy in either of the following ways. The information is (i) used in the sample design as well as in the estimator, (ii) used only in the estimator, (iii) used only in the sample design or (iv) not used at all.

The standard example of a strategy without use of auxiliary information is simple random sampling followed by straightforward estimation, [SRS, HT]. This strategy was included, not as a challenger to (6.20) though, mainly as a benchmark strategy. However, also the "naïve" SRS may be accompanied by sophisticated estimation, e.g. the strategy [SRS, GREG(γ)]. Then $\mathbf{x}$ as well as $\boldsymbol{\sigma}$ (or γ) enter in the estimation step, but not in the sample design.

A "traditional" πps strategy, as for instance [PAR(1), πps] employs only $\mathbf{x}$, as size measure in a $\pi\text{ps}(\mathbf{x})$ design, but neither $\mathbf{x}$ nor $\boldsymbol{\sigma}$ is used in the estimation step. A nearby thought is therefore that this strategy might be improved by use of auxiliary information in the estimation step, e.g.

by employing $[\text{PAR}(1),\text{GREG}(\gamma)]$. However, as discussed in Section 5.2.2, improvement along this line is a chimera. In the present notation, (5.7) says the following;

$$[\text{PAR}(1),\text{GREG}(\gamma)] = [\text{PAR}(1),\pi\text{ps}] \quad \text{for any } \gamma \geq 0. \tag{6.21}$$

Another possibility for employing just one of $\mathbf{x}$ and σ/γ is given by $[\text{PAR}(\gamma),\pi\text{ps}]$. This strategy borrows sample design from the presumed optimal strategy (6.20), but not estimator.

So far we have tacitly presumed "ideal" modeling conditions, with believed and true models equal, and the latter being linear. In practice model misspecifications occur, though. It is therefore of interest to try to find out how robust to model misspecification strategies are. This is the issue in questions (6.3) and (6.4), which both relate to cases with linear sampling - estimation model. Question (6.3) concerns misspecification of spread when also the true model is linear, while (6.4) concerns cases where the true model is non - linear. For misspecification of spread, we use the following terminology. Given that $\sigma \propto \mathbf{x}^{\gamma}$, spread is ***mildly*** respectively ***strongly overguestimated*** for $\delta \propto \mathbf{x}^{1.2\cdot\gamma}$ respectively $\delta \propto \mathbf{x}^{1.5\cdot\gamma}$. Analogously, it is ***mildly*** respectively ***strongly underguestimated*** for $\delta \propto \mathbf{x}^{0.8\cdot\gamma}$ respectively $\delta \propto \mathbf{x}^{0.5\cdot\gamma}$.

For the strategy in focus of interest, $[\text{PAR},\text{GREG}]$, the believed spread δ affects design as well as estimator. To investigate robustness against model misspecification the following strategies were considered; $\{[\text{PAR}(\kappa)),\text{GREG}(\kappa)] : \kappa = \gamma\cdot 1.2, \gamma\cdot 1.5, \gamma\cdot 0.8, \gamma\cdot 0.5\}$.

The strategies mentioned in the above discussion are listed in Table 6.2 below. Note that all of them are admissible and have fixed sample size.

<table>
<tr><th colspan="2">Table 6.2.</th><th colspan="4">Studied strategies. True spread σ is proportional to $\mathbf{x}^{\gamma}$.</th></tr>
<tr><th colspan="2" rowspan="2"></th><th colspan="4">Use of information about spread</th></tr>
<tr><th>In both steps</th><th>In the estimation step only</th><th>In the sampling step only</th><th>Not at all</th></tr>
<tr><td rowspan="4">Use of information about x</td><td>In both steps</td><td>$[\text{PAR}(1),\text{GREG}(\gamma)]$</td><td></td><td></td><td></td></tr>
<tr><td>In the estimation step only</td><td>Correct spread guestimate
$[\text{PAR}(\gamma),\text{GREG}(\gamma)]$
Overguestimated spread
$[\text{PAR}(\gamma\cdot 1.2)),\text{GREG}(\gamma\cdot 1.2)]$
$[\text{PAR}(\gamma\cdot 1.5),\text{GREG}(\gamma\cdot 1.5)]$
Underguestimated spread
$[\text{PAR}(\gamma\cdot 0.8),\text{GREG}(\gamma\cdot 0.8)]$
$[\text{PAR}(\gamma\cdot 0.5),\text{GREG}(\gamma\cdot 0.5)]$</td><td>Correct spread guestimate
$[\text{SRS},\text{GREG}(\gamma)]$</td><td></td><td></td></tr>
<tr><td>In the sampling step only</td><td>$[\text{PAR}(1),\pi\text{ps}]$</td><td></td><td>Correct spread
$[\text{PAR}(\gamma),\pi\text{ps}]$
Overguestimated spread
$[\text{PAR}(\gamma\cdot 1.2)),\pi\text{ps}]$
$[\text{PAR}(\gamma\cdot 1.5),\pi\text{ps}]$
Underguestimated spread
$[\text{PAR}(\gamma\cdot 0.8)),\pi\text{ps}]$
$[\text{PAR}(\gamma\cdot 0.5),\pi\text{ps}]$</td><td></td></tr>
<tr><td>Not at all</td><td></td><td></td><td></td><td>$[\text{SRS},\text{HT}]$</td></tr>
</table>

6.5 Conclusions from the numerical findings

Results from the numerical study are presented in the Appendix. Below we formulate our conclusions so that the reader may agree or disagree when examining the Appendix figures. Comparison of strategies is considered first, and approximation issues thereafter.

6.5.1 Comparison of strategies

The discussion is structured by the questions (6.1)-(6.4).

Conclusions relative to question (6.1) : Conjecture (6.1) concerns linear superpopulation model and correctly specified sampling-estimation model. Situations with linear superpopulation are those labeled by A, B ad C in Table 6.1. Corresponding RME and RVI values are presented in Tables A.1, A.2 and A.3 in the Appendix. As already stated, the conjecture cannot be proved, only supported or disproved. The crucial quantities are the RVI-values for strategies other than $[\mathrm{PAR}(\gamma), \mathrm{GREG}(\gamma)]$. Non-negative RVI values support the conjecture, while negative ones cast doubt over it.

The following is seen in Tables A.1, A.2 and A.3. For the considered strategies, sample sizes and spread alternatives almost all RVI are positive. Most of them solidly positive, but there are some exceptions. Firstly, in situation B $[\mathrm{PAR}(1), \mathrm{GREG}(\gamma)]$, $[\mathrm{PAR}(1), \pi ps]$ and $[\mathrm{PAR}(\gamma), \pi ps]$ are equally efficient as $[\mathrm{PAR}(\gamma), \mathrm{GREG}(\gamma)]$, all having RVI = 0. However, this is understood by what is stated in (6.21) and the fact that $\gamma = 1$ in situation B.

Secondly, and more surprising, negative RVI turn up for $[\mathrm{PAR}(\gamma \cdot 1.2)), \mathrm{GREG}(\gamma \cdot 1.2)]$. This strategy was included in the study on the "merit" model misspecification, which was expected to pull in the direction "positive RVI". However, whatever be the explanation for the negative RVI values, they are so small that our overall conclusion is as stated in (6.22) below. We do not have a clear understanding of the negative RVI, though. Possible explanations are: (i) Random disturbances due to the simulation approach. (ii). Perhaps slight over-guestimation of spread in fact is advantageous.

From survey practical point of view, the findings strongly support the conjecture that [Pareto $\pi ps(\sigma)$, $\hat{t}(y;\sigma)^{\pi ps}_{\mathrm{GREG}}$] is close to being an optimal strategy. (6.22)

Another observation from Tables A1, A2 and A3 is as follows. The "naïve" strategy [SRS, HT] is severely outperformed by all strategis which employ auxiliary information in some way. However, [SRS, GREG(γ)] yields substantial improvement of [SRS, HT], sometimes but not always it works better than $[\mathrm{PAR}(1), \pi ps]$.

Conclusions relative to question (6.2) : Also here the numerical background is given in Tables A.1, A.2 and A.3. As is seen, in line with conjecture (6.1) [Pareto $\pi ps(\sigma)$, $\hat{t}(y;\sigma)^{\pi ps}_{\mathrm{GREG}}$] never performs worse than [Pareto $\pi ps(\mathbf{x})$, $\hat{t}(y)_{\pi ps}$], and in many situations considerably better. Its superiority varies, though, from situation to situation. As already discussed, when $\gamma = 1$, the two strategies are equally good, while RVI for [Pareto $\pi ps(\mathbf{x})$, $\hat{t}(y)_{\pi ps}$] in some situations is as high as 50% and even more.

Conclusions relative to question (6.3) : Again the numerical background is provided by Tables A.1, A.2 and A.3. The findings are summarized as follows.

When true superpopulation model is linear, misspecification of spread shape has quite small adverse effect on the efficiency of the strategy [Pareto πps, $\hat{t}(y)^{\pi ps}_{\mathrm{GREG}}$]. (6.23)

Conclusions relative to question (6.4): The issue is behavior of [Pareto $\pi ps(\sigma)$, $\hat{t}(y;\sigma)^{\pi ps}_{\mathrm{GREG}}$] when the superpopulation model is judged to be linear, although it is not. Here Tables A.4, A.5

and A.6 numerical background. It is difficult to draw clear - cut conclusions, and we leave the figures to the reader's own reflections. A tentative conclusion may be as follows.

If one is in serious doubt about the shape of the trend in the superpopulation model, it may be wise to use simple random sampling with GREG estimation instead of trying to be optimal with a strategy of type [Pareto πps, $\hat{t}(y)^{\pi ps}_{GREG}$]. (6.24)

6.5.2 On approximation accuracy

We adhere to the disposition in Section 6.3.

On point estimator bias: Numerical background is given in tables A.7 - A.12. The clear message from the figures is that the point estimators under consideration work with negligible bias in all types of situations. Perhaps a warning should be issued for [SRS, GREG]. In particular is seen that the studied strategies live up to being *model assisted* in the sense that point estimators have negligible bias also under model misspecification.

On variance estimator bias: Numerical background is given in tables A.13 - A.18. Our overall conclusion is as follows. Even if the variance estimators are not exactly unbiased, they work in an acceptable way in all types of situations. A warning for the strategy [SRS, GREG] should perhaps be given also in this context.

On the basis of general experience from GREG estimation one believes that the V_2 - estimator in (5.6) should perform better than the V_1 - estimator in (5.4). On an overall basis this is confirmed by the numerical findings. The difference between the estimators is not very pronounced, though.

On accuracy in approximate formulas for theoretical estimator variances: From a survey practical point of view this is not an important issue. The most interesting aspect is perhaps if we could have dispensed of all the simulations, and based our comparisons of strategies on the approximate formulas for estimator variances. A look at the figures in Tables A.13 - A.15 shows that the approximate variance formulas sometimes can be misleading.

