Journal of Official Statistics, Vol. 13, No. 4, 1997, pp. 327–339

# A Sampling Scheme with Partial Replacement

*J.L. Sánchez-Crespo*

Unequal probability sampling *with* replacement is easier to handle, both theoretically and practically, than unequal probability sampling *without* replacement. However, sampling without replacement usually produces estimates that are more efficient. The sampling scheme due to Sánchez-Crespo and Gabeiras (1987), based on sampling with *partial replacement*, is a sort of compromise between sampling with and without replacement. It is an attempt to obtain a procedure that is more efficient than sampling with replacement while at the same time retaining some of the simplicity of sampling with replacement. In survey sampling practice the Sánchez-Crespo and Gabeiras scheme seems to provide a simple and potentially useful procedure that has always a smaller expected variance and *greater stability* for the variance estimator than unequal probability sampling with replacement. We have also found *strong evidence* of better stability for the variance estimator than some of the best procedures developed up to now for sampling without replacement and unequal probabilities.

*Key words:* Unequal probability sampling; variance estimation.

## 1. Introduction

The theory of unequal probability sampling *with* replacement was given by Hansen and Hurwitz (HH) (1943). Horvitz and Thompson (HT) (1952) developed a general theory for unequal probability sampling *without* replacement. Among the best sampling schemes for carrying out unequal probability sampling without replacement is the Brewer-Durbin (HT/B) method (Brewer 1963 and 1975; Durbin 1967).

Sánchez-Crespo and Gabeiras (SCG) (1987) presented a sampling scheme based on *partial* replacement. The objective was to develop a sampling scheme that would be easy to use and more efficient than the HH scheme, and would allow more stable variance estimates than existing without-replacement-schemes such as the Brewer-Durbin scheme. In the present article we demonstrate that the variance estimator for the SCG scheme is *always more stable* than the variance estimator for the HH scheme. Furthermore we have found *strong evidence* that the SCG variance estimator is more stable, than that of the Brewer-Durbin method. On the other hand the expected variance is usually greater with the SCG than with the Brewer-Durbin method. We have also found that for the 20 natural populations in the Rao and Bayless (1969) study the percentage loss in expected variance for SCG as compared to Brewer was on average equal to five per cent while the average gain in stability for the variance estimator of SCG over Brewer was at least 28 per cent.

**Acknowledgment:** The author wishes to express his gratitude to B.A. Bailar, R.K.W. Brewer, I.P. Fellegi, and J.N.K. Rao, for some very useful comments and suggestions. He also wishes to acknowledge the cooperation and help received from F. Azorín, J.de Parada, J. Porras, M. Herrador, and G.de Parada.

## 2. Selection Procedure

The basic context for the SCG sampling scheme is the selection of a sample of $n$ primary sampling units (PSUs) from a population of $N$ PSUs. Associated with the $i^{\text{th}}$ PSU is a characteristic $X_i$ and a size measure $M_i$, with

$$M = \sum_{i=1}^{N} M_i$$

We define the constant $b$ as

$$b = \left[\frac{\min(M_1, \ldots, M_N)}{n-1}\right] \quad (1)$$

where $[x]$ denotes the largest integer $\leq x$.

The selection procedure consists of $n$ selections of one PSU at a time. Each selection is made with probabilities proportional to size measures defined as follows. In the first selection we use the original size measures $M_1, \ldots, M_N$. Let us suppose that the $i^{\text{th}}$ PSU is selected. In the second selection, this PSU gets the reduced size measure $M_i - b$. In the following selections, each time a PSU is selected, its hitherto existing size measure is reduced by the subtraction of the number $b$. After all the $n$ selections, a sample of $n$ (not necessarily distinct) PSUs is obtained. Obviously, this is a with replacement sampling scheme. However, the probability of reselection is successively reduced each time a unit has been selected, which justifies the term *partial replacement*.

