# SOME RATIO-TYPE ESTIMATORS** 

R.P. Chakrabarty<br>Jackson State University, Jackson MS 39211 USA<br>(Received : August, 1977)

## 1. Introduction

One of the main objectives of a sample survey is the estimation of the population mean or total of a characteristic ' $y$ ' attached to the units in the population. Ratio estimators are among the most commonly used estimators of the population mean or total of ' $y$ ' utilizing an auxiliary characteristic ' $x$ ' that is positively correlated with ' $y$ '. The precision of the regression estimator is usually higher than that of the ratio estimator but in large-scale sample surveys, the ratio estimator is frequently employed because of its simplicity. In this paper, we develop some ratio-type estimators which will be more efficient than the customary ratio estimator and/or the unbiased estimator and yet computationally comparable to the customary ratio estimator.

We shall, without loss of generality, confine ourselves to the estimation of $\bar{T}$, the population mean of ' $y$ '. Further, to simplify the discussion, we shall confine ourselves to simple random sampling and assume the population size is infinite. From a simple random sample of $n$ pairs $\left(y_{i}, x_{i}\right)$ we have the unbiased estimator of $\bar{x}$, as

$$
\begin{equation*}
\bar{y}=\sum_{i=1}^{n} y_{i} / n \tag{1.1}
\end{equation*}
$$

The customary ratio estimator of $\bar{Y}$ is

$$
\begin{equation*}
\bar{y}_{\mathrm{r}}=(\bar{y} / \bar{x}) \bar{X}=r \bar{X} \tag{1.2}
\end{equation*}
$$

where $\bar{x}$ is the sample mean and $\bar{X}$ is the known population mean of $x$, and

$$
\begin{equation*}
r=\bar{y} / \bar{x} \tag{1.3}
\end{equation*}
$$

is the ratio estimator of the ratio $R=\bar{X} / \bar{X}$.
It is well known that the ratio estimator $\overline{\boldsymbol{y}}_{r}$ is more efficient than the unbiased estimator $\bar{y}$ in large samples if $\rho>C_{x} /\left(2 C_{y}\right)$

[^0]where $\rho$ is the coefficient of correlation between $y$ and $x$ and $C_{y}$ and $C_{x}$ are coefficients of variation of $y$ and $x$ respectively. The question of choice between $\bar{y}$ and $\bar{y}_{r}$ arises when it is suspected that $\rho(\geqslant 0)$ is not high and/or $C_{x}>C_{y}$. The customary procedure in such situations is to use $\bar{y}_{r}$ when $\rho>C_{x} /\left(2 C_{y}\right)$ otherwise use $\bar{y}$. It is, however, desirable to develop alternative ratio-type estimators which are more efficient than $\bar{y}_{r}$ as well as $\bar{y}$ and yet computationally comparable to $\bar{y}_{r}$. The two ratio-type estimators we propose are
\[

$$
\begin{equation*}
t_{1}=(1-W) \bar{y}+W \bar{y}_{r} ; W \geqslant 0 \tag{1.4}
\end{equation*}
$$

\]

and

$$
\begin{equation*}
t_{2}=(1-W) \bar{y}+W r^{*} \bar{X} ; W \geqslant 0 \tag{1.5}
\end{equation*}
$$

where $W$ is a constant weight to be determined and

$$
\begin{equation*}
r^{*}=2 r-\frac{1}{2}\left(r_{1}+r_{2}\right) \tag{1.6}
\end{equation*}
$$

is obtained by splitting the sample at random into two groups, each of size $n / 2$ when $n$ is even and $r_{j}=\tilde{y}_{j} / \bar{x}_{j},(j=1,2), \bar{y}_{j}$ and $\bar{x}_{j}$ are means of $y$ and $x$ respectively obtained from $j$ th half-sample. The estimator $t_{1}$ reduces to $\bar{y}$ and $\bar{y}_{r}$ when $W=0$ and 1 respectively. The estimator $t_{2}$ reduces to $\bar{y}$ when $W=0$ and when $W=1$ it reduces to $r^{*} \bar{X}$ which is the 'Jack-knife' ratio estimator of $\bar{Y}$. It may be mentioned here that by dividing the sample at random into $g(\leqslant n)$ groups, each of size $n / g$, a more general form of the estimator $t_{2}$ could be obtained as

