Biometrical Journal **42** (2000) 7, 865–876

A Generalized Class of Synthetic Estimators with Application to Crop Acreage Estimation for Small Domains

G. C. Tikkiwal and Alka Ghiya

Department of Mathematics and Statistics
J.N.V. University
Jodhpur
India

Summary

This paper defines and discusses a generalized class of synthetic estimators for small domains, using auxiliary information, under simple random sampling and stratified random sampling schemes. The generalized class of synthetic estimators, among others, includes the simple, ratio and product synthetic estimators. The proposed class of synthetic estimators gives consistent estimators if the synthetic assumption holds. Further, it demonstrates the use of the generalized synthetic and ratio synthetic estimators for estimating crop acreage for small domains and also compare their relative performance with direct estimators, empirically, through a simulation study.

Key words: Simulation-cum-Regression (SICURE) model; Synthetic Estimators; Small Domains; Inspector Land Revenue Circles (ILRCs); Simple Random Sampling without Replacement (SRSWOR)-design; Timely Reporting Scheme (TRS); Absolute Relative Bias (ARB); Simulated relative standard error (Srse).

1. Introduction

The common feature of small area estimation problem is that when large-scale sample surveys are designed to produce reliable estimates at the national or state level; generally they do not provide estimates of adequate precision at lower levels like District, Tehsil/County, and Inspector Land Revenue Circle. This is because the sample size at the lower levels are generally insufficient to provide reliable estimates using traditional estimators. Therefore, the need was felt to develop alternative estimators to provide small area statistics using the data already collected through large-scale sample surveys. The traditional design based and alternative estimators are also termed, in the literature of small area estimation, respectively as direct and indirect estimators.

The indirect estimators are based on methods which increase the effective sample size either by (i) simulating enough data through appropriate analysis of available data under appropriate modeling or (ii) by using data from other domains and/or time periods through models that assume similarities across domain and/or time periods. The only known method so far belonging to Category (i) is SICURE-modeling [TIKKIWAL (1993)]. The other methods of estimation like Synthetic, Composite, Generalized Regression belong to Category (ii). Among these the synthetic estimators are used for small area estimation, mainly because of its simplicity, applicability to general sampling design and potential to increase accuracy in estimation. However, if the implicit model assumption of similarities across domain and/or time period fails, the synthetic estimator may be badly design biased. GONZALEZ (1973), GONZALEZ and WAKSBERG (1973) and GHANGURDE and SINGH (1977, 1978) among others study the synthetic estimator based on auxiliary variable viz. the ratio synthetic estimator. These studies show that synthetic estimators provide reliable estimates to some extent.

In this paper we define a generalized class of synthetic estimators, using auxiliary information, under simple random sampling and stratified random sampling schemes. The generalized class of synthetic estimators, among others, include the simple, ratio and product synthetic estimators. Further, we demonstrate the use of estimators belonging to the generalized class for estimating crop acreage for small domains and also compare their relative performance with the corresponding direct estimators, empirically, through a simulation study.

2. Formulation of the Problem and Notations:

Suppose that a finite population $U = (1, \ldots, i, \ldots, N)$ is divided into 'A' non-overlapping small domains U_a of size N_a ($a = 1, \ldots, A$) for which estimates are required. We denote the characteristic under study by 'y'. We further assume that the auxiliary information is available and denote this by 'x'. A random sample s of size n is selected through Simple Random Sampling Without Replacement (SRSWOR) design from population U such that n_a units in the sample s comes from small domain U_a ($a = 1, \ldots, A$). It may be noted that if there is Simple Random Sampling With Replacement (SRSWR) design, it can be dealt with similarly.

Consequently,

$$\sum_{a=1}^{A} N_a = N \quad \text{and} \quad \sum_{a=1}^{A} n_a = n\,.$$

We denote the various population and sample means for characteristic $Z = X, Y$ by

$\bar{Z}$ = mean of the population based on N observations.