The following is an illustration, for the case $n = 2$, of the probability that the $i^{\text{th}}$ PSU will be included in the sample (that is, selected in at least one of the $n$ selections). Let the random variable $e_i$ denote the number of times that the $i^{\text{th}}$ PSU will appear in the sample, and let

$$P(e_i = t) = P(n, t); \quad t = 0, 1, \ldots, n \quad (2)$$

with

$$\sum_{t=0}^{n} P(n, t) = 1$$

In this example, with $n = 2$, these probabilities are

$$P(2,0) = \frac{M - M_i}{M} \cdot \frac{M - M_i - b}{M - b}$$

$$P(2,1) = 2 \cdot \frac{M_i}{M} \cdot \frac{M - M_i}{M - b}$$

$$P(2,2) = \frac{M_i}{M} \cdot \frac{M_i - b}{M - b}$$

The expected value of $e_i$ is

$$E(e_i) = \sum_{t=1}^{2} tP(2, t) = 2P_i$$

where $P_i = M_i/M$

The variance of $e_i$ and the covariance of $e_i$ and $e_j$ are respectively

$$V(e_i) = E(e_i^2) - (E(e_i))^2 = \frac{M-2b}{M-b} 2P_i(1-P_i)$$

$$Cov(e_i, e_j) = -\frac{M-2b}{M-b} 2P_iP_j$$

The density function of $e_i$ for a general $n$ is

$$P(n,t) = \binom{n}{t} \frac{W(M_i, b, t) \times W(M-M_i, b, (n-t))}{W(M, b, n)} \tag{3}$$

with $W(M, b, n) = M(M-b)(M-2b)\ldots(M-(n-1)b)$ (4)

The expected value, variance, and covariance are

$$E(e_i) = nP_i \tag{5}$$

$$V(e_i) = \frac{M-nb}{M-b} nP_i(1-P_i) \tag{6}$$

$$Cov(e_i, e_j) = -\frac{M-nb}{M-b} nP_iP_j \tag{7}$$

## 3. One-stage Sampling

### 3.1. An estimator of the total and its variance

An unbiased estimator of the population total $X = \Sigma_{i=1}^{N} X_i$ is

$$\hat{X}_{SCG} = \sum_{i=1}^{n} X_i/nP_i \tag{8}$$

where $nP_i$ is the expected number of times of appearances of PSU $i$ in the sample.

The variance of the estimator is

$$V(\hat{X}_{SCG}) = \frac{1}{n^2} \sum_{i}^{N} (X_i/P_i)^2 V(e_i) + \sum_{i \neq j}^{N} (X_iX_j/P_iP_j) Cov(e_ie_j)$$

and taking into consideration (5), (6), and (7) we have

$$V(\hat{X}_{SCG}) = \frac{M-nb}{M-b} \times \frac{1}{n} \sum_{i}^{N} ((X_i/P_i) - X)^2 P_i = \frac{M-nb}{M-b} V(\hat{X}_{HH}) \tag{9}$$

where $V(\hat{X}_{HH})$ is the variance of the estimator $\hat{X}_{HH}$ of the population total $X$ for unequal probability sampling with replacement (upswr). Therefore we *always have*

$$V(\hat{X}_{SCG}) < V(\hat{X}_{HH}) \tag{10}$$

### 3.2. *An unbiased non-negative estimator for the variance*

The equation

$$\hat{V}(\hat{X}_{SCG}) = \frac{M-nb}{M}\frac{1}{n(n-1)}\sum_{i=1}^{n}\left(\frac{X_i}{P_i} - \hat{X}_{SCG}\right)^2 \tag{11}$$

is an unbiased non-negative estimator of the variance.

Because of the invariance to a change of origin, we have

$$E(\hat{V}(\hat{X}_{SCG})) = \frac{M-nb}{M}\times\frac{1}{n(n-1)}E\left(\sum_{i}^{n}((X_i/P_i) - X - (\hat{X}_{SCG} - X))^2\right)$$

$$= \frac{M-nb}{Mn(n-1)}\left(\sum_{i}^{N}((X_i/P_i) - X)^2 E(e_i) - nV(\hat{X}_{SCG})\right)$$

$$= \frac{M-nb}{Mn(n-1)}\times\left(\frac{n^2(M-b)}{M-nb}\times V(\hat{X}_{SCG}) - nV(\hat{X}_{SCG})\right)$$

and therefore

$$E(\hat{V}(\hat{X}_{SCG})) = V(\hat{X}_{SCG}) \tag{12}$$

And for $n = 2$ we found the simple expression

$$\hat{V}(\hat{X}_{SCG}) = \frac{M-2b}{M}\times\frac{1}{4}(X_1/P_1 - X_2/P_2)^2 \tag{13}$$