$$
t_{2 g}=(1-W) \dot{y}+W\left[g r-\frac{g-1}{g} \sum_{j=1}^{g} r_{j}^{\prime}\right] \tilde{X}
$$

where $r_{g}^{\prime}$ is the customary ratio estimator calculated from the sample after omitting the $j$ th group. However, in this paper we shall consider the special case of $t_{2 g}$ given in (1.5). Srivastava (1967) proposed the estimator

$$
\begin{equation*}
t_{3}=\bar{y}(\bar{X} / \bar{x})^{W} \tag{1.7}
\end{equation*}
$$

where $W$ is a constant weight and obtained its asymptotic variance. The estimator $\dot{t}_{1}$ was suggested earlier by Chakrabarty (1968). In this paper these estimators will be compared regarding the properties of bias and efficiency. In section 2, we discuss the asymptotic theory and in section 3 we give the exact biases and variances of these estimators under a regression model.

## 2. Asymptotic Theory

### 2.1. Biases of the estimators

It is obvious that the estimators $t_{1}, t_{2}$, and $t_{3}$ are consistent but in general biased, like the ratio estimator $\bar{y}_{r}$. Now, as it is customary in the asymptotic theory of ratio method of estimation, we shall assume that the sample size $n$ is sufficiently large so that

$$
\begin{equation*}
\left|\delta_{\bar{x}}\right|=\left|\frac{\bar{x}-\bar{X}}{\bar{X}}\right| \ll 1 \tag{2.1}
\end{equation*}
$$

Under the above assumption, the expected value of $r$ is given by

$$
E(r)=R+\frac{R}{n}\left(C_{x}^{2}-\rho C_{y} C_{x}\right)+0\left(n^{-2}\right)
$$

Now, since $r_{1}$ and $r_{2}$ are independent

$$
E\left(r^{*}\right)=R+0\left(n^{-2}\right)
$$

Consequently, the biases of $t_{1}$ and $t_{2}$ are

$$
\begin{align*}
\cdot \operatorname{Bias}\left(t_{1}\right) & =W \operatorname{Bias}\left(\bar{y}_{r}\right) \\
& =\frac{W \bar{T}}{n}\left(C_{x}^{2}-\rho C_{y} C_{x}\right)+0\left(n^{-2}\right) \tag{2.2}
\end{align*}
$$

and

$$
\begin{equation*}
\operatorname{Bias}\left(t_{2}\right)=0+0\left(n^{-2}\right) \tag{2.3}
\end{equation*}
$$

respectively. From Srivastava (1967), the bias $t_{3}$ is given by

$$
\begin{equation*}
\operatorname{Bias}\left(t_{3}\right)=\frac{W \bar{T}}{n}\left[\frac{(W+1)}{2} C_{x}^{2}-\rho C_{y} C_{x}\right]+0\left(n^{-2}\right) \tag{2.4}
\end{equation*}
$$

Thus, the asymptotic bias of $t_{2}$ is of order $n^{-2}$ and hence smaller than that of $\bar{y}_{r}, t_{1}$ and $t_{3}$ whose biases are order $n^{-1}$. The bias of $t_{1}$ is smaller than that of $\bar{y}_{r}$ for $0<W<1$. We note that $C_{x}^{2}-{ }_{p} C_{v} C_{x}=0$ when the regression of $y$ on $x$ passes through the origin. Consequently, for the important case of regression through the origin the estimators $\bar{y}_{r}$ and $t_{1}$ are unbiased to terms of order $n^{-1}$ but the bias of $t_{3}$ is still of order $n^{-1}$. Further, substituting the formula for exact bias of $\bar{y}_{r}$ from Hartley and Ross (1954) we get the exact bias of $t_{1}$ as.

$$
\operatorname{Bias}\left(t_{1}\right)=-W \operatorname{Cov}(r, x)
$$

and

$$
\begin{equation*}
\frac{\left|\operatorname{Bias}\left(t_{1}\right)\right|}{\sigma \bar{y}_{r}} \leqslant \frac{W C_{n}}{\sqrt{n}} \tag{2.5}
\end{equation*}
$$

Thus if $\frac{W C_{x}}{\sqrt{n}} \leqslant 0.1$, the bias of $t_{1}$ is negligible in relation to the standard error of $\bar{y}_{r}$. No such upper bound to the bias of $t_{3}$ relative to its standard error could be obtained.