$\bar{Z}_a$ = population mean of domain 'a' based on N_a observations.

$\bar{z}$ = mean of the sample 's' based on n observation.

$\bar{z}_a$ = sample mean of domain 'a' based on n_a observations.

Also, the various mean squares and coefficient of variations of the population 'U' for characteristic Z are denoted by

$$S_z^2 = \frac{1}{N-1} \sum_{i=1}^{N} (z_i - \bar{Z})^2 , \qquad C_z = \frac{S_z}{\bar{Z}} .$$

The coefficient of covariance between X and Y is denoted by

$$C_{xy} = \frac{S_{xy}}{\bar{X}\bar{Y}}$$

where,

$$S_{xy} = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{Y}) (x_i - \bar{X}) .$$

The corresponding various mean squares and coefficient of variations of small domains U_a are denoted by

$$S_{z_a}^2 = \frac{1}{N_a - 1} \sum_{i=1}^{N_a} (z_{a_i} - \bar{Z}_a)^2$$

$$C_{z_a} = \frac{S_{z_a}}{\bar{Z}_a} \quad \text{and} \quad C_{x_a y_a} = \frac{S_{x_a y_a}}{\bar{X}_a \bar{Y}_a}$$

where,

$$S_{x_a y_a} = \frac{1}{N_a - 1} \sum_{i=1}^{N_a} (y_{a_i} - \bar{Y}_a) (x_{a_i} - \bar{X}_a)$$

and z_{a_i} ($a = 1, \ldots, A$ and $i = 1, \ldots, N_a$) denote the ith observation of the small domain 'a' for the characteristic $Z = X, Y$.

3. Generalized class of Synthetic Estimators

We define a generalized class of synthetic estimators of population mean $\bar{Y}_a$, based on the auxiliary information 'x' under SRSWOR design, as described in previous section, as follows:

$$\bar{y}_{\text{syn},\,a} = \bar{y} \left(\frac{\bar{x}}{\bar{X}_a} \right)^{\beta} \tag{3.1}$$

where, β is a suitably chosen constant.

The above estimator $\bar{y}_{\text{syn},\,a}$ may be heavily biased unless the following assumption is satisfied.

$$\bar{Y}_a (\bar{X}_a)^{\beta} \doteq \bar{Y} (\bar{X})^{\beta} . \tag{3.2}$$

It may be noted that if there is strict equality in the above relation; then the Estimator (3.1) is a consistent estimator. This is so as the estimator reduces to $\bar{Y}_a$ when $n_a = N_a$ for all a.

Remark 3.1: If $\beta = 0$, then the estimator given in (3.1) reduces to simple synthetic estimator

$$\bar{y}_{\text{syn}, s, a} = \bar{y}$$

and the corresponding assumption given in (3.2) reduces to $\bar{Y}_a \doteq \bar{Y}$

Remark 3.2: If $\beta = -1$, then the estimator given in (3.1) reduces to ratio synthetic estimator

$$\bar{y}_{\text{syn}, a} = \frac{\bar{y}}{\bar{x}} \bar{X}_a$$

and the assumption given in (3.2) reduces to

$$\frac{\bar{Y}_a}{\bar{X}_a} \doteq \frac{\bar{Y}}{\bar{X}} .$$

Further, for $\beta = 1$, the generalized synthetic estimator reduces to product synthetic estimator.

4. Design Bias and Mean Square Error of Generalized Synthetic Estimator

In order to obtain bias and mean square error of generalized synthetic estimator $\bar{y}_{\text{syn}, a}$, let

$$\bar{y} = \bar{Y}(1 + \varepsilon_1); \qquad \bar{x} = \bar{X}(1 + \varepsilon_2)$$

so that $E(\varepsilon_1) = E(\varepsilon_2) = 0$ and

$$E(\varepsilon_1^2) = \frac{N-n}{Nn} C_y^2, \qquad E(\varepsilon_2^2) = \frac{N-n}{Nn} C_x^2$$

$$E(\varepsilon_1 \varepsilon_2) = \frac{N-n}{Nn} C_{xy} .$$

The $\bar{y}_{\text{syn}, a}$ can be expressed as

$$\bar{y}_{\text{syn}, a} = \bar{Y} \left(\frac{\bar{X}}{\bar{X}_a}\right)^{\beta} (1 + \varepsilon_1)(1 + \varepsilon_2)^{\beta} .$$