## 4. Multistage Sampling

### 4.1. *An estimator of the total and its variance*

An unbiased estimator for the total is

$$\hat{\hat{X}}_{SCG} = \sum_{i}^{n}\hat{X}_i/nP_i \tag{14}$$

where $\hat{X}_i$ is an estimator of $X_i$ based on the sample selected in the $i^{\text{th}}$ first-stage sampling unit. If this estimator is unbiased, also $\hat{\hat{X}}_{SCG}$ would be unbiased.

Let us assume that the selection and estimation take place independently at the various stages, and that the first stage units are selected according to the SCG selection procedure.

The unconditional variance of $\hat{\hat{X}}_{SCG}$ is

$$V\left(\hat{\hat{X}}_{SCG}\right) = V_1\left(E_2\left(\hat{\hat{X}}_{SCG}\right)\right) + E_1\left(V_2\left(\hat{\hat{X}}_{SCG}\right)\right)$$

where the index 2 refers to the second and subsequent stages.

We have

$$E_2\left(\hat{\hat{X}}_{SCG}\right) = E_2\left(\sum_{i}^{n}\hat{X}_i/nP_i\right) = \sum_{i}^{n}E_2(\hat{X}_i)/nP_i = \hat{X}_{SCG}$$

$$V_1\left(E_2\left(\hat{\hat{X}}_{SCG}\right)\right) = V_1(\hat{X}_{SCG})$$

and in the same way for the second component we have

$$V_2\left(\hat{\hat{X}}_{SCG}\right) = V_2\left(\sum_i^n \hat{X}_i/nP_i\right) = \frac{1}{n^2}\sum_i^n V_2(\hat{X}_i)/P_i^2 = \frac{1}{n^2}\sum_i^N V_2(\hat{X}_i)e_i/P_i^2$$

$$E_1 V_2(\hat{\hat{X}}_{SCG}) = \sum_i^N V_2(\hat{X}_j)/nP_i$$

And so, for the unconditional variance we obtain

$$V\left(\hat{\hat{X}}_{SCG}\right) = V_1(\hat{\hat{X}}_{SCG}) + \sum_i^N V_2(\hat{X}_i)/nP_i \tag{15}$$

where

$$V_1(\hat{\hat{X}}_{SCG}) = \frac{M - nb}{M - b} \times \frac{\sum_i^N ((X_i/P_i) - X)^2 P_i}{n}$$

and the second component will depend on the sampling method used in the second and subsequent stages.

### 4.2. *An unbiased non-negative estimator for the total variance*

An unbiased estimator of the total variance $V(\hat{\hat{X}}_{SCG})$ is

$$\hat{V}\left(\hat{\hat{X}}_{SCG}\right) = \frac{M - nb}{M} \times \frac{1}{n(n-1)} \sum_{i=1}^n \left(\frac{\hat{X}_i}{P_i} - \hat{\hat{X}}_{SCG}\right)^2 + \frac{nb}{M}\hat{W}$$

where $\hat{W} = 1/n^2 \sum_{i=1}^n \hat{V}(\hat{X}_i)/P_i^2$ is an unbiased non-negative estimator of

$$W = \sum_{i=1}^N \frac{V_2(\hat{X}_i)}{nP_i}$$

## 5. The Superpopulation Model Approach

In the superpopulation model used here, the finite population is considered as being drawn from an infinite superpopulation and the results do not apply to any single finite population but only to the average of all finite populations that can be drawn from the superpopulation.

Rao and Bayless (1969) and Bayless and Rao (1970) have applied the following often used model

$$X_i = BM_i + z_i \quad (i = 1, 2, \ldots N)$$

with

$$E^*(z_i|M_i) = 0 \qquad E^*(z_i^2|M_i) = aM_i^g$$

where $g$ is a constant which in most practical situations lies in the range $1 \leq g \leq 2$ and $a > 0$. The symbol $E^*$ denotes the expectations over all possible finite populations that hypothetically can be drawn from the model, conditioned on a fixed set of $M_i$.