### 2.2. Variances of the estimators

In deriving the variances of estimators $t_{1}, t_{2}$ and $t_{3}$ we consider up to terms of $n^{-1}$ only and biases which are of order $n^{-1}$ are neglected. Expanding $r$ and $r_{j}$ by Tylor's series in terms of $\delta_{\bar{x}}, \delta_{\bar{y}}$ and $\delta_{\bar{x}_{j}}, \delta_{\bar{y}_{j}}(j=1,2)$ it can be shown that to terms of order $n^{-1}$ the variances of $t_{1}, t_{2}$ and $t_{3}$ are identical and are given by

$$
\begin{equation*}
V\left(t_{1}\right)=V\left(t_{2}\right)=V\left(t_{3}\right)=\frac{S_{y}^{2}}{n}[1+W K(W K-2 \mathrm{p})] \tag{2.6}
\end{equation*}
$$

where

$$
\begin{equation*}
K=C_{x} / C_{y} . \tag{2.7}
\end{equation*}
$$

The value of $W$ which minimizes this variance is

$$
\begin{equation*}
W_{o p t}=\mathrm{p} / K \tag{2.8}
\end{equation*}
$$

The minimum variance is given by

$$
\begin{equation*}
V_{m i n}=\frac{S_{y}^{2}}{n}\left(1-\rho^{2}\right) \tag{2.9}
\end{equation*}
$$

which is equal to the variance of the linear regression estimator up to terms of order $n^{-1}$. Substituting $W=1$ in (2.6) we get the variance of $\bar{y}_{r}$ as

$$
\begin{equation*}
V\left(\bar{y}_{r}\right)=\frac{S_{y}^{2}}{n}[1+K(K-2 \rho)] \tag{2.10}
\end{equation*}
$$

The asymptotic efficiencies of $t_{1}\left(t_{2}\right.$ and $\left.t_{3}\right)$ over $\bar{y}$ and $\bar{y}_{r}$ are given by

$$
\begin{equation*}
E_{1}=\frac{V(\bar{y})}{V\left(t_{1}\right)}=\frac{1}{[1+W K(W K-2 \rho)]} \tag{2.11}
\end{equation*}
$$

and

$$
\begin{equation*}
E_{2}=\frac{V\left(\bar{y}_{r}\right)}{V\left(t_{1}\right)}=\frac{[1+K(K-2 \rho)]}{[1+W K(W K-2 \rho)]} \tag{2.12}
\end{equation*}
$$

respectively. From (2.11) and (2.12) we get

$$
E_{1} \geqslant 1 \quad \text { if } \quad W \leqslant 2 \rho / K
$$

and

$$
\begin{equation*}
E_{2} \geqslant 1 \text { if }(2 \rho-K) / K \leqslant W \leqslant 1 \tag{2.13}
\end{equation*}
$$

Thus the estimators $t_{1}, t_{2}$ and $t_{3}$ are better than $\bar{y}$ and $\bar{y}_{r}$ for a wide range of $W$-values. For example, if $\rho=6, K=1$ and $W$ is between 0.2 and 1 estimators $t_{1}, t_{2}$ and $t_{3}$ are asymptotically more efficient than $\bar{y}$ and $\bar{y}_{r}$. The efficiencies $E_{1} \& E_{2}$ of the estimators $t_{1}, t_{2}$ and $t_{3}$ over $\bar{y}$ and $\bar{y}_{r}$ will depend on $\rho, K$ and the weight $W$. The numerical values of $E_{1}$ and $E_{2}$ for different values of $\rho, K$ and for
$W=\frac{1}{4}$ and $W=\frac{1}{2}$ are given as percentages in Tables 1 and 2 respectively. Comparing the results in the two tables we may conclude that if a good guess of $\rho / K$ is not available from a pilot sample survey, past data or experience (1) $W=\frac{1}{4}$ appears to be a good overall choice ${ }^{\text {for }}$ $t_{1}, t_{2}$ and $t_{3}$ for low correlation (. $2<p<4$ ) and/or $K>1$. (2) $W=\frac{1}{2}$ appears to be a good choice for moderate to high correlation ( $\rho>.4$ ) and $K>1$. (3) In cases where $p>.8$ and $K \leqslant 1$ it is preferable to use $\bar{y}_{r}$. The asymptotic variance given in (2.9) of the estimators $t_{1}, t_{2}$ and $t_{3}$ with optimum value of $W=\rho / K$ is equal to the asymptotic variance of the linear regression estimator

