MODIFICATIONS OF THE STRINGER-BOUND:

A SIMULATION STUDY ON THE PERFORMANCE OF AUDIT SAMPLING EVALUATION METHODS

by

Antoine Lucassen *

Hans Moors **

Paul van Batenburg *

Abstract

This paper presents three modifications of the Stringer bound in audit sampling evaluations. The behavior of the Stringer bound and of the suggested modifications is studied by means of simulation.
Two of the modifications proposed seem to yield better upper bounds for the misstatement percentage in a population: they satisfy the nominal confidence level and result in an upper bound below the Stringer bound. The modification Amount Tainting Order (ATO) is preferred by the authors.

---

* Deloitte & Touche
Computer Assurance Services Center
P.O. Box 58140
1040 HC Amsterdam, The Netherlands

** Tilburg University
Department of Econometrics
P.O. Box 90153
5000 LE Tilburg, The Netherlands

1. Introduction

Nowadays, monetary unit sampling is a widely used technique in auditing accounts (van Batenburg and Kriens (1991), Tamura (1989)). Sampling makes it possible to obtain an upper bound for the total misstatement percentage in an accounting population. One of these upper bounds is the Stringer bound, which takes into account the relative misstatement in sample items, expressed as "taintings" (the difference between the book value and the audited value as a percentage of the book value). These taintings are ranked according to their sizes in a descending order. In previous simulation studies the actual confidence level achieved by the Stringer bound exceeded the nominal confidence level, indicating that the Stringer bound might be conservative (Burdick and Reneau (1978), Leitch et al (1982), Plante et al (1985), Reneau (1987)).

The purpose of this paper is to present the results of a simulation study which was designed to:

- confirm that the Stringer bound is conservative;
- examine several modifications of the Stringer bound based on different rankings of the taintings.

In section 2 the term conservatism is defined. Section 3 describes the Stringer bound and several modifications. The study design is described in section 4. Section 5 summarizes the results from the simulation study. Finally, in section 6 the conclusions are drawn.

2. Conservatism

The Stringer bound is said to be conservative. Theoretically this property can be described as: " the probability, by which the calculated upper bound exceeds the actual misstatement percentage in an accounting population, exceeds the nominal confidence level for all (practical and theoretical) populations".

To verify this statement one has to find the population which minimizes the probability described above and to compare this minimal probability with the nominal confidence level. Formally this can be described as follows.

Define:	D	=	set of all possible accounting populations
	d	=	element (= population) of set D
	$(1-\alpha)$	=	nominal confidence level
	SB_d	=	Stringer bound (in %) for the total misstatement percentage in a population d, based on a sample of size n, containing r errors with taintings t_i and a nominal confidence level $(1-\alpha)$
	ψ_d	=	total misstatement percentage in population d

To find the population d^* which minimizes the probability described above one has to find:

$$\min_{d \in D} \; P(\underline{SB}_d > \psi_d)$$

where underlining denotes random variables.

The minimum $(1-\alpha^*) = P(\underline{SB}_{d*} > \psi_{d*})$ is the actual confidence level. The following cases can occur:

- if $(1-\alpha^*) > (1-\alpha)$: the Stringer bound is conservative;
- if $(1-\alpha^*) = (1-\alpha)$: the Stringer bound is exact;
- if $(1-\alpha^*) < (1-\alpha)$: the Stringer bound does not satisfy the nominal confidence level.

To be able to conclude one of the above by a simulation study, all possible populations have to be screened. This is not possible. In this article only some populations are considered.

3. Stringer bound and modifications of the Stringer bound

The Stringer bound (introduced by K.W. Stringer in 1963 at the annual meeting of the American Statistical Association - see Stringer (1963) -and elaborated by a.o. Anderson and Teitlebaum (1973) and Leslie, Teitlebaum and Anderson (1980)) is a linear combination of the taintings in which the coefficients decrease with decreasing tainting. The Stringer bound is defined as:

$$SB = p_0(\alpha) + \sum_{i=1}^{r} t_{(i)}(p_i(\alpha) - p_{i-1}(\alpha))$$

with

SB = Stringer bound
$p_i(\alpha)$ = upper bound for the misstatement percentage in the accounting population if j errors were found in the sample of size n, j = 0,...,n[1]
r = number of misstatements found in the sample
$t_{(i)}$ = tainting of i-th misstatement with $1 \geq t_{(1)} \geq t_{(2)} \geq \ldots \geq t_{(r)} > 0$

Note that the incremental factors $p_i(\alpha) - p_{i-1}(\alpha)$ decline as i increases for $\alpha < 0.28$ and that the taintings are ranked in a declining order. Defining T as the set of all possible ways to rank the taintings in the sample, the Stringer bound therefore is the most conservative upper bound in T.