We also assume that

$$E^*(z_i z_j | M_i M_j) = 0 \qquad i \neq j$$

For the comparison of variance estimators Rao and Bayless further assume that the $z_i$'s are normally distributed.

For the Hansen and Hurwitz method we have

$$E^* V(\hat{X}_{HH}) = \frac{aM^g}{n^g} \sum_i^N \left(1 - \frac{\mu_i}{n}\right) \mu_i^{g-1} \tag{16}$$

where $\mu_i = np_i$ is the expected number of appearances of unit $i$ in the sample.

The relation between the expected variances of the procedures HH and SCG is

$$E^* V(\hat{X}_{SCG}) = \frac{M - nb}{M - b} E^* V(\hat{X}_{HH})$$

which is the same as the relation between $V(\hat{X}_{SCG})$ and $V(\hat{X}_{HH})$ in Equation (9) in Section 3.1. above.

The expected reduction in variance for the SCG scheme relative to the HH procedure is

$$R = 1 - \frac{E^* V(\hat{X}_{SCG})}{E^* V(\hat{X}_{HH})} = \frac{b(n-1)}{M - b} = \frac{n - 1}{\frac{N\overline{M}}{b} - 1}$$

If we take into consideration the definition of $b$ given in Section 2 we have

$$R < \frac{n - 1}{\frac{N\overline{M}}{M_0}(n - 1)}$$

where $M_0$ is the minimum value for the $M_i$. Because $\overline{M}/M_0 > 1$ we also have

$$R < \frac{n - 1}{N(n - 1) - 1} = R_{\max} \tag{17}$$

We also find that the expected reduction in variance for HT relative to HH is

$$R_1 = 1 - \frac{E^* V(\hat{X}_{HT})}{E^* V(\hat{X}_{HH})} = 1 - \frac{\sum_i^N (1 - nP_i)}{\sum_i^N (1 - P_i)} = \frac{n - 1}{N - 1} \tag{18}$$

where $g = 1$.

If we compare the results in (17) and (18), we have the following indirect comparison of SCG relative to HT which appears in Table 1.

## 6. A Measure for Comparing the Variances of the SCG and HT Procedures

For a comparison of the gain in expected variance of the SCG and HT procedures (both relative to the HH procedure) we introduce the measure

$$\rho = \frac{E^* V(\hat{X}_{HH}) - E^* V(\hat{X}_{SCG})}{E^* V(\hat{X}_{HH}) - E^* V(\hat{X}_{HT})} \tag{19}$$

which is easy to conceptualise. The value zero can never be reached because the numerator will always be positive. On the other hand, the value $\rho = 1$ could only be reached if $E^* V(\hat{X}_{SCG}) = E^* V(\hat{X}_{HT})$ but this would be a degenerate case corresponding to cluster sampling with equal probabilities with $b = \overline{M}$.

For $g = 1$ in the superpopulation model, we have

$$\rho = \frac{1 - \dfrac{M - nb}{M - b}}{1 - \dfrac{N - n}{N - 1}} = \frac{b(N - 1)}{M - b} \qquad (0 < \rho < 1)$$

and for $g = 2$:

$$\rho' = \frac{1 - \dfrac{M - nb}{M - b}}{1 - \dfrac{1 - nD}{1 - D}} = \frac{(D^{-1} - 1)b}{M - b} \qquad (0 < \rho' < \rho < 1)$$

with $D = \sum_i^N P_i^2$ and $N > D^{-1}$, therefore $g = 1$ is slightly more favourable than $g = 2$ for the SCG sampling scheme.

## 7. Stability of the Variance Estimator

Under a superpopulation model approach an appropriate measure for the stability of the variance estimator is the expected value of the square of the coefficient of variation of the variance estimator (see Rao and Bayless 1969, p. 552) namely,

$$E^*\{CV^2[\hat{V}(\hat{X})]\} = E^*\left[\frac{V[\hat{V}(\hat{X})]}{\{E[\hat{V}(\hat{X})]\}^2}\right] \tag{20}$$

which is not easy to deal with because it is a ratio of two random variables.