Assuming that $|\varepsilon_2| < 1$

$$\bar{y}_{\text{syn}, a} = \bar{Y} \left(\frac{\bar{X}}{\bar{X}_a}\right)^{\beta} \left(1 + \beta\varepsilon_2 + \frac{\beta(\beta-1)}{2} \varepsilon_2^2 + \varepsilon_1 + \beta\varepsilon_1\varepsilon_2 + \ldots\right) .$$

Therefore,

$$E(\bar{y}_{\text{syn}, a}) = \bar{Y} \left(\frac{\bar{X}}{\bar{X}_a}\right)^{\beta} \left[1 + \frac{N-n}{Nn} \left(\frac{\beta(\beta-1)}{2} C_x^2 + \beta C_{xy}\right)\right] .$$

Assuming further that the contribution of terms involving powers in ε_1 and ε_2 higher than the second to the value of $E(\bar{y}_{\text{syn},a})$ is negligible.

And the design bias of $\bar{y}_{\text{syn},a}$ is given by

$$B(\bar{y}_{\text{syn},a}) = \bar{Y}\left(\frac{\bar{X}}{\bar{X}_a}\right)^{\beta}\left[1+\frac{N-n}{Nn}\left(\frac{\beta(\beta-1)}{2}\,C_x^2+\beta C_{xy}\right)\right]-\bar{Y}_a\,. \quad (4.1)$$

If the synthetic assumption given in (3.2) satisfies than the above expression reduces to

$$B(\bar{y}_{\text{syn},a}) \;=\; \bar{Y}_a\left(\frac{N-n}{Nn}\right)\left[\left(\frac{1}{2}\,C_x^2\right)(\beta^2-\beta)+C_{xy}\beta\right].$$

And further design bias is zero either if $\beta = 0$ or $\beta = 1 - 2\,\dfrac{C_{xy}}{C_x^2}$.

The MSE of $\bar{y}_{\text{syn},a}$ is given by

$$\begin{aligned}\text{MSE}\,(\bar{y}_{\text{syn},a}) &= E(\bar{y}_{\text{syn},a}-\bar{Y}_a)^2\\ &= \bar{Y}^2\left(\frac{\bar{X}}{\bar{X}_a}\right)^{2\beta}\left[1+\frac{N-n}{Nn}\;\{(2\beta^2-\beta)\,C_x^2+C_y^2\;+4\beta C_{xy}\}\right]\\ &\quad -2\,\bar{Y}_a\,\bar{Y}\left(\frac{\bar{X}}{\bar{X}_a}\right)^{\beta}\left[1+\frac{N-n}{Nn}\left(\frac{\beta(\beta-1)}{2}\,C_x^2\;+\beta C_{xy}\right)\right]+\bar{Y}_a^2\,. \quad (4.2)\end{aligned}$$

The suitable value of β is the one for which MSE $(\bar{y}_{\text{syn},a})$ is minimum. So minimizing the MSE $(\bar{y}_{\text{syn},a})$ with respect to β, gives simplified expression for β, if $\bar{X}_a \doteq \bar{X}$ as follows

$$\beta = \frac{\bar{Y}(C_x^2-4C_{xy})-2\bar{Y}_a\left(\dfrac{C_x^2}{2}-C_{xy}\right)}{(4\bar{Y}C_x^2-2\bar{Y}_aC_x^2)}\,. \quad (4.3)$$

If the synthetic assumption given in (3.2) is satisfied then the above expression reduces to

$$\text{MSE}\,(\bar{y}_{\text{syn},a}) = \frac{N-n}{Nn}\;\bar{Y}_a^2(\beta^2C_x^2+C_y^2+2\beta C_{xy}) \quad (4.4)$$

and the value of β for which this expression of MSE $(\bar{y}_{\text{syn},a})$ minimizes, is given by $\beta = -\dfrac{C_{xy}}{C_x^2}$.