The Stringer bound is generally considered to be conservative; e.g. Neter et. al. (1978) state: "All known simulation studies, however, do indicate that the method is conservative". Therefore, three modifications are presented here and their degree of conservatism is investigated. All of them are based on different rankings of the taintings:

[1] $p_i(\alpha)$ is that p which solves $P[\, \underline{r} \leq i \mid n,p\,] = \alpha$ for $i \leq r$

ITO (Increased Tainting Order): In set T ranking the taintings in an increasing order should result in the least conservative way of calculating an upper bound because the smallest taintings are connected with the highest incremental factors. So, an obvious modification of the Stringer bound is obtained by ranking the taintings in an increasing order:

$$ITO = p_0(\alpha) + \sum_{i=1}^{r} t_{(r-i+1)}(p_i(\alpha) - p_{i-1}(\alpha))$$

RTO (Random Tainting Order): The following variant is chosen here to examine whether ranking of the taintings is really necessary. Why not use the order in which the misstatements (e.g. taintings) were discovered?
Setting the taintings in a random order results in the upper bound:

$$RTO = p_0(\alpha) + \sum_{i=1}^{r} t_i(p_i(\alpha) - p_{i-1}(\alpha))$$

ATO (Amount Tainting Order): The following example explains why this variant is examined: what is more important to the auditor, a 1% misstatement in an item of 10000 monetary units or a 100% misstatement in an item of one monetary unit ? Probably the auditor will consider the former more important.
Therefore, we propose as third variant: rank the taintings according to descending misstatement amounts.
Define: y_i = book value of item i
x_i = audited value of item i
Rank the misstatement amounts $f_i = y_i - x_i$ ($y_i > x_i$ so $f_i > 0$) in a descending order. This results in:

$$f_{(1)} \geq f_{(2)} \geq \ldots \geq f_{(r)} > 0.$$

Calculating taintings $t^*_{(i)}$ by dividing $f_{(i)}$ by the corresponding book value yields the upper bound:

$$ATO = p_0(\alpha) + \sum_{i=1}^{r} t^*_{(i)}(p_i(\alpha) - p_{i-1}(\alpha))$$

These variants plus the Stringer bound itself will be examined in the simulation study described in the following sections. Logically, the relations

$$SB \geq RTO \geq ITO$$

and

$$SB \geq ATO$$

hold for all samples.

4. Study design

The simulation study is based on the populations of the AICPA-study of Neter and Loebbecke (1975). These populations are said to be representative for the accounting populations founded in practice. This simulation study only uses the populations 3 and 4 of the AICPA-study. The other two populations of this study (1 and 2) were not used because they contain understatements (book value < audited value) which can not be evaluated in a straightforward manner using monetary unit sampling. In total sixteen different study populations were used:

- ten study populations which already had been constructed by Neter and Loebbecke. In both their populations 3 and 4 five different error percentages measured in items were used. Neter and Loebbecke obtained the corresponding audited values using predefined error pool tables;

- the remaining six study populations were constructed from the populations 3 and 4 using a uniform tainting distribution. First it was determined whether a specific item would contain a misstatement (using probabilities of 5%, 10% and 30% respectively) and if so, a misstatement percentage was generated using a uniform probability distribution U(0,1), resulting in a corresponding audited value. Else the audited value equalled the book value.

This simulation study uses three different sample sizes that are often used in practice, namely n=60, n=150 and n=300. If no misstatements are found in the sample, they lead to upper bounds $p_0(0.05)$ of 5%, 2% and 1%, respectively, using a Poisson approximation and a confidence level of 95%. Hence, upper bounds for populations with ψ smaller than these minimal levels have a confidence level equal to 1. This applies for a number of study populations for several sample sizes; some of them have been omitted in the following tables.

Table 1 shows some characteristics of the study populations used including the populations 3 and 4 of the AICPA-study (TBV = total book value (relating to the original population), TAV = total audited value (relating to the study populations), ψ = misstatement percentage; μ = mean; σ = standard deviation ; ε = skewness ; ϕ = kurtosis).
Each study population has a 2 or 3 character code: the first character representing the original population of the AICPA-study, the next representing a code for the error percentage measured in items (C=5%, D=10%, E=30%) and a final character (U) if the study population was constructed using a uniform tainting distribution.