Rao and Bayless (1969) and Bayless and Rao (1970) suggested the following alternative measure that we will represent by $I^2\hat{V}(\hat{X})$, where

$$I^2\hat{V}(\hat{X}) = \frac{E^*E(\hat{V}^2(\hat{X}))}{(E^*E(\hat{V}(\hat{X})))^2} - 1 \tag{21}$$

It actually measures the variability of the estimator of the variance around the average variance with the model considered.

*Table 1. The value of the ratio $R_{max}/R_1$ [$R_{max}$ given by Equation (17) and $R_1$ by Equation (18)] for different values of N and n.*

| $N \backslash n$ | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| 3 | 1.00 | – | – | – |
| 4 | 1.00 | 0.43 | – | – |
| 5 | 1.00 | 0.44 | 0.29 | – |
| 6 | 1.00 | 0.45 | 0.29 | 0.22 |
| 7 | 1.00 | 0.46 | 0.30 | 0.22 |
| 8 | 1.00 | 0.47 | 0.30 | 0.23 |
| 9 | 1.00 | 0.47 | 0.31 | 0.23 |
| 10 | 1.00 | 0.49 | 0.31 | 0.24 |
| 20 | 1.00 | 0.49 | 0.32 | 0.24 |

The authors mentioned above have used the following formulae for the second order moments, in which we consider only the case $g = 2$ that is slightly less favourable for the SCG scheme (see Section 6).

To evaluate (21) we need the second order moments and the square of the variance estimators for the procedures HT/B (Horvitz-Thompson/Brewer), HH, and SCG ($n = 2$ and $g = 2$).

For the first two procedures we will follow Rao and Bayless (1969), Equation 32, p. 551; Equations 37 and 39, p. 552, and Equation 43, p. 553:

$$E^*E\left(\hat{V}^2\left(\hat{X}_{HT/B}\right)\right) = \frac{3a^2M^4}{4}\sum_{i<j}^{N}\frac{\left(4P_iP_j - \pi_{ij}\right)^2}{\pi_{ij}} \tag{22}$$

$$E^*E\left(\hat{V}^2\left(\hat{X}_{HH}\right)\right) = \frac{3a^2M^4}{2}\sum_{i<j}^{N}P_iP_j \tag{23}$$

$$\left(E^*V\left(\hat{X}_{HT/B}\right)\right)^2 = \left(\frac{aM^2}{2}\sum_{i}^{N}(1 - 2P_i)P_i\right)^2 \tag{24}$$

$$\left(E^*V\left(\hat{X}_{HH}\right)\right)^2 = \left(\frac{aM^2}{2}\sum_{i}^{N}(1 - P_i)P_i\right)^2 \tag{25}$$

For the SCG sampling scheme we have

$$E^*E\left(\hat{V}^2\left(\hat{X}_{SCG}\right)\right) = \frac{(M-2b)^2}{M^2}\times E^*E\left(\hat{V}^2_{HH}\right) = \frac{(M-2b)^2}{M^2}\times\frac{3a^2M^4}{2}\sum_{i<j}^{N}P_iP_j$$

and

$$\left(E^*V\left(\hat{X}_{SCG}\right)\right)^2 = \left(\frac{M-2b}{M-b}\times\frac{aM^2}{2}\sum_{i}^{N}(1 - P_i)P_i\right)^2$$

With the above formulae we can obtain the Rao and Bayless indicators

$$I^2\hat{V}\left(\hat{X}_{HT/B}\right) = \frac{3\sum_{i<j}^{N}\frac{\left(4P_iP_j - \pi_{ij}\right)^2}{\pi_{ij}}}{\left(1 - 2\sum_{i}^{N}P_i^2\right)^2} - 1 \tag{26}$$

$$I^2\hat{V}\left(\hat{X}_{HH}\right) = \frac{6\sum_{i<j}^{N}P_iP_j}{\left(1 - \sum_{i}^{N}P_i^2\right)^2} - 1 \tag{27}$$

$$I^2\hat{V}\left(\hat{X}_{SCG}\right) = \left[\frac{(M-b)^2}{M^2}\times 6\sum_{i<j}^{N}P_iP_j\Big/\left(1 - \sum_{i}^{N}P_i^2\right)^2\right] - 1 \tag{28}$$

Rao and Bayless (1969) have computed the percentage gains in average efficiency for various variance estimators and for several values of the constant $g$. These results were given in Rao and Bayless (1969, Table 5, p. 555) for the procedure HT/B under a superpopulation model.