5. Generalized Class of Synthetic Estimators under Stratification

Suppose that the finite population $U = (1, \ldots, i, \ldots, N)$ is divided into 'A' non-overlapping domains $U_{.a}$, of size $N_{.a}$ $(a = 1, \ldots, A)$, for which estimates are required as discussed in Section 2. The population is also divided along a second dimension into 'H' non-overlapping categories (called groups) $U_{h.}$ of size $N_{h.}$ $(h = 1, \ldots, H)$. As a result, the population is cross classified into HA cells, U_{ha}, of respective sizes N_{ha}.. Consequently,

$$N = \sum_{h=1}^{H} N_{h.} = \sum_{a=1}^{A} N_{.a} = \sum_{h=1}^{H} \sum_{a=1}^{A} N_{ha} . \tag{5.1}$$

We assume that N_{ha} are known from a previous census or other reliable sources.

Further, we assume that simple random samples of predetermined size $n_{h.}$ $(h = 1, \ldots, H)$ are selected from group h such that $\sum_{h=1}^{H} n_{h.} = n$. That is, n is the size of the random sample selected using stratified random sampling. Also let $n_{.a}$ and n_{ha} $(a = 1, \ldots, A;\ h = 1, \ldots, H)$ are the units of the sample that belongs to domain $U_{.a}$ and cell (h, a). So $n_{.a}$ and n_{ha} are random.

Denoting $y_{ha_i}(i = 1, \ldots, N_{ha}$, the ith observation of the characteristic under study of the cell (h, a), we define various population and sample means as follows, using capital letters for population means and small letters for sample means.

$$\bar{Y}_{a.} = \frac{1}{N_{.a}} \sum_{h=1}^{H} N_{ha}\bar{Y}_{ha} , \qquad \text{Population mean of small area '}a\text{'}$$

where

$$\bar{Y}_{ha} = \frac{1}{N_{ha}} \sum_{i=1}^{N_{ha}} y_{ha_i}$$

$$\bar{y}_{a.} = \frac{1}{n_{.a}} \sum_{h=1}^{H} n_{ha}\bar{y}_{ha} , \qquad \text{sample mean of small area '}a\text{'}$$

where

$$\bar{y}_{ha} = \frac{1}{n_{ha}} \sum_{i=1}^{n_{ha}} y_{ha_i}$$

$$\bar{Y}_{h.} = \frac{1}{N_{h.}} \sum_{a=1}^{A} N_{ha}\bar{Y}_{ha} , \qquad \text{population mean of the } h\text{th group}$$

$$\bar{y}_{h.} = \frac{1}{n_{h.}} \sum_{a=1}^{A} n_{ha}\bar{y}_{ha} , \qquad \text{sample mean of the } h\text{th group}$$

Similar notations are used, for various means for auxiliary characteristic x, just replacing 'y' with 'x' symbol.

We then, following (3.1), define Generalized Synthetic estimator under stratification as follows.

$$\bar{y}_{S,\,\text{syn},\,a} = \sum_{h=1}^{H} w_{ha}\bar{y}_{h.} \left(\frac{\bar{x}_{h.}}{\bar{X}_{ha}}\right)^{\beta} . \tag{5.2}$$