TABLE 1: **CHARACTERISTICS OF (STUDY)POPULATIONS**

POPULATION	**TBV/TAV (in $)**	**ψ (in %)**	μ	σ	ε	ϕ
3	13671500	-	1946	7022	7.9	78.1
3C U	13370471	2.20	1903	6885	7.9	78.1
3D U	13024516	4.73	1854	6797	8.1	80.8
3E	13509839	1.18	1993	7023	7.9	78.0
3E U	11385326	16.72	1620	6036	8.5	91.8
4	7502957	-	1860	3865	3.2	11.4
4C	7402350	1.34	1835	3855	3.3	11.5
4C U	7260465	3.23	1800	3766	3.3	11.8
4D	7237279	3.54	1795	3814	3.3	11.7
4D U	7146865	4.75	1772	3744	3.3	12.0
4E	6442371	14.14	1597	3608	3.5	13.3
4E U	6526820	13.01	1618	3576	3.5	13.6

For each study population and sample size 1500 replications were made. For each simulation run several measures were calculated:

- coverage: the proportion of the upper bounds exceeding the total misstatement percentage in the population;
- mean of the upper bounds found;
- standard deviation of the upper bounds found.

5. Results

In table 2 the results are shown of the coverages of the Stringer bound and of the modifications proposed in section 3.

The first conclusion from table 2 is that once again the Stringer bound appears to be (very) conservative: all coverages found exceed the nominal confidence level of 95% (here by even more than 2 percentage points).
The second conclusion from table 2 is the rejection of ITO as a correct method for calculating an upper bound for the misstatement percentage in a population. Some coverages found do not fulfill the necessary requirement that coverages should exceed the nominal confidence level.
The second modification, RTO, does not fulfill the necessary requirement in two cases (population 4E, n=150 and n=300). However, even in these two cases the average coverage in this simulation study of 1500 replications is not significantly different (α=0.05 one-sided Student-test) from the required value of 0.95. Therefore RTO can not be rejected as a valid way of calculating an upper bound for the misstatement percentage in a population.
Finally, it can be concluded that ATO satisfies the nominal confidence level.

TABLE 2: **COVERAGES OF STRINGER BOUND AND MODIFICATIONS**

POPULATION	n	SB	ITO	RTO	ATO
3C U	150	100.00	100.00	100.00	100.00
3C U	300	99.60	99.27	99.47	99.40
3D U	150	98.73	97.73	98.33	98.67
3D U	300	98.20	95.33	97.07	97.47
3E	150	100.00	100.00	100.00	100.00
3E	300	99.80	99.60	99.80	99.60
3E U	60	99.60	99.07	99.53	99.60
3E U	150	99.13	96.67	98.27	98.87
3E U	300	99.27	94.80	97.80	98.67
4C	150	100.00	100.00	100.00	100.00
4C	300	97.40	97.13	97.20	97.33
4C U	150	99.67	99.47	99.47	99.60
4C U	300	98.73	97.60	98.13	98.67
4D	150	98.80	95.60	96.53	98.27
4D	300	97.47	91.73	95.87	97.07
4D U	150	98.87	98.47	98.73	98.87
4D U	300	98.87	96.73	97.67	98.67
4E	60	98.27	96.40	97.07	98.00
4E	150	97.73	89.60	94.87	97.67
4E	300	97.07	87.47	94.60	96.80
4E U	60	99.40	98.93	99.20	99.33
4E U	150	99.20	96.47	98.33	98.93
4E U	300	98.73	94.50	97.53	98.20

One can not compare variants on their coverages alone since a coverage of 96% is not worse than one of 99%. In fact, a coverage of 96% would be preferable to one of 99% if it implied that the actual bounds were smaller in the first case than in the second. Hence, table 3 shows the mean (m) and the standard deviation (s) of the upper bounds found. Here the Stringer bound SB is compared with ATO and RTO.

Table 3 shows that the three upper bounds (on average) clearly overstate the true value of the misstatement percentage. The upper bounds improve with increasing sample size. This is logical since the more observations one has, the more accurate the estimation will tend to be.