References

Aires, N. & Rosén, B. (2000). On Inclusion Probabilities and Estimator Bias for Pareto πps Sampling. Stat Sweden R&D Report 2000:2.

Cassel, C. M., Särndal C. - E. & Wretman, J. (1976). Some results on generalized difference estimation and generalized regression estimation for finite populations. *Biometrika* **63**, 615-620.

Cassel, C. M., Särndal C. - E. & Wretman, J. (1977). *Foundations of Inference in Survey Sampling.* Wiley & Sons, New York.

Rosén, B. (1977). On Sampling with Probability Proportional to Size. *J of Statist Planning and Inference*, **62**, 159 - 191.

Rosén, B. (1998). On Inclusion Probabilities for Order πps Sampling. *J of Statist Planning and Inference*, **90**, 117 - 143.

Särndal C. - E., Swensson, B. & Wretman, J. (1992). *Model Assisted Survey Sampling.* Springer Verlag, New York.

Appendix. Numerical results

The numerical findings are collected in Tables A.1 - A.18.

Section A.1 contains findings on performances of the different strategies. As regards "absolute" performance, the measure relative margin of error (RME) is most relevant. However, when comparing strategies it is easier to look at relative variance increase (RVI). Both are reported.

Tables A.1 - A.14 include both the strategies [PAR(1), GREG(γ)] and [PAR(1), πps] although we know they are equal, see (6.21), for the simple reason that it is easy to forget about their equivalence.

Section A.2 contains findings on approximation accuracy. Tables A.7 - A.12 concern relative point estimator bias (RBPE) and Tables A.13 - A.18 relative variance estimator bias (RBVE). Tables A.13 - A.18 also present relative errors for the approximate variance formulas.

In Tables A.13 - A.18 equivalent strategies are not duplicated. Only values for [PAR(1), πps] are listed, not for its equivalents [PAR(1), GREG(γ)] / V_1 and [PAR(1), GREG(γ)] / V_2. The reason for blank columns under RESTD for spread magnitude c_2 is that we (in last minute) came to suspect a program bug, which could not be sorted out.

A.1 Relative margins of error (RME) and relative variance increase (RVI)

Table A.1. RVI and RME in % [see (6.5) and (6.6)] for test situations of Type A [see Table 6.1]. True superpopulation model is linear ($\alpha = 1$) and spread is proportional to $\sqrt{x}$ ($\gamma=0.5$).

Strategy	n=10			n=25			n=50			n=80		
	c			c			c			c		
	0.9	1.8	3.5	0.9	1.8	3.5	0.9	1.8	3.5	0.9	1.8	3.5
				RVI in %								
Correct spread guestimate												
[PAR(γ), GREG(γ)]	0	0	0	0	0	0	0	0	0	0	0	0
[SRS, GREG(γ)]	16.8	21.9	15.6	15.4	19.9	13.3	16.5	24.2	27.2	18.7	32.0	20.2
[PAR(1), GREG(γ)]	5.8	2.7	12.4	10.3	9.1	9.3	12.1	7.8	13.0	18.8	8.3	11.7
[PAR(1), πps]	5.8	2.7	12.4	10.3	9.1	9.3	12.1	7.8	13.0	18.8	8.3	11.7
[PAR(γ), πps]	944	147	56.6	927	157	58.0	1105	184	59.6	1159	199	66.5
[SRS, HT]	4667	776	293	4797	889	288	5350	965	327	5530	976	316
Overguestimated spread												
[PAR($\gamma\cdot$1.2)), GREG($\gamma\cdot$1.2)]	0.5	-1.2	2.1	-0.4	-1.4	1.6	0	-1.6	0.4	-1.0	1.2	-0.3
[PAR($\gamma\cdot$1.5), GREG($\gamma\cdot$1.5)]	1.7	0.7	5.4	2.6	1.6	2.7	4.3	0.5	2.0	4.1	3.2	1.1
[PAR($\gamma\cdot$1.2)), πps]	578	83.9	35.7	581	93.5	37.8	680	109	37.7	701	115	40.1
[PAR($\gamma\cdot$1.5), πps]	208	26.7	14.3	211	33.7	16.7	244	35.7	15.6	275	42.7	16.6
Underguestimated spread												
[PAR($\gamma\cdot$0.8), GREG($\gamma\cdot$0.8)]	0.3	2.4	1.8	1.9	1.3	0.9	0.5	1.3	2.5	-0.9	3.2	0.1
[PAR($\gamma\cdot$0.5), GREG($\gamma\cdot$0.5)]	4.6	5.9	4.2	2.5	6.4	1.9	3.7	6.2	7.1	1.3	11.1	3.7
[PAR($\gamma\cdot$0.8)), πps]	1413	225	85.7	1405	245	84.7	1633	276	90.3	1706	296	94.9
[PAR($\gamma\cdot$0.5), πps]	2352	383	147	2352	421	137	2696	464	156	2841	500	157
				RME in %								
Correct spread guestimate												
[PAR(γ), GREG(γ)]	5.0	12.1	20.4	3.0	6.9	12.8	1.9	4.5	8.1	1.3	3.1	5.7
[SRS, GREG(γ)]	5.4	13.4	21.9	3.2	7.6	13.6	2.1	5.0	9.1	1.4	3.6	6.3
[PAR(1), GREG(γ)]	5.1	12.3	21.6	3.1	7.2	13.4	2.0	4.7	8.6	1.4	3.2	6.0
[PAR(1), πps]	5.1	12.3	21.6	3.1	7.2	13.4	2.0	4.7	8.6	1.4	3.2	6.0
[PAR(γ), πps]	16.1	19.0	25.5	9.6	11.1	16.1	6.6	7.6	10.2	4.7	5.4	7.4
[SRS, HT]	34.3	35.8	40.4	20.9	21.8	25.2	14.1	14.7	16.8	9.8	10.2	11.7
Overguestimated spread												
[PAR($\gamma\cdot$1.2)), GREG($\gamma\cdot$1.2)]	5.0	12.0	20.6	3.0	6.9	12.9	1.9	4.5	8.1	1.3	3.1	5.7
[PAR($\gamma\cdot$1.5), GREG($\gamma\cdot$1.5)]	5.0	12.1	20.9	3.0	7.0	13.0	2.0	4.5	8.2	1.3	3.2	5.8
[PAR($\gamma\cdot$1.2)), πps]	13.0	16.4	23.7	7.8	9.6	15.0	5.3	6.5	9.5	3.7	4.6	6.8
[PAR($\gamma\cdot$1.5), πps]	8.7	13.6	21.8	5.3	8.0	13.8	3.6	5.2	8.7	2.5	3.7	6.2
Underguestimated spread												
[PAR($\gamma\cdot$0.8), GREG($\gamma\cdot$0.8)]	5.0	12.2	20.6	3.0	7.0	12.9	1.9	4.5	8.2	1.3	3.2	5.7
[PAR($\gamma\cdot$0.5), GREG($\gamma\cdot$0.5)]	5.1	12.4	20.8	3.0	7.1	12.9	1.9	4.6	8.4	1.3	3.3	5.8
[PAR($\gamma\cdot$0.8)), πps]	19.4	21.8	27.8	11.6	12.9	17.4	8.0	8.7	11.2	5.6	6.2	8.0
[PAR($\gamma\cdot$0.5), πps]	24.6	26.6	32.0	14.8	15.8	19.7	10.1	10.7	13.0	7.1	7.6	9.2

Table A.2. RVI and RME in % [see (6.5) and (6.6)] for test situations of Type B [see Table 6.1]. True superpopulation model is linear ($\alpha = 1$) and spread is proportional to x ($\gamma = 1$).

Strategy	$n=10$			$n=25$			$n=50$			$n=80$		
	c			c			c			c		
	0.06	0.12	0.25	0.06	0.12	0.25	0.06	0.12	0.25	0.06	0.12	0.25
					RVI in %							
Correct spread guestimate												
[PAR(γ), GREG(γ)]	0	0	0	0	0	0	0	0	0	0	0	0
[SRS, GREG(γ)]	38.8	46.8	32.6	38.8	44.6	37.9	45.3	60.2	64.2	56.7	89.6	69.4
[PAR(1), GREG(γ)]	0	0	0	0	0	0	0	0	0	0	0	0
[PAR(1), πps]	0	0	0	0	0	0	0	0	0	0	0	0
[PAR(γ), πps]	0	0	0	0	0	0	0	0	0	0	0	0
[SRS, HT]	9850	1560	517	10029	1769	543	11586	2025	632	12645	2298	695
Overguestimated spread												
[PAR(γ·1.2)), GREG(γ·1.2)]	1.9	-0.9	-0.9	1.7	0.5	-1.1	3.3	-0.8	0.1	1.4	-1.6	-1.2
[PAR(γ·1.5), GREG(γ·1.5)]	4.4	1.4	4.2	6.4	2.6	2.8	11.5	2.7	5.0	10.4	3.7	6.2
[PAR(γ·1.2)), πps]	324	58.0	16.4	342	64.4	13.5	412	76.5	20.8	419	77.8	19.5
[PAR(γ·1.5), πps]	2074	350	114	2184	393	106	2617	470	130	2823	519	149
Underguestimated spread												
[PAR(γ·0.8), GREG(γ·0.8)]	3.0	2.2	-0.6	0.7	1.1	3.1	2.2	3.9	2.3	2.8	9.6	4.5
[PAR(γ·0.5), GREG(γ·0.5)]	8.1	10.5	3.8	8.7	10.0	10.0	8.7	12.8	14.2	10.5	21.5	19.6
[PAR(γ·0.8)), πps]	295	42.6	12.4	297	46.4	20.3	344	54.5	18.6	404	73.7	25.8
[PAR(γ·0.5), πps]	2028	316	102	1979	331	115	2439	410	127	2686	491	158
					RME in %							
Correct spread guestimate												
[PAR(γ), GREG(γ)]	3.4	8.6	15.4	2.1	5.0	9.4	1.3	3.1	5.9	0.9	2.1	3.9
[SRS, GREG(γ)]	4.1	10.5	17.7	2.4	6.0	11.0	1.6	4.0	7.6	1.1	2.8	5.1
[PAR(1), GREG(γ)]	3.4	8.6	15.4	2.1	5.0	9.4	1.3	3.1	5.9	0.9	2.1	3.9
[PAR(1), πps]	3.4	8.6	15.4	2.1	5.0	9.4	1.3	3.1	5.9	0.9	2.1	3.9
[PAR(γ), πps]	3.4	8.6	15.4	2.1	5.0	9.4	1.3	3.1	5.9	0.9	2.1	3.9
[SRS, HT]	34.4	35.2	38.3	20.9	21.4	23.9	14.2	14.5	16.0	9.9	10.1	11.1
Overguestimated spread												
[PAR(γ·1.2)), GREG(γ·1.2)]	3.5	8.6	15.3	2.1	5.0	9.4	1.3	3.1	5.9	0.9	2.0	3.9
[PAR(γ·1.5), GREG(γ·1.5)]	3.5	8.7	15.7	2.1	5.0	9.5	1.4	3.2	6.0	0.9	2.1	4.0
[PAR(γ·1.2)), πps]	7.1	10.9	16.6	4.4	6.4	10.0	3.0	4.2	6.5	2.0	2.7	4.3
[PAR(γ·1.5), πps]	16.1	18.3	22.6	9.9	11.0	13.5	6.8	7.5	8.9	4.7	5.1	6.2
Underguestimated spread												
[PAR(γ·0.8), GREG(γ·0.8)]	3.5	8.7	15.4	2.1	5.0	9.6	1.3	3.2	6.0	0.9	2.2	4.0
[PAR(γ·0.5), GREG(γ·0.5)]	3.6	9.1	15.7	2.2	5.2	9.9	1.4	3.3	6.3	0.9	2.3	4.3
[PAR(γ·0.8)), πps]	6.8	10.3	16.3	4.1	6.0	10.3	2.8	3.9	6.4	2.0	2.7	4.4
[PAR(γ·0.5), πps]	15.9	17.6	21.9	9.5	10.3	13.8	6.6	7.1	8.9	4.6	5.0	6.3