Under the synthetic assumption

$$\bar{Y}_{ha}\,(\bar{X}_{ha})^{\beta} \doteq \bar{Y}_{h.}(\bar{X}_{h.})^{\beta} \tag{5.3}$$

where, $w_{ha} = \dfrac{N_{ha}}{N_{.a}}$

Now the design bias of $\bar{y}_{S,\text{syn},a}$ is given by

$$B(\bar{y}_{S,\,\text{syn},\,a}) = E(\bar{y}_{S,\,\text{syn},\,a}) - \bar{Y}_{a.}$$

$$= \sum_{h=1}^{H} w_{ha}\left(\bar{Y}_{h.}\left(\frac{\bar{X}_{h.}}{\bar{X}_{ha}}\right)^{\beta}\left(1+\frac{N_{h.}-n_{h.}}{N_{h.}n_{h.}}\left(\frac{\beta(\beta-1)}{2}\,C_{hx}^{2}+\beta C_{hxy}\right)\right)-\bar{Y}_{ha}\right). \tag{5.4}$$

If the synthetic assumption (5.3) satisfies, the expression of the bias reduces to

$$B(\bar{y}_{S,\,\text{syn},\,a}) \;=\; \sum_{h=1}^{H} w_{ha}\left(\bar{Y}_{ha}\left(\frac{N_{h.}\;-n_{h.}}{N_{h.}n_{h.}}\right)\left(\frac{\beta(\beta-1)}{2}\,C_{hx}^{2}\;+\;\beta\,C_{hxy}\right)\right). \tag{5.5}$$

Similarly, expression of mean square error of the $\bar{y}_{S,\,\text{syn},\,a}$ can be obtained, using the expression (4.2) of MSE $(\bar{y}_{\text{syn},\,a})$.

Remark 5.1: For $\beta = 0$, the estimator (5.2) reduces to

$$\bar{y}'_{S,\,\text{syn},\,a} = \sum_{h=1}^{H} w_{ha}\bar{y}_{h.}\,, \qquad \text{which in turn gives}$$

$$\hat{T}_{S,\,\text{syn},\,a} = \sum_{h=1}^{H} N_{ha}\bar{y}_{h.}$$

the estimator of population total 'T_a' of small area 'a', discussed by SARNDAL [1984, Eq. (3.1), p. 625].

Also the expression (5.4) of bias for $\beta = 0$ reduces to

$$B(\bar{y}'_{S,\,\text{syn},\,a}) = \sum_{h=1}^{H} w_{ha}(\bar{Y}_{h.} - \bar{Y}_{ha})$$

which in turn gives the expression of bias of $\hat{T}_{S,\,\text{syn},\,a}$

$$B(\hat{T}_{S,\,\text{syn},\,a}) = \sum_{h=1}^{H} N_{ha}(\bar{Y}_{h.} - \bar{Y}_{ha})$$

[see SARNDAL 1984, Eq. (6.1), p. 628].

Remark 5.2: CASSEL et al. (1987) uses the above estimator $\bar{y}'_{S,\,\text{syn},\,a}$ to provide estimates of unemployment at municipal level in Swedish Labor Force Survey. The performance of this estimator proves to be better than the corrected synthetic estimator.

Remark 5.3: The ratio and product synthetic estimators under stratification can be obtained for different values of β as obtained, for estimator $\bar{y}_{\text{syn},\,a}$. For example, for $\beta = -1$, Generalized estimator (5.2) reduces to

$$\bar{y}''_{S,\,\mathrm{syn},\,a} = \sum_{h=1}^{H} w_{ha} \frac{\bar{y}_{h.}}{\bar{x}_{h.}} \bar{X}_{ha} .$$

This estimator is currently in use to provide improved estimator of States' income in USA [see SCHAIBLE (1996), pp. 28–57].

6. Crop Acreage Estimation for Small Domains – A Simulation Study

In this section we demonstrate the use of the generalized synthetic and ratio synthetic estimators to obtain crop acreage estimates for small domains and also compare their relative performance with the corresponding direct estimators empirically, through a simulation study. This we do by taking up the state of Rajasthan, one of the states in India, for case study.