The percentage gain in average efficiency for a variance estimator $\hat{V}(\hat{X})$, under an arbitrary sampling scheme, negative to the variance estimator of the HT/B procedure is defined as

$$\left(\frac{I^2\hat{V}(\hat{X}_{HT/B})}{I^2\hat{V}(\hat{X})} - 1\right)100 \tag{29}$$

## 8. Efficiency of the SCG Sampling Scheme Relative to the HH Method

### 8.1. Expected variance

From Sections 4 and 5, we always have for both the variance and the expected variance that they are smaller with the SCG scheme than with the HH method.

### 8.2. Stability of the variance estimator

From (27) and (28) we have

$$I^2\hat{V}(\hat{X}_{SCG}) = \left(\frac{(M-b)^2(1+I^2\hat{V}(\hat{X}_{HH}))}{M^2}\right) - 1$$

and as $(M-b)^2/M^2 < 1$, we have

$$I^2\hat{V}(\hat{X}_{SCG}) < I^2\hat{V}(\hat{X}_{HH}) \tag{30}$$

Therefore we arrive at the conclusion: The variance estimator with the SCG sampling scheme is always more stable than the variance estimator with the HH method.

## 9. Efficiency of the SCG and HH Procedures Relative to the HT/B Method

### 9.1. Expected variance

Following Rao and Bayless (1969), p. 553, we will define the percentage gain in expected variance of $\hat{X}_{SCG}$ relative to $\hat{X}_{HT/B}$ as follows

$$e' = \left(\frac{E^*V(\hat{X}_{HT/B})}{E^*V(\hat{X}_{SCG})} - 1\right) \times 100 \tag{31}$$

We have

a) Both methods are equally efficient when $e' = 0$.
b) SCG is less efficient, in terms of expected variance, than HTB when $e' < 0$.
c) SCG is more efficient, in terms of expected variance, than HTB when $e' > 0$.

In a similar way we define the percentage gain in expected variance for the HH method relative to the HT/B procedure

$$e_1' = \left( \frac{E^* V(\hat{X}_{HT/B})}{E^* V(\hat{X}_{HH})} - 1 \right) \times 100 \tag{32}$$

The relationship between $e'$ and $e_1'$ is

$$e' = \left( \left( 1 + \frac{e_1'}{100} \right) \frac{M-b}{M-2b} - 1 \right) \times 100 \tag{33}$$

From (33) we can obtain the $e'$ values as a function of the $e_1'$ values computed by Rao and Bayless (1969, Table 4, p. 554) for the 20 natural populations included in their study.

### 9.2. *Stability of the variance estimator*

In accordance with Equation (29), let us now define the percentage gain in stability for the variance estimator with the SCG sampling scheme relative to the variance estimator of the HT/B procedure

$$e = \left( \frac{I^2 \hat{V}(\hat{X}_{HT/B})}{I^2 \hat{V}(\hat{X}_{SCG})} - 1 \right) \times 100 \tag{34}$$

In the same way we define the percentage gain in stability for the HH method relative to the HT/B procedure

$$e_1 = \left( \frac{I^2 \hat{V}(\hat{X}_{HT/B})}{I^2 \hat{V}(\hat{X}_{HH})} - 1 \right) \times 100 \tag{35}$$

The relationship between $e$ and $e_1$ is

$$e > \left( \frac{M^2}{(M-b)^2} \left( 1 + \frac{e_1}{100} \right) - 1 \right) \times 100 \tag{36}$$

as we will see now.