Table A.3. RVI and RME in % [see (6.5) and (6.6)] for test situations of Type C [see Table 6.1]. True superpopulation model is linear ($\alpha = 1$) and spread is proportional to $x^{1/4}$ ($\gamma = 0.25$).												
Strategy	n=10			n=25			n=50			n=80		
	c			c			c			c		
	3	7	13	3	7	13	3	7	13	3	7	13
				RVI in %								
Correct spread guestimate												
[PAR(γ), GREG(γ)]	0	0	0	0	0	0	0	0	0	0	0	0
[SRS, GREG(γ)]	6.5	10.7	5.2	8.3	8.1	5.4	5.6	10.5	11.7	8.8	10.9	7.7
[PAR(1), GREG(γ)]	37.9	20.0	46.4	47.8	36.0	45.1	52.2	40.7	61.1	81.1	46.2	67.7
[PAR(1), πps]	37.9	20.0	46.4	47.8	36.0	45.1	52.2	40.7	61.1	81.1	46.2	67.7
[PAR(γ), πps]	2027	239	98.5	2091	268	97	2326	291	103	2505	298	107
[SRS, HT]	4044	495	197	4295	576	206	4634	619	220	4917	595	216
Overguestimated spread												
[PAR(γ· 1.2)), GREG(γ· 1.2)]	-0.7	-0.4	0.2	2.0	-0.9	0.7	0.2	-1.2	0.7	0.7	-2.1	-0.2
[PAR(γ· 1.5), GREG(γ· 1.5)]	-0.5	-0.2	1.0	3.6	-1.4	2.6	1.6	-1.0	0.3	3.4	-0.4	1.9
[PAR(γ· 1.2)), πps]	1717	199	82.7	1786	225	84.4	2030	251	90.6	2124	256	91.1
[PAR(γ· 1.5), πps]	1334	152	63.7	1370	170	68.0	1555	190	67.9	1645	195	73.2
Underguestimated spread												
[PAR(γ· 0.8), GREG(γ· 0.8)]	-0.4	1.5	0.3	2.0	1.4	0.3	0.4	1.5	1.0	1.1	0.5	0
[PAR(γ· 0.5), GREG(γ· 0.5)]	2.1	3.4	2.5	2.7	2.8	1.5	1.5	3.0	2.9	4.1	1.9	1.3
[PAR(γ· 0.8)), πps]	2383	283	115	2486	322	116	2673	338	118	2919	348	125
[PAR(γ· 0.5), πps]	2909	349	142	3071	402	143	3311	425	150	3574	432	154
				RME in %								
Correct spread guestimate												
[PAR(γ), GREG(γ)]	5.3	14.9	24.5	3.1	8.5	15.1	2.1	5.5	9.8	1.4	3.9	6.9
[SRS, GREG(γ)]	5.5	15.7	25.2	3.3	8.9	15.5	2.1	5.8	10.3	1.5	4.1	7.2
[PAR(1), GREG(γ)]	6.2	16.3	29.7	3.8	9.9	18.2	2.5	6.6	12.4	1.9	4.8	9.0
[PAR(1), πps]	6.2	16.3	29.7	3.8	9.9	18.2	2.5	6.6	12.4	1.9	4.8	9.0
[PAR(γ), πps]	24.5	27.4	34.6	14.7	16.3	21.2	10.1	11.0	13.9	7.1	7.9	10.0
[SRS, HT]	34.2	36.3	42.2	20.9	22.1	26.4	14.1	14.9	17.5	9.9	10.4	12.3
Overguestimated spread												
[PAR(γ· 1.2)), GREG(γ· 1.2)]	5.3	14.9	24.6	3.2	8.5	15.2	2.1	5.5	9.8	1.4	3.9	6.9
[PAR(γ· 1.5), GREG(γ· 1.5)]	5.3	14.9	24.7	3.2	8.5	15.3	2.1	5.5	9.8	1.4	3.9	7.0
[PAR(γ· 1.2)), πps]	22.7	25.8	33.2	13.7	15.3	20.5	9.5	10.4	13.5	6.6	7.4	9.6
[PAR(γ· 1.5), πps]	20.1	23.6	31.4	12.1	14.0	19.6	8.4	9.4	12.7	5.8	6.8	9.1
Underguestimated spread												
[PAR(γ· 0.8), GREG(γ· 0.8)]	5.3	15.0	24.6	3.2	8.6	15.1	2.1	5.6	9.8	1.4	4.0	6.9
[PAR(γ· 0.5), GREG(γ· 0.5)]	5.4	15.1	24.8	3.2	8.6	15.2	2.1	5.6	9.9	1.4	4.0	7.0
[PAR(γ· 0.8)), πps]	26.5	29.1	36.0	16.0	17.5	22.2	10.8	11.6	14.4	7.6	8.3	10.4
[PAR(γ· 0.5), πps]	29.2	31.6	38.2	17.7	19.1	23.6	12.0	12.7	15.4	8.4	9.1	11.0

Table A.4. RVI and RME in % [see (6.5) and (6.6)] for test situations of Type D [see Table 6.1]. True superpopulation model is mildly convex ($\alpha = 1.2$) and spread is proportional to $\sqrt{x}$ ($\gamma = 0.5$).

Strategy	n=10			n=25			n=50			n=80		
	c			c			c			c		
	2.5	5	10	2.5	5	10	2.5	5	10	2.5	5	10
				RVI in %								
Correct spread guestimate												
[PAR(γ), GREG(γ)]	0	0	0	0	0	0	0	0	0	0	0	0
[SRS, GREG(γ)]	1.5	18.5	13.4	3.0	16.8	10.7	-3.3	19.1	24.8	3.4	23.4	17.7
[PAR(1), GREG(γ)]	13.4	5.7	11.8	18.8	12.8	9.3	20.7	10.8	13.6	27.7	10.2	11.7
[PAR(1), πps]	13.4	5.7	11.8	18.8	12.8	9.3	20.7	10.8	13.6	27.7	10.2	11.7
[PAR(γ), πps]	761	199	71.2	790	218	72.3	880	254	77.8	906	252	81
[SRS, HT]	2960	802	291	3183	927	288	3331	1017	327	3493	989	316
Overguestimated spread												
[PAR($\gamma\cdot$1.2)), GREG($\gamma\cdot$1.2)]	3.0	-1.3	2.1	4.3	-0.7	1.7	3.4	-0.8	0.7	1.8	0.3	-0.3
[PAR($\gamma\cdot$1.5), GREG($\gamma\cdot$1.5)]	5.7	1.1	4.4	9.8	3.6	2.7	8.3	2.0	2.8	13.6	3.9	1.2
[PAR($\gamma\cdot$1.2)), πps]	523	131	49.9	558	150	51.8	609	173	55.0	610	165	54
[PAR($\gamma\cdot$1.5), πps]	254	61.9	24.7	277	76.1	27.4	296	83.6	29.4	320	85	28
Underguestimated spread												
[PAR($\gamma\cdot$0.8), GREG($\gamma\cdot$0.8)]	-0.3	1.5	1.5	-0.6	1.0	-0.2	-3.7	-0.1	1.8	-3.3	1.3	-0.7
[PAR($\gamma\cdot$0.5), GREG($\gamma\cdot$0.5)]	-2.2	3.9	3.7	-1.6	4.8	0.2	-5.7	3.8	6.7	-5.9	7.6	1.9
[PAR($\gamma\cdot$0.8)), πps]	1054	278	99.5	1104	309	97.9	1192	347	108	1239	346	108
[PAR($\gamma\cdot$0.5), πps]	1611	431	157	1695	481	147	1798	529	169	1916	542	167
				RME in %								
Correct spread guestimate												
[PAR(γ), GREG(γ)]	7.0	13.3	22.9	4.1	7.6	14.3	2.7	4.9	9.1	1.9	3.5	6.4
[SRS, GREG(γ)]	7.1	14.5	24.4	4.2	8.2	15.1	2.7	5.4	10.2	1.9	3.9	7.0
[PAR(1), GREG(γ)]	7.5	13.7	24.3	4.5	8.1	15.0	3.0	5.2	9.7	2.1	3.6	6.8
[PAR(1), πps]	7.5	13.7	24.3	4.5	8.1	15.0	3.0	5.2	9.7	2.1	3.6	6.8
[PAR(γ), πps]	20.6	23.0	30.0	12.3	13.5	18.8	8.5	9.3	12.2	5.9	6.5	8.7
[SRS, HT]	38.7	40.0	45.3	23.6	24.4	28.2	16.0	16.4	18.9	11.1	11.4	13.1
Overguestimated spread												
[PAR($\gamma\cdot$1.2)), GREG($\gamma\cdot$1.2)]	7.1	13.2	23.2	4.2	7.6	14.5	2.8	4.9	9.2	1.9	3.5	6.4
[PAR($\gamma\cdot$1.5), GREG($\gamma\cdot$1.5)]	7.2	13.4	23.4	4.3	7.7	14.5	2.8	5.0	9.2	2.0	3.5	6.5
[PAR($\gamma\cdot$1.2)), πps]	17.5	20.3	28.1	10.5	12.0	17.7	7.3	8.1	11.4	4.9	5.6	8.0
[PAR($\gamma\cdot$1.5), πps]	13.2	17.0	25.6	8.0	10.1	16.2	5.4	6.7	10.4	3.8	4.7	7.3
Underguestimated spread												
[PAR($\gamma\cdot$0.8), GREG($\gamma\cdot$0.8)]	7.0	13.4	23.1	4.1	7.6	14.3	2.7	4.9	9.2	1.8	3.5	6.4
[PAR($\gamma\cdot$0.5), GREG($\gamma\cdot$0.5)]	6.9	13.6	23.4	4.1	7.8	14.4	2.6	5.0	9.4	1.8	3.6	6.5
[PAR($\gamma\cdot$0.8)), πps]	23.8	25.9	32.4	14.3	15.4	20.2	9.8	10.4	13.1	6.8	7.3	9.3
[PAR($\gamma\cdot$0.5), πps]	29.0	30.7	36.8	17.4	18.3	22.6	11.9	12.3	15.0	8.3	8.8	10.5

Table A.5. RVI and RME in % [see (6.5) and (6.6)] for test situations of Type E [see Table 6.1]. True superpopulation model is strongly convex ($\alpha = 1.5$)and spread is proportional to $\sqrt{x}$ ($\gamma = 0.5$).