6.1 *Existing methodology for estimation*

In order to improve timelines and quality of crop acreage statistics, a scheme known as Timely Reporting Scheme (TRS) has been in vogue since early seventies in most of the States of India. The TRS has the objective of providing quick and reliable estimates of crop acreage statistics and there-by productions of the principle crops during each agricultural season. Under the scheme the Patwari (Village Accountant) is required to collect acreage statistics on a priority basis in a 20 percent sample of villages, selected by stratified linear systematic sampling design taking Tehsil (a sub-division of the District) as a stratum. These statistics are further used to provide state level estimates using direct estimators viz. Unbiased (based on sample mean) and ratio estimators.

The performance of both the estimators in the State of Rajasthan, like in other states, is satisfactory at state level, as the sampling error is within 5 percent. However, the sampling error of both the estimators increases considerably, when they are used for estimating acreage statistics of various principle crops even at district level, what to speak of levels lower than a district. For example, the sampling error of direct ratio estimator for Kharif crops (the crop sown in June–July and harvested in October–November every year) of Jodhpur district (of Rajasthan State) for the agricultural season 1991–92 varies approximately between 6 to 68 percent. Therefore, there is need to use indirect estimators at district and lower levels for decentralized planning and other purposes like crop insurance.

6.2 *Details of the simulation study*

For the collection of revenue and other administrative purposes, the State of Rajasthan, like most of the other states of India, is divided into a number of districts.

Further, each district is divided into a number of Tehsils and each Tehsil is also divided into a number of Inspector Land Revenue Circles (ILRCs). Each ILRC consists of a number of villages. For the present study, we take ILRCs as small domains.

In the simulation study, we undertake the problem of crop acreage estimation for all Inspector Land Revenue Circles (ILRCs) of Jodhpur Tehsil of Rajasthan. They are seven in number. These ILRCs are small domains from the TRS point of view. The crop under consideration is Bajra (Indian corn or millet) for the agriculture season 1993–94. The bajra crop acreage for agriculture season 1992–93 is taken as the auxiliary characteristic x. The various information regarding the ILRCs of Jodhpur Tehsil are provided in Table 6.2.1.

We now give below the list of all those estimators, whose relative performance is to be assessed for estimating population total T_a of small domain 'a' for $a = 1, 2, \ldots, 7$.

Direct Estimators:

Direct ratio estimator $$t_{1,a} = N_a \left(\frac{\bar{y}_a}{\bar{x}_a}\right) \bar{X}_a$$

Direct general estimator $$t_{2,a} = N_a \, \bar{y}_a \left(\frac{\bar{x}_a}{\bar{X}_a}\right)^{\beta}$$

Indirect Estimators:

Ratio synthetic estimator $$t_{3,a} = N_a \left(\frac{\bar{y}}{\bar{x}}\right) \bar{X}_a$$

Generalized synthetic estimator $$t_{4,a} = N_a \, \bar{y} \left(\frac{\bar{x}}{\bar{X}_a}\right)^{\beta}$$

Table 6.2.1

Total area (Irrigated and Unirrigated) under Bajra crop in Inspector Land Revenue Circles (ILRCs) of Jodhpur Tehsil for Agricultural Season 1992–93 and 1993–94

S. No.	ILRC of Jodhpur Tehsil	No. of villages in ILRC	Total area (Irr. + U.Irr.) under the crop Bajra in 1992–93	Total area (Irr. + U.Irr.) under the crop Bajra in 1993–94
1	Jodhpur (1)	29	7799.5899	5696.5000
2	Keru (2)	44	21209.5880	15699.6656
3	Dhundhada (3)	32	19019.0288	16476.4863
4	Bisalpur (4)	30	15153.9248	14269.0000
5	Luni (5)	33	19570.1323	16821.4508
6	Dhava (6)	40	25940.0979	25075.5000
7	Jajiwal Kalan (7)	44	18007.4120	15875.0000
	Total	**252**	**126699.7737**	**109913.6027**