$$
\begin{equation*}
\bar{y}_{l r}=\bar{y}+b(\bar{X}-\bar{x}) \tag{2.14}
\end{equation*}
$$

where $b$ is the sample regression coefficient. Thus these estimators with constant weights ( $W=\frac{1}{4}$ or $\frac{1}{2}$ ) are asymptotically no more efficient than $\bar{y}_{l r}$. However, if the regression of $y$ on $x$ is not linear, Cochran (1963) has shown that the bias in $\bar{y}_{l r}$ is of order $n^{-1}$ and hence it is more biased than $t_{2}$ whose bias is of order $n^{-2}$. Thus $t_{2}$ may be preferable to $\bar{y}_{i r}$ in situations where freedom from bias is important. Moreover, computationally $t_{2}$ is simpler than $\bar{y}_{l r}$.

## 3. The Exact Theory

We assume the following model for the comparison of estimators :

$$
\begin{aligned}
y_{i} & =\alpha+\beta x_{i}+u_{i} ; \beta>0 \\
E\left(u_{i} \mid x_{i}\right) & =0, E\left(u_{i}, u_{j} \mid x_{i}, x_{j}\right)=0 \\
V\left(u_{i} \mid x_{i}\right) & =n \delta\left(\delta \text { is a constant of order } n^{-1}\right)(\mathbf{I})
\end{aligned}
$$

where the variates $x_{i} / n$ have the gamma distribution with parameter $h$ so that $\bar{x}=\Sigma x_{i} / n$ has the gamma distribution with the parameter $m=n h$. This model was used by Durbin (1959), and Rao and Webster (1966) to investigate the bias in estimation of ratios, and Chakrabarty and Rao (1967) to investigate the stability of the 'JackKnife' variance estimator in ratio estimation. Chakrabarty (1973) has used this model to investigate the exact efficiency of the ratio estimator $\bar{y}_{r}$ and stability of the variance estimator of $\bar{y}_{r}$ relative to that of $\bar{y}$. He has shown that for $\rho \geqslant .4$ and $K<2 \rho$ the ratio estimator is generally more efficient than the unbiased estimator $\bar{y}$ even in small samples, and that the variance estimator of the ratio estimator is generally more stable than the variance estimator of $\bar{y}$. It may be noted that all our results under this model are exact for any sample size, $n$.

### 3.1 The exact biases of the estimators

In terms of the model (1) we have

$$
\bar{y}=\alpha+\beta \bar{x}+\bar{u}
$$

$$
\begin{align*}
E(\bar{y}) & =\alpha+\beta m=\bar{x} \\
t_{1} & =\alpha\left(1-W+\frac{W m}{\bar{x}}\right)+\beta[(1-W) \bar{x}+W m] \\
& +\bar{u}\left\{(1-W)+\frac{W m}{\bar{x}}\right\} \tag{3.1}
\end{align*}
$$

Consequently, the bias of $t_{1}$ is

$$
\begin{align*}
\operatorname{Bias}\left(t_{1}\right) & =E\left(t_{1}\right)-(\alpha+\beta m) \\
& =\alpha W /(m-1)  \tag{3.2}\\
t_{2} & =\alpha\left[(1-W)+W m\left(\frac{2}{\bar{x}}-\frac{1}{2 \bar{x}_{1}}-\frac{1}{2 \bar{x}_{2}}\right)\right] \\
& +\beta[(1-\mathrm{W}) \bar{x}-W m]-\frac{W m}{2}\left(\frac{\bar{u}_{1}}{\bar{x}_{1}}+\frac{\bar{u}_{2}}{\bar{x}_{2}}\right) \\
& +\bar{u}\left[(1-W)+\frac{2 W m}{\bar{x}}\right] \\
E\left(t_{2}\right) & =\beta m+\alpha[1-2 W /(m-1)(m-2)]
\end{align*}
$$