Strategy	n=10			n=25			n=50			n=80		
	c			c			c			c		
	12	25	50	12	25	50	12	25	50	12	25	50
					RVI in %							
Correct spread guestimate												
[PAR(γ), GREG(γ)]	0	0	0	0	0	0	0	0	0	0	0	0
[SRS, GREG(γ)]	-3.6	11.6	10.6	-7.0	8.4	6.8	-11.8	9.4	20.0	-3.8	12.2	14.4
[PAR(1), GREG(γ)]	13.4	8.1	11.2	16.6	14.5	9.1	18.3	11.9	15.5	21.4	11.1	11.3
[PAR(1), πps]	13.4	8.1	11.2	16.6	14.5	9.1	18.3	11.9	15.5	21.4	11.1	11.3
[PAR(γ), πps]	344	163	66.2	362	177	67.8	390	198	75.0	410	198	74.4
[SRS, HT]	1171	568	237	1257	646	237	1294	691	278	1406	686	261
Overguestimated spread												
[PAR($\gamma\cdot$ 1.2)), GREG($\gamma\cdot$ 1.2)]	3.3	-0.7	2.2	5.8	1.5	2.0	4.9	1.6	1.9	2.4	-0.2	0.1
[PAR($\gamma\cdot$ 1.5), GREG($\gamma\cdot$ 1.5)]	6.5	2.2	4.2	10.9	6.1	3.4	10.7	4.5	5.2	14.0	5.4	2.1
[PAR($\gamma\cdot$ 1.2)), πps]	249	114	48.7	270	130	51.0	285	144	56.4	290	135	51.8
[PAR($\gamma\cdot$ 1.5), πps]	135	60.3	26.2	149	73.7	29.1	156	77.9	34.0	170	78.0	29.6
Underguestimated spread												
[PAR($\gamma\cdot$ 0.8), GREG($\gamma\cdot$ 0.8)]	-1.4	0.1	0.9	-2.3	-0.2	-1.2	-5.2	-2.5	0.8	-3.7	-0.8	-1.6
[PAR($\gamma\cdot$ 0.5), GREG($\gamma\cdot$ 0.5)]	-5.3	0.3	2.6	-5.7	0.9	-1.6	-9.1	-1.7	4.7	-7.2	3.0	0
[PAR($\gamma\cdot$ 0.8)), πps]	458	217	89.1	486	241	88.5	507	258	98.8	542	261	95.7
[PAR($\gamma\cdot$ 0.5), πps]	669	321	134	709	356	128	740	379	150	806	394	143
					RME in %							
Correct spread guestimate												
[PAR(γ), GREG(γ)]	12.6	17.9	28.7	7.4	10.3	17.8	4.9	6.7	11.2	3.3	4.7	8.0
[SRS, GREG(γ)]	12.3	18.9	30.2	7.1	10.7	18.4	4.6	7.0	12.3	3.2	5.0	8.6
[PAR(1), GREG(γ)]	13.4	18.6	30.3	8.0	11.0	18.6	5.4	7.1	12.1	3.6	4.9	8.4
[PAR(1), πps]	13.4	18.6	30.3	8.0	11.0	18.6	5.4	7.1	12.1	3.6	4.9	8.4
[PAR(γ), πps]	26.5	29.0	37.0	15.8	17.1	23.1	10.9	11.6	14.8	7.4	8.1	10.6
[SRS, HT]	44.8	46.2	52.8	27.2	28.0	32.8	18.4	18.9	21.8	12.8	13.1	15.2
Overguestimated spread												
[PAR($\gamma\cdot$ 1.2)), GREG($\gamma\cdot$ 1.2)]	12.8	17.8	29.0	7.6	10.3	18.0	5.1	6.8	11.3	3.3	4.7	8.0
[PAR($\gamma\cdot$ 1.5), GREG($\gamma\cdot$ 1.5)]	13.0	18.1	29.3	7.8	10.6	18.1	5.2	6.9	11.5	3.5	4.8	8.1
[PAR($\gamma\cdot$ 1.2)), πps]	23.5	26.1	35.0	14.2	15.6	21.9	9.7	10.5	14.0	6.5	7.2	9.9
[PAR($\gamma\cdot$ 1.5), πps]	19.2	22.6	32.3	11.6	13.5	20.3	7.9	9.0	13.0	5.4	6.2	9.1
Underguestimated spread												
[PAR($\gamma\cdot$ 0.8), GREG($\gamma\cdot$ 0.8)]	12.5	17.9	28.9	7.3	10.3	17.7	4.8	6.6	11.3	3.2	4.7	7.9
[PAR($\gamma\cdot$ 0.5), GREG($\gamma\cdot$ 0.5)]	12.2	17.9	29.1	7.2	10.3	17.7	4.7	6.7	11.5	3.2	4.8	8.0
[PAR($\gamma\cdot$ 0.8)), πps]	29.7	31.8	39.5	17.9	18.9	24.5	12.2	12.7	15.8	8.3	8.9	11.2
[PAR($\gamma\cdot$ 0.5), πps]	34.8	36.7	44.0	21.0	21.9	26.9	14.3	14.7	17.7	9.9	10.4	12.5

Table A.6. RVI and RME in % [see (6.5) and (6.6)] for test situations of Type F [see Table 6.1]. True superpopulation model is concave ($\alpha = 0.7$) and spread is proportional to $\sqrt{x}$ ($\gamma = 0.5$).

Strategy	n=10			n=25			n=50			n=80		
	c			c			c			c		
	0.2	**0.4**	**0.7**	**0.2**	**0.4**	**0.7**	**0.2**	**0.4**	**0.7**	**0.2**	**0.4**	**0.7**
				RMI in %								
Correct spread guestimate												
[PAR(γ), GREG(γ)]	0	0	0	0	0	0	0	0	0	0	0	0
[SRS, GREG(γ)]	-20.5	0.5	5.6	-29.5	-6.1	3.8	-30.2	-6.5	9.8	-33.1	1.5	8.4
[PAR(1), GREG(γ)]	41.7	19.3	23.0	47.1	26.0	20.5	56.7	32.1	23.5	67.6	41.5	29.3
[PAR(1), πps]	41.7	19.3	23.0	47.1	26.0	20.5	56.7	32.1	23.5	67.6	41.5	29.3
[PAR(γ), πps]	-35.6	-24.0	-10.5	-36.2	-26.3	-2.7	-30.8	-26.1	-12.1	-25.6	-17.5	-2.5
[SRS, HT]	579	273	175	611	307	194	632	316	195	674	349	209
Overguestimated spread												
[PAR($\gamma\cdot$1.2)), GREG($\gamma\cdot$1.2)]	8.2	4.2	3.7	10.2	4.7	4.4	11.9	5.2	3.1	9.9	7.8	3.1
[PAR($\gamma\cdot$1.5), GREG($\gamma\cdot$1.5)]	22.6	12.3	13.0	28.7	15.1	9.7	34.1	17.8	9.5	34.2	23.1	11.9
[PAR($\gamma\cdot$1.2)), πps]	-70.3	-40.8	-19.5	-70.1	-42.9	-13.1	-70.2	-44.6	-21.9	-68.8	-39.2	-15.9
[PAR($\gamma\cdot$1.5), πps]	-76.2	-40.8	-19.2	-74.7	-41.7	-16.3	-75.3	-43.7	-21.9	-75.8	-39.6	-18.0
Underguestimated spread												
[PAR($\gamma\cdot$0.8), GREG($\gamma\cdot$0.8)]	-9.1	-2.7	-0.9	-8.3	-4.1	-0.4	-11.3	-5.0	-1.1	-11.2	-3.5	-1.7
[PAR($\gamma\cdot$0.5), GREG($\gamma\cdot$0.5)]	-16.8	-4.7	-3.0	-20.7	-7.4	-2.1	-21.3	-8.0	-1.9	-21.7	-4.8	-1.4
[PAR($\gamma\cdot$0.8)), πps]	25.6	5.4	7.5	26.0	4.7	15.8	36.3	6.0	7.4	45.7	19.9	17.7
[PAR($\gamma\cdot$0.5), πps]	172	75.2	52.3	174	79.5	58.6	193	82.4	54.2	216	110	67.3
				RME in %								
Correct spread guestimate												
[PAR(γ), GREG(γ)]	10.3	14.8	19.2	6.1	8.6	11.7	4.1	5.8	7.7	2.8	3.9	5.3
[SRS, GREG(γ)]	9.2	14.9	19.7	5.2	8.4	11.9	3.4	5.6	8.1	2.3	3.9	5.5
[PAR(1), GREG(γ)]	12.2	16.2	21.3	7.4	9.7	12.8	5.1	6.6	8.6	3.6	4.6	6.0
[PAR(1), πps]	12.2	16.2	21.3	7.4	9.7	12.8	5.1	6.6	8.6	3.6	4.6	6.0
[PAR(γ), πps]	8.3	12.9	18.1	4.9	7.4	11.5	3.4	4.9	7.3	2.4	3.5	5.2
[SRS, HT]	26.8	28.6	31.8	16.4	17.4	20.0	11.1	11.7	13.3	7.8	8.2	9.3
Overguestimated spread												
[PAR($\gamma\cdot$1.2)), GREG($\gamma\cdot$1.2)]	10.7	15.1	19.5	6.4	8.8	11.9	4.3	5.9	7.9	2.9	4.0	5.4
[PAR($\gamma\cdot$1.5), GREG($\gamma\cdot$1.5)]	11.4	15.7	20.4	7.0	9.2	12.2	4.7	6.2	8.1	3.2	4.3	5.6
[PAR($\gamma\cdot$1.2)), πps]	5.6	11.4	17.2	3.4	6.5	10.9	2.2	4.3	6.8	1.6	3.0	4.9
[PAR($\gamma\cdot$1.5), πps]	5.0	11.4	17.2	3.1	6.6	10.7	2.0	4.3	6.8	1.4	3.0	4.8
Underguestimated spread												
[PAR($\gamma\cdot$0.8), GREG($\gamma\cdot$0.8)]	9.8	14.6	19.1	5.9	8.4	11.7	3.9	5.6	7.7	2.6	3.8	5.3
[PAR($\gamma\cdot$0.5), GREG($\gamma\cdot$0.5)]	9.4	14.5	18.9	5.5	8.3	11.6	3.6	5.5	7.7	2.5	3.8	5.3
[PAR($\gamma\cdot$0.8)), πps]	11.5	15.2	19.9	6.9	8.8	12.6	4.8	5.9	8.0	3.4	4.2	5.8
[PAR($\gamma\cdot$0.5), πps]	16.9	19.6	23.7	10.2	11.5	14.7	7.0	7.8	9.6	5.0	5.6	6.9