Before simulation, we first examine the following assumption, given earlier in (3.1), for ratio synthetic estimator $t_{3,a}$ and generalized synthetic estimator $t_{4,a}$ with respect to the seven small domains under study:

$$\bar{Y}_a(\bar{X}_a)^{\beta} \doteq \bar{Y}(\bar{X})^{\beta}.$$

For $\beta = -1$ for estimator $t_{3,a}$ and for the optimum value of β given in (4.4) for estimator $t_{4,a}$, the Tables 6.2.2 and 6.2.3 provide below absolute differences between $\bar{Y}_a/\bar{X}_a$ and $\bar{Y}/\bar{X}$, and between $\bar{Y}_a(\bar{X}_a)^{\beta}$ and $\bar{Y}(\bar{X})^{\beta}$ for all the seven ILRCs respectively. From the examination of these tables we note in this study that both the assumptions closely meet in ILRCs 3, 5 and 7, deviate moderately in ILRCs 4 and 6, but deviate considerably in ILRCs 1 and 2.

Table 6.2.2

Absolute Difference under Synthetic Assumption of Ratio Synthetic Estimator for various ILRCs

ILRC	$(\bar{Y}_a/\bar{X}_a)$	$(\bar{Y}/\bar{X})$	Absolute Difference $\lvert(\bar{Y}_a/\bar{X}_a) - (\bar{Y}/\bar{X})\rvert$
(1)	0.73036	0.86751	0.1371
(2)	0.7402	0.86751	0.1273
(3)	0.8663	0.86751	0.00119
(4)	0.9416	0.86751	0.0741
(5)	0.8595	0.86751	0.0079
(6)	0.9666	0.86751	0.099
(7)	0.8815	0.86751	0.014

Table 6.2.3

Absolute Difference under Synthetic Assumption of Generalized Synthetic Estimator for various ILRCs

ILRC	$\bar{Y}_a\,(\bar{X}_a)^{\beta}$	$\bar{Y}\,(\bar{X})^{\beta}$	Absolute difference $\lvert\bar{Y}_a(\bar{X}_a)^{\beta} - \bar{Y}(\bar{X})^{\beta}\rvert$
(1)	3.31157	4.6578	1.346
(2)	2.11349	2.4947	0.3812
(3)	0.77584	0.7791	0.00331
(4)	1.23143	1.1343	0.0971
(5)	0.81360	0.82231	0.008706
(6)	2.40008	2.44412	0.04403
(7)	0.14251	0.13789	0.00462

Now taking villages as sampling units for simulation purposes and otherwise, 500 independent simple random samples for each size of 25, 50, 63, 76 and 88 are selected from the population of 252 villages of Jodhpur Tehsil. Then, to assess the relative performance of the estimators under consideration, their Absolute Relative Bias (ARB) and Simulated relative standard error (Srse) or simply coefficient of variation are calculated for each ILRC as follows:

$$\text{ARB}\ (t_{k,a}) = \frac{\left| \dfrac{1}{500} \sum_{s=1}^{500} t_{k,a}^{s} - T_a \right|}{T_a} \times 100 \tag{6.2.1}$$

and

$$\text{Srse}\ (t_{k,a}) = \frac{\sqrt{\text{ASE}\ (t_{k,a})}}{E(t_{k,a})} \times 100 \tag{6.2.2}$$

where

$$\text{ASE}\ (t_{k,a}) = \frac{1}{500} \sum_{s=1}^{500} (t_{k,a}^{s} - T_a)^2 \tag{6.2.3}$$

for $k = 1, 2, \ldots, 7$ and $a = 1, 2, \ldots, 7$.

6.3 *Results*

We present the results of ARB and Srse in Table (6.3.1) only for $n = 50$ (a sample of 20 percent villages, as presently adopted in TRS) as the findings from other tables are similar.