TABLE 3: **MEAN (m) AND STANDARD DEVIATION (s) OF UPPER BOUNDS SB, RTO AND ATO vs ψ**

POPULATION	n	SB m	in % s	RTO m	in % s	ATO m	in % s	ψ in %
3C U	150	5.29	1.21	5.15	1.18	5.15	1.17	2.20
3C U	300	4.13	0.81	3.99	0.80	3.99	0.78	2.20
3D U	150	8.70	1.84	8.37	1.79	8.52	1.81	4.73
3D U	300	7.26	1.26	6.96	1.23	7.11	1.25	4.73
3E	150	3.88	1.09	3.81	1.07	3.69	1.02	1.18
3E	300	2.79	0.71	2.71	0.70	2.60	0.67	1.18
3E U	60	27.73	4.50	26.71	4.43	27.17	4.41	16.72
3E U	150	22.92	2.72	22.11	2.70	22.48	2.69	16.72
3E U	300	20.85	1.80	20.19	1.79	20.51	1.78	16.72
4C	150	4.16	1.35	4.07	1.32	4.10	1.33	1.34
4C	300	3.00	0.85	2.91	0.83	2.95	0.84	1.34
4C U	150	6.72	1.56	6.50	1.51	6.68	1.55	3.23
4C U	300	5.54	1.05	5.23	1.02	5.41	1.05	3.23
4D	150	7.27	1.96	6.96	1.93	7.18	1.96	3.54
4D	300	5.88	1.26	5.61	1.26	5.81	1.26	3.54
4D U	150	8.67	1.79	8.34	1.74	8.58	1.76	4.75
4D U	300	7.31	1.21	7.01	1.19	7.22	1.20	4.75
4E	60	25.13	5.60	23.94	5.60	24.99	5.61	14.14
4E	150	20.13	3.17	19.19	3.19	20.04	3.18	14.14
4E	300	18.22	2.18	17.49	2.20	18.15	2.18	14.14
4E U	60	23.11	4.23	22.28	4.14	22.69	4.14	14.14
4E U	150	18.65	2.47	17.95	2.43	18.29	2.43	13.01
4E U	300	16.70	1.72	16.13	1.70	16.41	1.70	13.01

6. Conclusions

Three modifications of the Stringer bound SB were presented: ITO uses the taintings in the sample in an increasing order and RTO in a random order, while ATO ranks the taintings according to the corresponding misstatement amount order. The following relations between these upper bounds exist:

$$SB \geq RTO \geq ITO, \; SB \geq ATO.$$

A simulation study showed that ITO did not meet the nominal confidence level uniformly; hence ITO is not a correct upper bound. In all (23) cases considered SB and ATO satisfied the nominal confidence level; in two cases the coverage of RTO was slightly (but not significantly) below this value.

The coverage of SB exceeded 97% throughout, confirming the general conviction that SB is conservative. The coverage of both ATO and RTO was uniformly below that of SB.
Our overall conclusion is that ATO and RTO both are serious competitors to SB. A disadvantage of RTO may be that the random tainting order might be manipulated. An advantage of ATO is that it probably results from a more realistic method of expressing the auditor's ideas about the severeness of the misstatements. For these reasons ATO is preferred; in fact, ATO is already recommended as an alternative to SB in the Deloitte Touche Tohmatsu International manual on Audit Sampling.

References

Anderson, R. and Teitlebaum, A.D. "Dollar Unit Sampling: a solution to the audit sampling dilemma" in Canadian Chartered Accountant, April 1973.

Batenburg, P.C. van and Kriens, J. "Applications of statistical methods and techniques to auditing and accounting" in Bulletin of the International Statistical Institute, Proceedings of the 48th session, vol 54, book 1, theme 6.

Burdick, R.K. and Reneau, J.H. "The impact of different error distributions on the performance of selected sampling estimators in accounting populations" in Proceedings of the Business and Economical Statistics Section, 1978, p. 779-781.

Neter, J., Leitch, R.A. and Fienberg, S.E. "Dollar Unit Sampling: multinomial bounds for total overstatement and understatement errors" in The Accounting Review, no 1, 1978, p. 77-93.

Neter, J. and Loebbecke, J.K. "Behavior of major statistical estimators in sampling accounting populations: an empirical study" in Auditing Research Monograph no 2, 1975.

Leitch, R.A., Neter, J., Plante, R. and Sinha, P. "Modified multinomial bounds for larger number of errors in audits" in The Accounting Review, no 2, April 1982, p. 384-400.

Leslie, D.A., Teitlebaum, A.D. and Anderson, R.J. "Dollar Unit Sampling: a practical guide for auditors", 1980.

Plante, R., Neter, J. and Leitch, R.A. "Comparative performance of multinomial, cell and Stringer bounds" in Auditing: a journal of practice and theory, 1985, p. 40-56.

Reneau, J.H. "C.A.V. bounds in dollar unit sampling: some simulation results" in The Accounting Review, no 3, 1987, p. 669-680.

Stringer, K.W. "Practical aspects of statistical sampling in auditing" in Proceedings of the Business and Economical Section, 1963, p. 405-411.

Tamura, H. "Statistical models and Analysis in auditing" in Statistical Science, vol 4, no 1, February 1989, p. 2-33.

Ontvangen: 6-3-1995
Geaccepteerd: 9-10-1995