A.2 On approximation accuracy

A.2.1 Point estimator bias

Table A.7 **RBPE according to (6.7)) for test situations of Type A [see Table 6.1].**

Strategy	n=10			n=25			n=50			n=80		
	c=.9	c=1.8	c=3.5	c=0.9	c=1.8	c=3.5	c=0.9	c=1.8	c=3.5	c=0.9	c=1.8	c=3.5
Correct spread guestimate												
[PAR(γ), GREG(γ)]	0.0	0.1	-0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[SRS, GREG(γ)]	0.0	0.1	0.0	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
[PAR(1), GREG(γ)]	0.0	0.0	-0.1	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
[PAR(1), πps]	0.0	0.0	-0.1	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
[PAR(γ), πps]	-0.1	-0.1	-0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[SRS, HT]	-0.1	0.0	-0.1	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
Overguestimated spread												
[PAR($\gamma\cdot$1.2)), GREG($\gamma\cdot$1.2)]	0.0	0.1	-0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot$1.5), GREG($\gamma\cdot$1.5)]	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot$1.2)), πps]	-0.1	0.0	-0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot$1.5), πps]	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Underguestimated spread												
[PAR($\gamma\cdot$0.8), GREG($\gamma\cdot$0.8)]	0.0	0.1	-0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot$0.5), GREG($\gamma\cdot$0.5)]	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot$0.8)), πps]	-0.1	-0.1	-0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot$0.5), πps]	-0.2	-0.2	-0.3	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0

Table A.8 **RBPE according to (6.7)) for test situations of Type B [see Table 6.1]**

Strategy	n=10			n=25			n=50			n=80		
	c=.06	c=.12	c=.25	c=.06	c=.12	c=.25	c=.06	c=.12	c=.25	c=.06	c=.12	c=.25
Correct spread guestimate												
[PAR(γ), GREG(γ)]	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[SRS, GREG(γ)]	0.0	0.1	0.0	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
[PAR(1), GREG(γ)]	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR(1), πps]	-0.1	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR(γ), πps]	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[SRS, HT]	-0.1	0.0	-0.1	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
Overguestimated spread												
[PAR($\gamma\cdot$1.2)), GREG($\gamma\cdot$1.2)]	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot$1.5), GREG($\gamma\cdot$1.5)]	0.0	0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot$1.2)), πps]	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot$1.5), πps]	0.0	0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Underguestimated spread												
[PAR($\gamma\cdot$0.8), GREG($\gamma\cdot$0.8)]	0.0	0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot$0.5), GREG($\gamma\cdot$0.5)]	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot$0.8)), πps]	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot$0.5), πps]	-0.1	-0.1	-0.3	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0

Table A.9	**RBPE according to (6.7)) for test situations of Type C [see Table 6.1]**											
Strategy	n=10			n=25			n=50			n=80		
	c=3	c=7	c=13	c=3	c=7	c=13	c=3	c=7	c=13	c=3	c=7	c=13
Correct spread guestimate												
[PAR(γ),GREG(γ)]	0.0	0.1	-0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[SRS,GREG(γ)]	0.0	0.1	0.0	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
[PAR(1),GREG(γ)]	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR(1),πps]	0.0	0.0	-0.1	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
[PAR(γ),πps]	-0.2	-0.2	-0.3	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[SRS,HT]	0.0	0.0	-0.1	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
Overguestimated spread												
[PAR($\gamma\cdot 1.2$)),GREG($\gamma\cdot 1.2$)]	0.0	0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot 1.5$),GREG($\gamma\cdot 1.5$)]	0.0	0.1	-0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot 1.2$)),πps]	-0.1	-0.1	-0.2	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot 1.5$),πps]	-0.1	-0.1	-0.2	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
Underguestimated spread												
[PAR($\gamma\cdot 0.8$),GREG($\gamma\cdot 0.8$)]	0.0	0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot 0.5$),GREG($\gamma\cdot 0.5$)]	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot 0.8$)),πps]	-0.1	-0.1	-0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot 0.5$),πps]	-0.1	0.0	-0.2	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0

Table A.10	**RBPE according to (6.7)) for test situations of Type D [see Table 6.1]**											
Strategy	n=10			n=25			n=50			n=80		
	c=2.5	c=5	c=10	c=2.5	c=5	c=10	c=2.5	c=5	c=10	c=2.5	c=5	c=10
Correct spread guestimate												
[PAR(γ),GREG(γ)]	-0.2	-0.1	-0.5	-0.1	0.0	0.0	0.0	0.0	-0.1	0.0	0.0	0.0
[SRS,GREG(γ)]	-0.3	-0.2	-0.4	-0.1	-0.1	-0.2	-0.1	0.0	-0.1	0.0	0.0	0.0
[PAR(1),GREG(γ)]	0.0	0.0	-0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR(1),πps]	0.0	0.0	-0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR(γ),πps]	-0.1	-0.1	-0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[SRS,HT]	-0.1	0.0	-0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Overguestimated spread												
[PAR($\gamma\cdot 1.2$)),GREG($\gamma\cdot 1.2$)]	-0.2	-0.1	-0.4	-0.1	0.0	0.0	0.0	-0.1	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot 1.5$),GREG($\gamma\cdot 1.5$)]	-0.1	-0.1	-0.2	-0.1	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot 1.2$)),πps]	-0.1	-0.1	-0.3	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0
[PAR($\gamma\cdot 1.5$),πps]	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Underguestimated spread												
[PAR($\gamma\cdot 0.8$),GREG($\gamma\cdot 0.8$)]	-0.2	-0.2	-0.5	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot 0.5$),GREG($\gamma\cdot 0.5$)]	-0.2	-0.2	-0.4	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot 0.8$)),πps]	-0.1	-0.1	0.4	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
[PAR($\gamma\cdot 0.5$),πps]	-0.2	-0.2	-0.4	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table A.11	RBPE according to (6.7)) for test situations of Type E [see Table 6.1]											
Strategy	n=10			n=25			n=50			n=80		
	c=12	c=25	c=50	c=12	c=25	c=50	c=12	c=25	c=50	c=12	c=25	c=50
Correct spread guestimate												
[PAR(γ),GREG(γ)]	-0.4	-0.8	-0.7	-0.2	-0.1	-0.1	-0.1	0.0	-0.1	0.0	0.0	0.0
[SRS,GREG(γ)]	-0.7	-0.5	-0.8	-0.2	-0.2	-0.3	-0.1	-0.1	-0.1	-0.1	0.0	-0.1
[PAR(1),GREG(γ)]	0.0	0.0	-0.2	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.1
[PAR(1),πps]	0.0	0.0	-0.2	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.1
[PAR(γ),πps]	-0.2	-0.1	-0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
[SRS,HT]	0.0	0.0	-0.1	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
Overguestimated spread												
[PAR(γ· 1.2)),GREG(γ· 1.2)]	-0.4	-0.4	-0.7	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	0.0	0.0	0.0
[PAR(γ· 1.5),GREG(γ· 1.5)]	-0.3	-0.3	-0.5	-0.1	-0.1	-0.1	-0.1	0.0	-0.1	0.0	0.0	0.0
[PAR(γ· 1.2)),πps]	-0.1	-0.1	-0.4	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
[PAR(γ· 1.5),πps]	-0.1	-0.1	-0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Underguestimated spread												
[PAR(γ· 0.8),GREG(γ· 0.8)]	-0.5	-0.4	-0.8	-0.2	-0.1	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0
[PAR(γ· 0.5),GREG(γ· 0.5)]	-0.6	-0.5	-0.8	-0.2	-0.2	-0.1	-0.1	0.0	-0.1	0.0	0.0	0.0
[PAR(γ· 0.8)),πps]	-0.2	-0.1	-0.4	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.1
[PAR(γ· 0.5),πps]	-0.2	-0.2	-0.4	-0.1	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0

Table A.12	RBPE according to (6.7)) for test situations of Type F [see Table 6.1]											
Strategy	n=10			n=25			n=50			n=80		
	c=0.2	c=0.4	c=0.7	c=0.2	c=0.4	c=0.7	c=0.2	c=0.4	c=0.7	c=0.2	c=0.4	c=0.7
Correct spread guestimate												
[PAR(γ),GREG(γ)]	0.4	0.5	0.3	0.2	0.2	0.2	0.0	0.1	0.0	0.0	0.0	0.0
[SRS,GREG(γ)]	0.5	0.6	0.4	0.1	0.2	0.1	0.1	0.1	0.1	0.0	0.0	0.0
[PAR(1),GREG(γ)]	0.0	0.0	-0.1	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
[PAR(1),πps]	0.0	0.0	-0.1	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
[PAR(γ),πps]	0.1	0.1	-0.2	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
[SRS,HT]	0.0	0.1	-0.1	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
Overguestimated spread												
[PAR(γ· 1.2)),GREG(γ· 1.2)]	0.4	0.4	0.2	0.1	0.2	0.2	0.1	0.1	0.1	0.0	0.0	0.0
[PAR(γ· 1.5),GREG(γ· 1.5)]	0.4	0.4	0.3	0.1	0.1	0.2	0.1	0.1	0.1	0.0	0.0	0.0
[PAR(γ· 1.2)),πps]	0.0	0.0	-0.2	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
[PAR(γ· 1.5),πps]	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Underguestimated spread												
[PAR(γ· 0.8),GREG(γ· 0.8)]	0.4	0.5	0.2	0.1	0.2	0.2	0.0	0.1	0.0	0.0	0.0	0.0
[PAR(γ· 0.5),GREG(γ· 0.5)]	0.4	0.5	0.3	0.1	0.2	0.2	0.1	0.1	0.1	0.0	0.0	0.1
[PAR(γ· 0.8)),πps]	0.0	0.0	-0.2	0.1	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
[PAR(γ· 0.5),πps]	-0.1	-0.1	-0.2	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

A.2.2 Variance estimator bias and approximation accuracy for theoretical variances

Table A.13. RBVE and RESTD in % for test situations of Type A. See (6.8), (6.9) and Table 6.1.