For assessing relative performance of the various estimators, we have to adopt some rule of thumb. Here, we adopt the rule that at the ILRC level, an estimator should not have Srse more than 10 percent and bias more than 5 percent. We note from the table that none of the estimators satisfy the rule in ILRCs 1 and 2. This is happening because, in these circles, there is considerable deviation from the

Table 6.3.1

Simulated Relative Standard Errors and Absolute Relative Biases, in percentage, of various Estimators for $n = 50$ in Different ILRCs

Estimator	ILRC						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$t_{1,a}$	37.37	17.46	8.51	16.29	12.73	12.28	15.29
	(0.21)	(2.28)	(0.76)	(0.13)	(2.41)	(0.32)	(2.78)
$t_{2,a}$	18.55	18.32	6.56	15.43	1.27	13.68	1.34
	(0.96)	(1.50)	(0.12)	(0.18)	(1.12)	(0.54)	(0.61)
$t_{3,a}$	40.54	21.00	5.96	10.17	5.95	12.14	7.16
	(39.67)	(19.84)	(0.11)	(8.68)	(0.18)	(11.03)	(3.97)
$t_{4,a}$	19.11	20.67	5.71	10.11	5.71	8.43	5.85
	(17.90)	(19.50)	(0.72)	(8.66)	(0.05)	(4.06)	(1.02)

Note: The figures shown in parenthesis are the absolute relative biases in percentage.

synthetic assumptions, as observed earlier. In ILRCs 4 and 6, where the assumptions deviate moderately, $t_{4,a}$ alone satisfies the rule to some extent. In ILRCs 3, 5 and 7, where the synthetic assumptions closely meet, both $t_{3,a}$ and $t_{4,a}$ satisfy the rule but $t_{4,a}$'s performance is slightly better than $t_{3,a}$.

From the above analysis it is clear that if the synthetic estimators do not deviate considerably from their corresponding synthetic assumptions then, performance of the synthetic estimators $t_{3,a}$ and $t_{4,a}$, based on a sample of 20 percent villages (as presently being taken under TRS), is satisfactory at the level of ILRCs. Therefor, these estimators are also likely to perform better both at Tehsil and district levels. When the synthetic estimators deviate considerably from their corresponding synthetic assumptions then we should look for other types of estimators such as those obtained through the SICURE MODEL [TIKKIWAL, 1993] and assess their relative performance through studies of the kind, in series, over some years for crop acreage estimation.

References

CASELL, C. M., KRISTIANSSON, E. K., RABACK, G., and WAHLSTROM, 1987: *Using model-biased estimation to improve the estimate of unemployment on a regional level in the Swedish Labor Force survey.* Small Area Statistics, John Wiley and Sons, New York.

GHANGURDE, P. D. and SINGH, M. P., 1977: Synthetic estimates in periodic household surveys. *Survey Methodology*, **3**, 151–181.

GHANGURDE, P. D. and SINGH, M. P., 1978: Evaluation of efficiency of synthetic estimates. *Proceeding of the Social Statistical Section of the American Statistical Association,* 53–61.

Gonzalez, M. E., 1973: Use and evaluation of synthetic estimates. Proceedings of the Social Statistical Section of American Statistical Association, 33–36.

GONZALEZ, M. E. and WAKSBERG, J. 1973: *Estimation of the error of synthetic estimates.* Paper presented at First Meeting of the International Association of Survey Statisticians, Vienna, Austria, 18–25.

SARNDAL, C. E., 1984: Design – consistent versus model – dependent estimators for small domains. *Journal of the American Statistical Association,* **79**, 624–531.

SCHAIBLE, W. L., 1996: *Indirect Estimation in U.S. Federal Programs.* Research Monograph Springer-Verlag.

TIKKIWAL, B. D., 1993: Modelling through survey data for small domains. *Proceedings International Scientific conference on Small Area Statistics and Survey Design* held in September, 1992 at Warsaw, Poland.

Dr. G. C. TIKKIWAL
Department of Mathematics
and Statistics (New Campus)
J. N. V. University,
Jodhpur – 342 001
India

Received, April 1999
Revised, June 2000
Accepted, July 2000

Ms. ALKA GHIYA
Department of Mathematics
and Statistics (New Campus)
J. N. V. University,
Jodhpur – 342 001
India