From (34) we have

$$1 + \frac{e}{100} = \frac{I^2 \hat{V}(\hat{X}_{HT/B})}{I^2 \hat{V}(\hat{X}_{SCG})}$$

and from (35) we have

$$1 + \frac{e_1}{100} = \frac{I^2 \hat{V}(\hat{X}_{HT/B})}{I^2 \hat{V}(\hat{X}_{HH})}$$

and therefore

$$\frac{1 + \dfrac{e}{100}}{1 + \dfrac{e_1}{100}} = \frac{I^2 \hat{V}(\hat{X}_{HH})}{I^2 \hat{V}(\hat{X}_{SCG})} = \frac{I^2 \hat{V}(\hat{X}_{HH})}{\dfrac{(M-b)^2}{M^2} I^2 \hat{V}(\hat{X}_{HH}) + \dfrac{(M-b)^2}{M^2} - 1}$$

### *9.3. Joint consideration of the stability for the variance estimator, and the expected variance of the total estimator, for the HH and SCG procedures relative to the HT/B method*

Table 2 shows both aspects for the 20 natural populations of Rao and Bayless (1969) for $g = 1$ and $n = 2$.

Positive values of the percentages of Table 2 indicate an advantage for SCG relative to HT/B. On the other hand, negative values indicate an advantage for HT/B relative to SCG.

In Table 2 we can observe that in all the 20 populations the values are positive for the stability of the variance estimator while all the values are negative for the expected variance. However, in quantitative terms for the 20 populations, we have an average percentage gain in expected variance in favour of HT/B, and an average gain in stability of the variance estimator in favour of SCG. The first amounts to five per cent and the second to 28 per cent. We think that we have found strong evidence of better stability for the variance estimator with the SCG sampling scheme relative to the HT/B procedure.

## 10. Some Conclusions About the SCG Sampling Scheme

a) The selection procedure is simpler, even for $n = 2$ than most of the alternatives (Section 2).
b) A finite population correction is available (Section 3.1).
c) The expected variance for the total, under a superpopulation model, is always smaller than in multinomial sampling (Section 5).
d) An unbiased non-negative estimator for the variance is always available (Section 3.2).
e) The variance estimator is always more stable than in multinomial sampling (Section 8.1).
f) We have found strong evidence of better stability for the variance estimator than with the Brewer-Durbin method (Section 9.3).
g) The SCG procedure has the Rotability, and the Ratio estimator properties (see Appendix).

Perhaps the mentioned advantages should be considered in the light of the following comments:

Cassel et al. (1977), p. 148. "... the HT strategy for example, may have to be rejected as impractical even though it is optimal by certain mathematical criteria and some suboptimal alternative may be preferred," and "practical advantages of a certain strategy may outweigh its loss of efficiency relative to the best strategy." And on p. 165 "The practical advantages of $M_{RHC}$ and $M_{HH}$ may therefore in certain situations outweigh their efficiency losses."

Brewer and Hanif (1983), p. 109 read: "... since there is no ideal way of proceeding using upswor, some may still prefer to use multinomial sampling and the Hansen-Hurwitz estimator. The variance reduction represented by the finite population correction is then entirely lost but the simplicity of the selection and estimation procedures and the further simplicity and stability of the variance estimation procedure leave little to be desired."

*Table 2. Percentage gains in stability and in expected variance relative to the HT/B method. (*)*