Strategy and variance estimator	n=10			n=25			n=50			n=80		
	c=1	c=2.5	c=5	c=1	c=2.5	c=5	c=1	c=2.5	c=5	c=1	c=2.5	c=5
					RBVE							
Correct spread guestimate												
[PAR(γ),GREG(γ)]/V_1	-2.9	-5.3	0.4	-2.4	5.8	-8.4	-1.3	4.0	-4.2	0.4	1.8	-8.1
[PAR(γ),GREG(γ)]/V_2	-0.4	-3.9	2.7	-1.7	6.4	-7.6	-0.9	4.2	-3.8	0.7	2.0	-7.9
[SRS,GREG(γ)]/V_1	-12.0	-14.8	-7.6	-5.1	2.9	-9.2	-1.3	2.3	-12.1	3.7	0.6	-6.3
[SRS,GREG(γ)]/V_2	-8.7	-13.8	-4.9	-3.7	3.9	-7.7	-0.8	2.7	-11.8	4.1	0.8	-6.0
[PAR(1),πps]	0.8	-3.2	-1.5	-2.0	2.2	-6.3	-0.4	3.0	-2.1	-1.2	2.5	-1.3
[PAR(γ),πps]	-0.4	-2.0	3.3	4.2	6.4	-5.6	-6.4	-1.7	-0.6	-1.2	2.5	-1.3
[SRS,HT]	2.9	2.0	4.4	2.3	1.5	-1.7	-4.2	-4.2	-4.4	-3.9	-3.6	-6.5
Overguestimated spread												
[PAR(γ·1.2)),GREG(γ·1.2)]/V_1	-2.6	-3.7	-0.6	-1.6	7.0	-9.1	-1.0	5.0	-3.9	1.9	-0.6	-7.2
[PAR(γ·1.2)),GREG(γ·1.2)]/V_2	-0.1	-2.1	1.9	-0.8	7.6	-8.2	-0.5	5.3	-3.5	2.2	-0.4	-7.0
[PAR(γ·1.5),GREG(γ·1.5)]/V_1	-0.3	-3.8	0.3	-1.8	4.7	-7.0	-2.5	3.9	-2.2	0.1	-1.5	-4.8
[PAR(γ·1.5),GREG(γ·1.5)]/V_2	2.7	-1.8	3.4	-0.7	5.6	-5.7	-1.8	4.5	-1.4	0.6	-1.1	-4.3
[PAR(γ·1.2)),πps]	-0.4	-0.5	2.7	2.6	6.9	-6.8	-5.7	0.3	-1.2	-2.1	-0.5	-5.4
[PAR(γ·1.5),πps]	0.1	-1.4	2.3	2.0	4.9	-6.6	-3.1	3.3	-0.8	-6.4	-1.0	-4.2
Underguestimated spread												
[PAR(γ·0.8),GREG(γ·0.8)]/V_1	-3.8	-7.2	-1.9	-4.0	5.8	-8.8	-1.1	4.5	-5.9	2.4	1.2	-7.2
[PAR(γ·0.8),GREG(γ·0.8)]/V_2	-1.2	-5.9	0.0	-3.2	6.4	-8.1	-0.7	4.8	-5.5	2.6	1.3	-7.1
[PAR(γ·0.5),GREG(γ·0.5)]/V_1	-6.8	-8.7	-3.7	-2.4	4.3	-7.7	-1.0	4.5	-6.9	4.8	0.0	-6.4
[PAR(γ·0.5),GREG(γ·0.5)]/V_2	-3.8	-7.2	-1.3	-1.4	5.1	-6.8	-0.6	4.8	-6.5	5.1	0.2	-6.2
[PAR(γ·0.8)),πps]	-0.6	-2.1	-0.3	2.9	4.7	0.0	-5.5	-1.7	0.0	-2.7	-3.1	0.1
[PAR(γ·0.5),πps]	0.0	-1.2	1.2	2.7	3.8	-1.4	-5.3	-2.0	-2.3	-3.2	-3.7	-2.8
					RESTD							
Correct spread guestimate												
[PAR(γ),GREG(γ)]	-6.0		-8.7	-5.3		-12.3	-4.5		-10.2	-3.3		-11.7
[SRS,GREG(γ)]	-8.0		-10.2	-6.4		-12.4	-5.0		-14.6	-2.8		-11.8
[PAR(1),πps]	-1.6		-7.5	-2.6		-9.9	-1.2		-7.5	-0.5		-6.3
[PAR(γ),πps]	3.1		-1.2	5.5		-5.2	-0.2		-3.0	0.8		-6.1
[SRS,HT]	3.4		4.2	3.2		1.4	-0.2		-0.1	1.3		1.4
Overguestimated spread												
[PAR(γ·1.2)),GREG(γ·1.2)]	-5.8		-9.2	-4.7		-12.6	-4.0		-9.9	-2.2		-11.0
[PAR(γ·1.5),GREG(γ·1.5)]	-5.0		-9.3	-4.6		-11.6	-4.1		-9.9	-2.1		-9.4
[PAR(γ·1.2)),πps]	3.7		-2.9	5.1		-7.2	0.6		-4.7	2.2		-6.8
[PAR(γ·1.5),πps]	4.8		-4.7	5.6		-9.3	2.8		-6.4	0.6		-8.1
Underguestimated spread												
[PAR(γ·0.8),GREG(γ·0.8)]	-6.0		-9.4	-6.1		-12.5	-4.6		-11.2	-2.7		-11.6
[PAR(γ·0.5),GREG(γ·0.5)]	-7.0		-9.5	-5.3		-12.0	-4.8		-11.9	-2.1		-11.6
[PAR(γ·0.8)),πps]	2.7		0.0	4.6		-3.3	-0.1		-1.8	1.2		-3.7
[PAR(γ·0.5),πps]	2.4		1.1	3.9		-0.3	-0.4		-0.9	0.7		-1.3

Table A.14. RBVE and RESTD in % for test situations of Type B. See (6.8), (6.9) and Table 6.1.												
Strategy and	n=10			n=25			n=50			n=80		
variance estimator	c=.06	c=.12	c=.25	c=.06	c=.12	c=.25	c=.06	c=.12	c=.25	c=.06	c=.12	c=.25
Correct spread guestimate												
[PAR(γ),GREG(γ)]/V_1	0.2	-3.1	-1.3	-1.4	5.2	-5.4	-0.3	3.6	-3.2	-0.9	2.3	-3.9
[PAR(γ),GREG(γ)]/V_2	0.2	-3.1	-1.3	-1.4	5.2	-5.4	-0.3	3.6	-3.2	-0.9	2.3	-3.9
[SRS,GREG(γ)]/V_1	-11.8	-15.5	.7.5	-5.7	2.8	-8.0	-1.3	1.0	-14.2	3.1	-0.6	-5.5
[SRS,GREG(γ)]/V_2	-10.8	-17.0	-7.1	-4.9	3.2	-7.1	-1.1	1.2	-13.9	3.4	-0.4	-5.4
[PAR(1),πps]	0.2	-3.1	-1.3	-1.4	5.2	-5.4	-0.3	3.6	-3.2	-0.3	3.6	-3.2
[PAR(γ),πps]	0.2	-3.1	-1.3	-1.4	5.2	-5.4	-0.3	3.6	-3.2	-0.3	3.6	-3.2
[SRS,HT]	2.8	2.3	4.5	2.0	1.5	-1.4	-4.4	-4.5	-5.3	-1.8	-1.9	-1.5
Overguestimated spread												
[PAR(γ·1.2)),GREG(γ·1.2)]/V_1	-3.4	-4.4	-1.9	-3.0	3.2	-4.7	-3.0	2.8	-3.2	-1.2	2.1	-2.1
[PAR(γ·1.2)),GREG(γ·1.2)]/V_2	-2.8	-3.7	-1.4	-2.8	3.5	-4.7	-3.0	2.9	-3.3	-1.2	2.1	-2.3
[PAR(γ·1.5),GREG(γ·1.5)]/V_1	-7.4	-9.3	-8.9	-6.4	0.0	-7.7	-7.0	0.2	-5.3	-3.9	-1.1	-4.4
[PAR(γ·1.5),GREG(γ·1.5)]/V_2	-5.6	-7.0	-7.0	-5.7	1.0	-7.2	-6.4	0.6	-5.4	-3.4	-1.1	-4.7
[PAR(γ·1.2)),πps]	2.0	-1.3	-1.4	1.6	5.0	-2.1	-3.4	0.0	-4.2	3.4	2.3	-1.5
[PAR(γ·1.5),πps]	2.5	1.5	-3.2	1.5	3.5	-0.4	-3.3	-2.5	-2.8	0.7	0.1	-1.6
Underguestimated spread												
[PAR(γ·0.8),GREG(γ·0.8)]/V_1	-1.6	-3.0	0.9	-0.9	6.7	-6.8	-1.0	3.1	-3.8	-1.7	-2.1	-5.8
[PAR(γ·0.8),GREG(γ·0.8)]/V_2	-1.6	-3.4	0.8	-0.9	6.5	-6.9	-1.0	3.1	-3.7	-1.7	-2.1	-5.8
[PAR(γ·0.5),GREG(γ·0.5)]/V_1	-2.5	-5.5	0.9	-3.3	5.5	-7.8	0.0	4.7	-7.1	1.1	1.8	-8.4
[PAR(γ·0.5),GREG(γ·0.5)]/V_2	-2.2	-6.0	1.0	-3.2	5.4	-7.7	0.0	4.6	-7.1	1.2	1.8	-8.4
[PAR(γ·0.8)),πps]	-0.1	-2.4	-0.1	0.6	4.8	0.0	-2.2	2.7	0.0	-5.6	-2.2	0.0
[PAR(γ·0.5),πps]	-0.1	-1.7	4.0	4.0	6.5	-4.4	-6.7	-3.3	-2.8	-4.0	-3.9	-6.8
					RESTD							
Correct spread guestimate												
[PAR(γ),GREG(γ)]	-4.7		-9.7	-5.5		-11.6	-5.0		-10.6	-5.7		-11.0
[SRS,GREG(γ)]	-6.1		-9.1	-5.5		-11.3	-4.0		-15.1	-2.2		-11.2
[PAR(1),πps]	-4.7		-9.7	-5.5		-11.8	-5.0		-10.6	-5.7		-11.0
[PAR(γ),πps]	-4.7		-9.7	-5.5		-11.8	-5.0		-10.6	-5.7		-11.0
[SRS,HT]	3.0		5.9	2.7		3.1	-0.6		0.9	0.8		2.9
Overguestimated spread												
[PAR(γ·1.2)),GREG(γ·1.2)]	-4.9		-8.7	-5.6		-10.4	-5.6		-9.8	-5.1		-9.2
[PAR(γ·1.5),GREG(γ·1.5)]	-2.9		-8.0	-4.3		-8.9	-5.0		-7.9	-3.3		-6.9
[PAR(γ·1.2)),πps]	-4.0		-9.7	-4.7		-9.8	-6.9		-10.5	-2.6		-8.2
[PAR(γ·1.5),πps]	-0.1		-5.7	-0.4		-3.8	-2.8		-3.9	-1.0		-2.5
Underguestimated spread												
[PAR(γ·0.8),GREG(γ·0.8)]	-5.6		-9.1	-5.4		-12.5	-5.4		-11.1	-6.2		-12.2
[PAR(γ·0.5),GREG(γ·0.5)]	-5.3		-8.5	-6.1		-12.6	-4.7		-12.6	-4.5		-13.4
[PAR(γ·0.8)),πps]	4.7		-3.7	5.2		-8.5	3.4		-6.1	1.1		-8.1
[PAR(γ·0.5),πps]	2.7		2.4	5.0		-1.5	-0.7		-0.8	0.5		-2.9