Thus the bias of $t_{2}$ is

$$
\begin{align*}
\operatorname{Bias}\left(t_{2}\right) & =-2 W \alpha /\{m-1)(m-2)  \tag{3.3}\\
t_{3} & =(\alpha+\beta \bar{x}+\bar{u}) m^{W} \bar{x}^{-W} \\
E\left(t_{3}\right) & =\frac{m^{W}}{\Gamma(m)}[\alpha \Gamma(m-W)+\beta \Gamma(m-W+1)]
\end{align*}
$$

Consequently, the bias of $t_{3}$ is

$$
\begin{align*}
\operatorname{Bias}\left(t_{3}\right) & =\alpha\left[\frac{m^{W} \Gamma(m-W)}{\Gamma(m)}-1\right] \\
& +\beta\left[\frac{m^{W} \Gamma(m-W+1)}{\Gamma(m)}-m\right] \tag{3.4}
\end{align*}
$$

Now. putting, $W=1$ in either (3.2) or (3.4) we get the bias of $\bar{y}_{r}$ as

$$
\begin{equation*}
\operatorname{Bias}\left(y_{r}\right)=\alpha /(m-1) \tag{3.5}
\end{equation*}
$$

From (3.2) through (3.5) it can be seen that the bias of $t_{2}$ is of order $n^{-2}$ while those of $\bar{y}_{r}, t_{1}$ and $t_{3}$ are of order $n^{-1}$ since $m=n h$ in our model. Also, the bias of $t_{1}$ is less than the bias of $\bar{y}_{r}$ if $W<1$. Further, for the special case of the linear regression through the origin (i.e. $\alpha=0$ in model I) the estimators $\bar{y} r, t_{1}$ and $t_{2}$ are unbiased but $t_{3}$ is still biased. A numerical evaluation of the biases of these estimators is given in the next section.

### 3.2 The exact variances of the estimators

The method of obtaining exact expressions for the variances of these estimators under model I is similar to that of Rao and Webster (1966). The details of evaluating these variances, which involve
some algebra, are omitted and only the final results are given here. The variance of $t_{1}$ can be shown to be

$$
\begin{align*}
V\left(t_{1}\right) & =\frac{W^{2} m^{2}}{(m-1)^{2}(m-2)} \alpha^{2}+(1-W)^{2} m \beta^{2} \\
& +\left[\frac{W^{2} m^{2}}{(m-1)(m-2)}+\frac{W(1-W)(m+1)}{(m-1)}+(1-W)\right] \delta \\
& -\frac{2 W(1-W) m}{(m-1)} \alpha \beta \tag{3.6}
\end{align*}
$$

Putting $W=1$ and $W \circ 0$ in (3.6) the variance of $\bar{y}_{r}$ and $\bar{y}$ are obtained as :

$$
\begin{equation*}
V\left(\bar{y}_{r}\right)=\frac{m^{2} \alpha^{2}}{(m-1)^{2}(m-2)}+\frac{m^{2} \delta}{(m-1)(m-2)} \tag{3.7}
\end{equation*}
$$

and

$$
\begin{equation*}
V(\bar{y})=\delta+\beta^{2} m \tag{3.8}
\end{equation*}
$$

respectively. The variance of $t_{2}$ is obtained as

$$
\begin{align*}
V\left(t_{2}\right) & =\frac{W^{2} m^{2}\left(m^{2}-6 m+17\right)}{(m-1)^{2}(m-2)^{2}(m-4)} \alpha^{2} \\
& -\frac{2 W(1-W) m(m-3)}{(m-1)(m-2)} \alpha \beta+(1-W)^{2} m \beta^{2} \\
& +\left[(1-W)^{2}+\frac{W^{2}\left(m^{2}-7 m+18\right) m^{2}}{(m-1)(m-2)^{2}(m-4)}\right. \\
& \left.+\frac{2 W(1-W) m(m-3)}{(m-1)(m-2)}\right] \delta \tag{3.9}
\end{align*}
$$