| $N^o$ | HH | $(e_1 =)$ | SCG | $(e >)$ | HH | $(e' =)$ | SCG | $(e' =)$ |
|---|---|---|---|---|---|---|---|---|
| 1 | − | 3 | + | 1.15 | − | 5 | − | 2.95 |
| 2 | − | 3 | + | 1.59 | − | 5 | − | 2.73 |
| 3 | − | 3 | + | 3.82 | − | 8 | − | 4.71 |
| 4 | + | 71 | + | 72.74 | − | 11 | − | 10.55 |
| 5 | + | 284 | + | 287.25 | − | 11 | − | 10.63 |
| 6 | − | 12 | + | 3.37 | − | 11 | − | 2.20 |
| 7 | − | 5 | + | 0.05 | − | 5 | − | 2.44 |
| 8 | + | 8 | + | 8.16 | − | 6 | − | 5.93 |
| 9 | + | 4 | + | 7.73 | − | 7 | − | 5.32 |
| 10 | + | 5 | + | 5.61 | − | 6 | − | 5.73 |
| 11 | − | 13 | + | 2.59 | − | 11 | − | 2.68 |
| 12 | + | 20 | + | 23.06 | − | 7 | − | 5.75 |
| 13 | + | 54 | + | 54.38 | − | 7 | − | 6.88 |
| 14 | + | 12 | + | 15.50 | − | 6 | − | 4.52 |
| 15 | + | 8 | + | 16.23 | − | 11 | − | 7.55 |
| 16 | + | 35 | + | 38.52 | − | 11 | − | 8.83 |
| 17 | + | 2 | + | 3.80 | − | 3 | − | 2.15 |
| 18 | − | 0 | + | 2.88 | − | 5 | − | 3.62 |
| 19 | − | 1 | + | 2.17 | − | 5 | − | 3.47 |
| 20 | + | 16 | + | 26.31 | − | 13 | − | 9.04 |

## Appendix

*Rotability and ratio estimator properties*

In the SCG procedure the unconditional probability of drawing unit $i$ is at any draw equal to $P_i$. This property could be important in schemes using rotation as was pointed out by Fellegi (1963).

The probability of unit $i$ at the first draw is $P_i$.

The probability of being selected at the second draw is

$$P_i \frac{M_i - b}{M - b} + (1 - P_i) \frac{M_i}{M - b} = P_i$$

and so on.

The estimator $\hat{X}_{SCG}$ also has the property (the ''ratio estimator property'') that if the $M_i$ are exactly proportional to the $X_i$ values, then the variance of $\hat{X}_{SCG}$ is equal to zero, because in this case $X_i/P_i$ is equal to $X$ which makes $V(\hat{X}_{SCG})$ equal to zero, see Equation (9).

## 11. References

Bayless, D.L. and Rao, J.N.K. (1970). An Empirical Study of the Stabilities of Estimators and Variance Estimators in Unequal Probability Sampling. Journal of the American Statistical Association, 65, 1645–1667.

* The per cent values of Table 2 have been taken from Rao and Bayless (1969), Tables 4 and 5. The values for the Sánchez-Crespo and Gabeiras (SCG) sampling scheme have been computed with formulae (33) and (36) of this article.

Brewer, K.R.W. (1963). A Model of Systematic Sampling with Unequal Probabilities. Australian Journal of Statistics, 5, 5–13.

Brewer, K.R.W. (1975). A Simple Procedure for Sampling Upswor. Australian Journal of Statistics, 17, 166–172.

Brewer, K.R.W. and Hanif, M. (1983). Sampling with Unequal Probabilities. New York, Springer-Verlag.

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

Cochran, W.G. (1977). Sampling Techniques. New York, J. Wiley and Sons.

Durbin, J. (1967). Design of Multi-stage Surveys for the Estimation of Sampling Errors. Applied Statistics, 16, 152–164.

Fellegi, I.P. (1963). Sampling with Varying Probabilities Without Replacement: Rotating and Non-rotating Samples. Journal of the American Statistical Association, 58, 183–201.

Hansen, M.H. and Hurwitz, W.N. (1943). On the Theory of Sampling from a Finite Population. Annals of Mathematical Statistics, 14, 333–362.

Horvitz, D.G. and Thompson, D.J. (1952). A Generalization of Sampling Without Replacement from a Finite Universe. Journal of the American Statistical Association, 47, 663–685.

Rao, J.N.K. and Bayless, D.L. (1969). An Empirical Study of the Stability of Estimators and Variance Estimators in Unequal Probability Sampling of Two Units per Stratum. Journal of the American Statistical Association, 64, 540–559.

Sánchez-Crespo, J.L. and Gabeiras, J.M. (1987). Un Esquema Mixto de Muestreo con Probabilidades Desiguales. Estadística Española, 115, 5–39. [In Spanish].

Sánchez-Crespo, J.L. (1989). Dos Nuevos Resultados en la Estrategia Mixta de Sánchez-Crespo y Gabeiras. Estadística Española, 120, 108–118. [In Spanish].

Received January 1990
Revised July 1997