Table A.15. RBVE and RESTD in % for test situations of Type C. See (6.8), (6.9) and Table 6.1.												
Strategy and	**n=10**			**n=25**			**n=50**			**n=80**		
variance estimator	**c=3**	**c=7**	**c=13**	**c=3**	**c=7**	**c=13**	**c=3**	**c=7**	**c=13**	**c=3**	**c=7**	**c=13**
Correct spread guestimate												
[PAR(γ), GREG(γ)] / V_1	-5.9	-8.3	-4.1	-1.9	4.4	-6.7	-2.0	4.2	-4.7	5.4	0.9	-6.3
[PAR(γ), GREG(γ)] / V_2	-1.9	-5.7	-0.5	-0.6	5.6	-5.4	-1.5	4.7	-4.2	5.8	1.2	-6.0
[SRS, GREG(γ)] / V_1	-10.6	-14.3	-7.4	-5.3	2.8	-8.0	-1.5	2.7	-10.1	4.7	1.7	-6.8
[SRS, GREG(γ)] / V_2	-6.2	-11.9	-3.5	-3.7	4.1	-6.4	-0.8	3.2	-9.5	5.2	2.1	-6.5
[PAR(1), πps]	1.9	-1.1	-2.1	-3.4	0.1	-5.7	1.0	2.0	-2.2	-1.2	2.3	0.5
[PAR(γ), πps]	0.4	-1.0	0.7	2.6	3.2	-1.1	-5.9	-1.5	-1.3	-3.3	-3.6	-3.8
[SRS, HT]	3.2	2.0	4.0	1.9	0.8	-2.1	-4.8	-4.3	-4.0	-1.9	-1.9	-2.9
Overguestimated spread												
[PAR(γ· 1.2)), GREG(γ· 1.2)] / V_1	-4.8	-7.9	-3.8	-3.8	5.1	-7.1	-2.1	4.9	-5.2	4.6	2.3	-5.8
[PAR(γ· 1.2)), GREG(γ· 1.2)] / V_2	-0.9	-5.3	-0.3	-2.4	6.2	-5.8	-1.5	5.4	-4.6	5.0	2.6	-5.5
[PAR(γ· 1.5), GREG(γ· 1.5)] / V_1	-3.8	-7.3	-2.4	-4.2	5.8	-7.7	-2.7	4.5	-3.8	2.7	0.2	-6.8
[PAR(γ· 1.5), GREG(γ· 1.5)] / V_2	0.2	-4.5	1.2	-3.0	6.8	-6.4	-2.2	4.9	-3.3	3.0	0.4	-6.5
[PAR(γ· 1.2)), πps]	0.8	-0.6	1.2	2.3	3.9	-2.2	-7.9	-2.5	-2.7	-2.7	-4.3	-3.3
[PAR(γ· 1.5), πps]	-0.2	-1.3	1.7	2.6	4.1	-3.6	-7.1	-1.3	-0.7	-2.8	-3.6	-4.2
Underguestimated spread												
[PAR(γ· 0.8), GREG(γ· 0.8)] / V_1	-5.7	-9.5	-4.7	-3.6	3.8	-7.1	-1.9	3.7	-5.5	4.8	1.8	-5.9
[PAR(γ· 0.8), GREG(γ· 0.8)] / V_2	-1.7	-7.0	-1.2	-2.3	4.8	-5.8	-1.3	4.2	-5.0	5.2	2.1	-5.6
[PAR(γ· 0.5), GREG(γ· 0.5)] / V_1	-8.2	-10.5	-6.5	-3.3	3.7	-7.2	-1.8	4.4	-6.1	3.6	3.1	-5.8
[PAR(γ· 0.5), GREG(γ· 0.5)] / V_2	-4.0	-8.0	-2.9	-1.9	4.8	-5.9	-1.2	4.9	-5.5	4.0	3.4	-5.5
[PAR(γ· 0.8)), πps]	-0.6	-1.3	-0.2	0.7	1.3	0.0	-4.8	-1.1	0.0	-3.4	-3.5	0.0
[PAR(γ· 0.5), πps]	1.7	0.9	1.6	1.4	1.7	-1.5	-4.6	-1.5	-1.5	-2.5	-3.2	-3.3
					RESTD							
Correct spread guestimate												
[PAR(γ), GREG(γ)]	-6.1		-9.3	-3.9		-10.7	-3.8		-9.9	-0.1		-10.5
[SRS, GREG(γ)]	-7.5		-10.2	-6.0		-11.5	-4.5		-13.0	-1.7		-11.5
[PAR(1), πps]	-0.2		-6.5	0.2		-6.1	2.0		-7.1	2.3		-4.8
[PAR(γ), πps]	2.4		-1.1	3.6		-2.2	-0.8		-2.5	0.4		-3.9
[SRS, HT]	3.3		2.7	2.8		-0.2	-0.6		-1.3	0.9		-0.8
Overguestimated spread												
[PAR(γ· 1.2)), GREG(γ· 1.2)]	-5.5		-9.3	-4.6		-10.8	-3.7		-10.0	-0.1		-10.1
[PAR(γ· 1.5), GREG(γ· 1.5)]	-5.1		-9.1	-4.8		-11.1	-3.6		-9.1	-0.5		-10.2
[PAR(γ· 1.2)), πps]	2.7		-1.3	3.6		-3.2	-1.8		-3.7	0.8		-4.3
[PAR(γ· 1.5), πps]	2.6		-2.1	4.1		-4.7	-1.1		-3.6	0.8		-5.6
Underguestimated spread												
[PAR(γ· 0.8), GREG(γ· 0.8)]	-5.9		-9.5	-4.8		-10.9	-4.0		-10.4	-0.7		-10.5
[PAR(γ· 0.5), GREG(γ· 0.5)]	-6.8		-10.2	-4.9		-11.1	-4.2		-10.9	-1.6		-10.7
[PAR(γ· 0.8)), πps]	2.0		-0.5	2.6		-2.2	-0.3		-1.5	0.2		-3.3
[PAR(γ· 0.5), πps]	2.8		0.6	2.8		-1.1	-0.3		-1.2	0.7		-2.2

Table A.16. RBVE and RESTD in % for test situations of Type D. See (6.8), (6.9) and Table 6.1.												
Strategy and	**n=10**			**n=25**			**n=50**			**n=80**		
variance estimator	**c=3**	**c=7**	**c=13**	**c=3**	**c=7**	**c=13**	**c=3**	**c=7**	**c=13**	**c=3**	**c=7**	**c=13**
Correct spread guestimate												
[PAR(γ),GREG(γ)]/V_1	-4.1	-5.9	0.0	0.3	5.4	-8.0	-5.0	4.7	-4.6	-2.4	-0.5	-8.7
[PAR(γ),GREG(γ)]/V_2	-0.7	-4.1	2.4	1.5	6.2	-7.2	-4.5	5.0	-4.2	-2.0	-0.2	-8.5
[SRS,GREG(γ)]/V_1	-15.3	-16.2	-8.0	-7.5	0.9	-8.5	-3.2	2.7	-12.9	-1.9	0.0	-7.0
[SRS,GREG(γ)]/V_2	-11.5	-14.7	-5.1	-5.6	2.1	-7.1	-2.4	3.2	-12.2	-1.5	0.3	-6.8
[PAR(1),πps]	-2.8	-3.4	-1.1	-1.2	1.9	-4.8	-4.7	5.2	-2.2	-5.0	3.2	-1.3
[PAR(γ),πps]	0.4	-1.1	3.0	4.0	5.4	-4.6	-7.6	-3.3	-2.6	-3.3	-3.2	-6.6
[SRS,HT]	2.7	2.1	3.9	2.2	1.3	-1.7	-4.8	-4.6	-5.2	-1.8	-1.7	-2.5
Overguestimated spread												
[PAR(γ·1.2)),GREG(γ·1.2)]/V_1	-3.4	-3.2	-0.5	-0.5	6.9	-8.4	-4.8	6.1	-4.3	-0.8	-0.5	-7.6
[PAR(γ·1.2)),GREG(γ·1.2)]/V_2	0.0	-1.3	2.1	0.8	7.8	-7.5	-4.1	6.5	-3.8	-0.5	-0.3	-7.3
[PAR(γ·1.5),GREG(γ·1.5)]/V_1	0.2	-2.5	1.0	0.8	5.1	-5.7	-3.1	5.9	-2.6	-4.9	-1.2	-4.7
[PAR(γ·1.5),GREG(γ·1.5)]/V_2	3.7	-0.2	4.1	2.1	6.1	-4.4	-2.3	6.5	-1.8	-4.4	-0.8	-4.3
[PAR(γ·1.2)),πps]	0.4	0.4	2.4	2.2	5.3	-5.6	-7.1	-1.4	-2.9	-1.2	0.0	-5.3
[PAR(γ·1.5),πps]	1.5	0.5	2.8	2.4	4.3	-5.0	-4.6	1.3	-2.3	-5.2	-1.4	-4.2
Underguestimated spread												
[PAR(γ·0.8),GREG(γ·0.8)]/V_1	-7.4	-8.0	-2.5	-2.2	4.7	-7.9	-3.9	5.6	-6.1	-1.4	-0.5	-7.6
[PAR(γ·0.8),GREG(γ·0.8)]/V_2	-4.0	-6.3	-0.3	-1.0	5.4	-7.1	-3.4	5.9	-5.8	-1.1	-0.3	-7.4
[PAR(γ·0.5),GREG(γ·0.5)]/V_1	-9.6	-9.5	-4.6	-4.1	3.0	-6.7	-3.9	4.9	-8.1	0.6	-1.9	-6.5
[PAR(γ·0.5),GREG(γ·0.5)]/V_2	-5.8	-7.8	-2.2	-2.7	3.9	-5.8	-3.3	5.2	-7.7	0.9	-1.6	-6.3
[PAR(γ·0.8)),πps]	0.1	-1.1	-0.4	2.6	3.8	0.1	-6.1	-2.5	0.0	-2.4	-2.6	0.1
[PAR(γ·0.5),πps]	0.7	-0.4	1.4	2.6	3.3	-0.8	-4.9	-2.2	-3.5	-3.1	-3.7	-3.4