Finally, the variance of $t_{3}$ is given by

$$
\begin{align*}
{\left[m^{-2 W} \Gamma^{2}(m)\right] V\left(t_{3}\right) } & =\left[\Gamma(m-2 W) \Gamma(m)-\Gamma^{2}(m-W)\right] \dot{\alpha}^{2} \\
& +\left[\Gamma(m+2-2 W) \Gamma(m)-\Gamma^{2}(m+1-W)\right] \beta^{2} \\
& +2[\Gamma(m+1-2 W) \Gamma(m)-\Gamma(m-1-W) \Gamma(m-W)] \alpha \beta \\
& +[\Gamma(m-2 W) \Gamma(m)] \delta \tag{3.10}
\end{align*}
$$

We note that in terms of the model I

$$
\begin{align*}
& \alpha=\bar{Y}[(K-\rho) / K] \\
& \beta=\bar{Y}[\rho /(K m)] \\
& \delta=\bar{\gamma}^{2}\left[\left(1-\rho^{2}\right) /\left(K^{2} m\right)\right] \tag{3.11}
\end{align*}
$$

and $K=C_{x} / C_{y}$
The exact efficiencies of $\bar{y}_{r}$ and $t_{i}(i=1,2$ and 3 ), relative to that of $\bar{y}$ are given by

$$
\begin{align*}
& E_{r}^{\prime}=V(\overline{\mathrm{y}}) / \operatorname{MSE}\left(\overline{\mathrm{y}}_{\mathrm{r}}\right) \\
& E_{t}^{\prime}=V(\overline{\mathrm{y}}) / \operatorname{MSE}\left(t_{t}\right) \quad i=1,2 \& 3 \tag{3.12}
\end{align*}
$$

Now, using (3.2) through (3.10) and substituting the values of $a, \beta$ and $\delta$ given by (3.11) efficiencies $E_{r}^{\prime}$ and $E_{i}^{\prime}(i=1,2 \& 3)$ can be expressed explicitly as functions of $K=C_{x} / C_{y}, m=n h, \rho$ and weight $W$. However, it is difficult to investigate analytically the efficiencies of the estimators from the resulting expressions.
Therefore, we have evaluated the values of $E_{r}^{\prime}$ and $E_{i}^{\prime}$ (percentages) for selected values of $\rho, K$ and $m$ and for $W=\frac{1}{4}$ and $\frac{1}{2}$. The results are given in Tables 3 and 4 respectively. The results of Table 3 may be summarized as follows: (1) The ratio estimator $\bar{y}_{r}$ is less efficient than $\bar{y}$ for low correlation ( $\rho \leqslant 4$ ) except when $\rho=.4, K<1$ and $m \geqslant 20$. (2) The estimators $t_{1}, t_{2}$ and $t_{3}$ with $W=\frac{1}{4}$ are more efficient than both $\overline{\mathrm{y}}$ and $\overline{\mathrm{y}} r$ for the following values of $\rho, K$ and $m,(a) .2 \leqslant \rho \leqslant 4, K \leqslant 1, m \geqslant 16$. (b) $.2<\rho<.4, K>1$, $m \geqslant 32$. Noting that in our model $C_{x}=h^{-1 / 2} C \bar{x}=m^{-1 / 2}$ and $n \leqslant m$ if $h \geqslant 1$ we may conclude that for low correlation $(.2 \leqslant \rho \leqslant .4), W=\frac{1}{4}$ appears to be a good choice for estimators $t_{1}, t_{2} \& t_{3}$ even in small samples if $K \leqslant 1$ and in large samples only when $K>1$. Further, the exact efficiencies of these estimators with $W=\frac{1}{4}$ are of the same order as judged by their mean square errors.