Table A.17. RBVE and RESTD in % for test situations of Type E. See (6.8), (6.9) and Table 6.1.												
Strategy and variance estimator	n=10			n=25			n=50			n=80		
	c=3	c=7	c=13	c=3	c=7	c=13	c=3	c=7	c=13	c=3	c=7	c=13
Correct spread guestimate												
$[PAR(\gamma), GREG(\gamma)]/V_1$	-3.4	-5.6	-0.8	0.8	3.9	-7.6	-6.5	0.8	-1.9	-1.6	-1.8	-8.4
$[PAR(\gamma), GREG(\gamma)]/V_2$	0.3	-3.1	1.8	2.2	5.0	-6.7	-5.9	1.2	-1.5	-1.2	-1.5	-8.1
$[SRS, GREG(\gamma)]/V_1$	-19.0	-18.0	-9.7	-7.5	-2.1	-8.3	-6.0	-2.0	-10.3	-3.0	-1.0	-7.3
$[SRS, GREG(\gamma)]/V_2$	-14.4	-15.2	-6.5	-4.9	-0.4	-6.7	-4.8	-1.2	-9.7	-2.4	-0.5	-7.0
$[PAR(1), \pi ps]$	-2.7	-2.7	-0.8	0.9	2.0	-3.6	-5.2	3.6	-0.7	-3.1	3.3	-1.1
$[PAR(\gamma), \pi ps]$	0.9	-0.8	2.5	3.4	4.6	-4.2	-7.3	-3.6	-1.5	-2.4	-2.9	-6.2
$[SRS, HT]$	2.1	1.4	3.1	2.1	1.1	-1.7	-4.9	-4.8	-4.8	-1.1	-1.0	-2.3
Overguestimated spread												
$[PAR(\gamma \cdot 1.2)), GREG(\gamma \cdot 1.2)]/V_1$	-1.9	-2.2	-0.8	-0.2	4.8	-7.6	-6.5	1.3	-2.2	0.0	0.1	-7.2
$[PAR(\gamma \cdot 1.2)), GREG(\gamma \cdot 1.2)]/V_2$	1.6	0.3	1.9	1.1	5.9	-6.7	-5.9	1.8	-1.7	0.4	0.4	-6.9
$[PAR(\gamma \cdot 1.5), GREG(\gamma \cdot 1.5)]/V_1$	1.7	-0.3	1.2	2.0	4.5	-4.7	-5.1	2.6	-1.1	-4.1	-1.0	-4.7
$[PAR(\gamma \cdot 1.5), GREG(\gamma \cdot 1.5)]/V_2$	4.7	2.2	4.3	3.2	5.5	-3.5	-4.5	3.2	-0.4	-3.7	-0.6	-4.2
$[PAR(\gamma \cdot 1.2)), \pi ps]$	0.9	0.6	1.9	1.6	4.1	-5.1	-7.1	-2.7	-2.0	-0.5	0.4	-4.9
$[PAR(\gamma \cdot 1.5), \pi ps]$	1.7	1.2	2.5	2.4	3.8	-4.1	-5.6	-0.3	-1.5	-4.2	-1.3	-4.1
Underguestimated spread												
$[PAR(\gamma \cdot 0.8), GREG(\gamma \cdot 0.8)]/V_1$	-7.0	-7.8	-3.3	-1.4	2.6	-7.1	-5.5	2.2	-3.3	-1.2	-1.5	-7.0
$[PAR(\gamma \cdot 0.8), GREG(\gamma \cdot 0.8)]/V_2$	-3.1	-5.4	-0.9	0.0	3.7	-6.3	-4.8	2.7	-2.9	-0.9	-1.2	-6.9
$[PAR(\gamma \cdot 0.5), GREG(\gamma \cdot 0.5)]/V_1$	-10.0	-9.7	-5.9	-3.9	0.8	-6.3	-6.2	1.7	-5.6	-0.9	-3.3	-6.0
$[PAR(\gamma \cdot 0.5), GREG(\gamma \cdot 0.5)]/V_2$	-5.4	-7.0	-3.2	-2.0	2.1	-5.3	-5.5	2.2	-5.1	-0.5	-3.0	-5.8
$[PAR(\gamma \cdot 0.8)), \pi ps]$	0.5	-0.8	-0.4	2.0	3.0	0.1	-5.9	-2.5	0.0	-1.7	-2.1	0.1
$[PAR(\gamma \cdot 0.5), \pi ps]$	0.8	-0.3	1.0	2.2	2.7	-1.0	-5.8	-2.8	-2.7	-2.7	-3.4	-3.2

Table A.18. RBVE and RESTD in % for test situations of Type F. See (6.8), (6.9) and Table 6.1.												
Strategy and	n=10			n=25			n=50			n=80		
variance estimator	c=3	c=7	c=13	c=3	c=7	c=13	c=3	c=7	c=13	c=3	c=7	c=13
Correct spread guestimate												
[PAR(γ), GREG(γ)]/V_1	-0.7	-3.0	-1.2	-0.6	3.4	-4.4	-5.4	-2.8	-8.5	-0.3	1.9	-6.2
[PAR(γ), GREG(γ)]/V_2	6.3	0.9	2.1	1.8	4.8	-3.2	-4.3	-2.2	-8.0	0.4	2.3	-5.8
[SRS, GREG(γ)]/V_1	-17.7	-15.4	-10.1	-3.4	2.5	-7.0	-6.3	0.1	-13.1	5.3	1.7	-5.8
[SRS, GREG(γ)]/V_2	-10.2	-12.5	-6.6	-0.3	4.3	-5.3	-5.0	0.8	-12.4	6.0	2.1	-5.5
[PAR(1), πps]	1.8	-0.4	-3.5	2.8	3.3	-3.5	-2.7	-3.2	-4.3	2.1	-0.4	-1.5
[PAR(γ), πps]	-1.1	-4.2	3.1	4.0	7.6	-7.3	-6.5	2.2	-0.9	-4.6	-1.6	-7.9
[SRS, HT]	3.8	2.5	5.6	2.3	1.7	-2.3	-4.5	-4.0	-4.6	-2.9	-2.4	-3.1
Overguestimated spread												
[PAR(γ·1.2)), GREG(γ·1.2)]/V_1	2.2	-1.7	-1.4	0.2	3.7	-5.3	-5.5	-2.9	-8.0	1.7	-0.6	-5.5
[PAR(γ·1.2)), GREG(γ·1.2)]/V_2	9.8	2.6	2.1	2.9	5.4	-3.8	-4.2	-2.0	-7.2	2.5	-0.1	-5.0
[PAR(γ·1.5), GREG(γ·1.5)]/V_1	8.1	1.1	0.1	0.4	2.7	-3.4	-6.9	-4.5	-6.3	-0.2	-3.8	-5.1
[PAR(γ·1.5), GREG(γ·1.5)]/V_2	17.0	6.8	5.4	4.0	5.3	-1.0	-5.1	-3.2	-4.9	1.1	-2.8	-4.2
[PAR(γ·1.2)), πps]	-2.8	-3.1	1.2	0.3	8.4	-8.3	-3.6	4.8	-2.1	-2.1	-0.2	-7.2
[PAR(γ·1.5), πps]	1.6	-4.2	-0.9	-2.6	4.5	-6.1	-5.1	1.6	-3.0	2.2	-2.3	-5.3
Underguestimated spread												
[PAR(γ·0.8), GREG(γ·0.8)]/V_1	-1.5	-4.9	-3.0	-1.7	3.8	-6.1	-3.4	-1.5	-9.3	1.2	1.9	-6.3
[PAR(γ·0.8), GREG(γ·0.8)]/V_2	5.3	-1.4	-0.1	0.6	5.1	-5.0	-2.4	-0.9	-8.8	1.7	2.2	-6.1
[PAR(γ·0.5), GREG(γ·0.5)]/V_1	-6.0	-8.0	-3.6	0.1	3.6	-5.6	-4.4	-1.2	-8.9	1.1	1.3	-6.2
[PAR(γ·0.5), GREG(γ·0.5)]/V_2	1.2	-4.4	-0.4	2.6	5.1	-4.3	-3.4	-0.6	-8.4	1.7	1.6	-5.9
[PAR(γ·0.8)), πps]	-1.2	-4.1	-0.2	2.9	6.0	0.1	-6.4	1.1	0.0	-3.5	-2.1	0.0
[PAR(γ·0.5), πps]	-0.1	-2.2	2.2	3.0	4.7	-2.4	-5.6	0.1	-1.8	-3.8	-3.3	-3.9

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 1998 and onwards:

1998:1	Preliminär statistik: Nybyggnadskostnader - en simuleringstudie *(Catarina Elffors)*
1998:2	On Inclusion Probabilities for Order Sampling *(Bengt Rosén)*
1998:3	On the Stratification of Highly Skewed Populations *(Dan Hedlin)*
1998:4	Bortfallsbarometer nr 13 *(Per Nilsson, Antti Ahtiainen, Stefan Berg, Mats Bergdahl, Monica Rennermalm och Marcus Vingren)*
1998:5	Estimation from Order πps Samples with Non - Response *(Bengt Rosén and Pär Lundqvist)*
1998:6	On variance estimation for measures of change when samples are coordinated by a permanent random numbers technique *(Lennart Nordberg)*
1999:1	Täckningsproblem i Registret över totalbefolkning RTB. Skattning av övertäckning med en indirekt metod *(Jan Qvist)*
1999:2	Bortfallsbarometer nr 14 *(Per Nilsson, Antti Ahtiainen, Mats Bergdahl, Tomas Garås, Jan Qvist och Charlotte Strömstedt)*
1999:3	Att mäta statistikens kvalitet *(Claes Andersson, Håkan L. Lindström och Thomas Polfeldt)*
2000:1	Kalibrering av vikter – beskrivning av tekniken och de SCB-fall den prövats i *(Sixten Lundström et al)*
2000:2	On Inclusion Probabilities and Estimator Bias for Pareto πps Sampling *(Nibia Aires and Bengt Rosén)*
2000:3	Bortfallsbarometer nr 15 *(Per Nilsson, Ann-Louise Engstrand, Sara Tångdahl, Stefan Berg, Tomas Garås och Arne Holmqvist)*
2000:4	Bortfallsanalys av SCB-undersökningarna HINK och ULF *(Jan Qvist)*
2000:5	Generalized Regression Estimation and Pareto πps *(Bengt Rosén)*

ISSN 0283-8680

Tidigare utgivna **R&D Reports** kan beställas genom Katarina Klingberg, SCB, MET, Box 24 300, 104 51 STOCKHOLM (telefon 08-506 942 82, fax 08-506 945 99, e-post katarina.klingberg@scb.se).
R&D Reports from 1988-1997 can - in case they are still in stock - be ordered from Statistics Sweden, attn. Katarina Klingberg, MET, Box 24 300, SE-104 51 STOCKHOLM (telephone +46 8 506 942 82, fax +46 8 506 945 99, e-mail katarina.klingberg@scb.se).

SCB-Tryck, Örebro 2000.11