From table 4, it can be seen that the estimators $t_{1}, t_{2}$ and $t_{3}$ with $W=\frac{1}{2}$ are more efficient than both $\bar{y}$ and $\overline{\mathrm{y}} r$ for $\rho \geqslant .5$, $.25 \leqslant K \leqslant 1.50$ and $m \geqslant 16$. However, the ratio estimator $\overrightarrow{\mathrm{y}}_{r}$ is most efficient when $\rho=.9$ and $.5 \leqslant K \leqslant 1$. Thus, $W=\frac{1}{2}$ appears to be a good choice for estimators $t_{1}, t_{2}$ and $t_{3}$ for moderate to high correlation ( $\rho>.4$ ), except when $\rho=.9$ and $.5 \leqslant K \leqslant 1$. The exact efficiencies of $t_{1}, t_{2}$ and $t_{3}$ with $W=\frac{1}{2}$ are again generally of the same order. It is interesting to note that under model I the exact efficiencies of the estimators $t_{1}, \mathfrak{t}_{2}$ and $t_{3}$ approach the asymptotic efficiency when $m=n h \geqslant 32$. For example when $\rho=.4 \& K=1.0, E_{1}=116$ (table 1) $\& E_{1}^{\prime}=114, E_{2}^{\prime}=E_{3}^{\prime}=115$ for $m=32$ (table 3 ).

We note from tables 3 and 4 that it is difficult to choose among the estimators $t_{1}, t_{2}$ and $t_{3}$ on the basis of their exact mean square errors. The absolute biases of estimators $\bar{y}_{r}$ and $t_{i}$ relative to their mean square errors are given by
and

$$
B_{r}=\left|\operatorname{Bias}\left(\bar{y}_{r}\right)\right| /\left[\operatorname{MSE}\left(\bar{y}_{r}\right)\right]^{1 / 2}
$$

$$
\begin{equation*}
B_{i}=\left|\operatorname{Bias}\left(t_{i}\right)\right| /\left[\operatorname{MSE}\left(t_{i}\right)\right]^{1 / 2}, i=1,2 \& 3 \tag{3.13}
\end{equation*}
$$

respectively. The numerical values of $B_{r}$ and $B_{i}(i=1,2 \& 3)$ for $W=\frac{1}{4}$ and $W=\frac{1}{2}$ are given in tables 5 and 6 respectively for selected values of $m, K \& \rho$. From table 5, it can be seen that $B_{2}$ is generally less than $1 \% ; B_{1}$ is slightly greater than $B_{3}$ but $B_{1}$ is still less than
$10 \%$ for $m=n h \geqslant 16$. The ratio estimator $\bar{y}_{r}$ is generally badly biased ( $B_{r}>10 \%$ for $K \geqslant 1$ ). From table 6, we find that $B_{2}<1 \%$ for $K \leqslant 1$ and for $K>1, B_{2}<2.5 \%$ when $m \geqslant 16$. Turning to the relative biases of $t_{1}$ and $t_{3}$ we find that $B_{1}<B_{3}$ for $K<1$ and $B_{1}>B_{3}$ for $K>1$. It is also interesting to note that although $\operatorname{MSE}\left(\bar{y}_{r}\right)<\operatorname{MSE}\left(t_{i}\right)$ for $\rho=.9$ and $. t \leqslant K \leqslant 1$ (table 4), $B_{r}$ in this case exceeds $10 \%$ and is considerably higher than $B_{i}$. Thus, for $\rho=.9$ and $.5 \leqslant K \leqslant 1$, although $\operatorname{MSE}\left(\overline{\mathrm{y}}_{r}\right)<\operatorname{MSE}\left(t_{i}\right)$, the estimators $i_{i}$ 's may be preferable in situations where the freedom from bias is desirable.

It may be noted that in surveys with many strata and small samples within strata the bias of the ratio estimator relative to its standard error may be considerable if it is appropriate to use 'separate' ratio estimators [see Cochran (1963)]. In such situations it may be of great advantage to use the proposed estimators $t_{i}(i=1,2$ and 3 ). These estimators not only reduce the bias but also increase the precision.

In light of the above results we conclude that the three ratiotype estimators $t_{1}, t_{2}$ and $t_{3}$ are preferable to both $\bar{y}$ and $\bar{y}_{r}$. The efficiencies of these estimators are the same in large samples and are practically of the same order in small samples. Computationally $t_{1}$ is simplest and the bias of $t_{2}$ is least.

The author wishes to thank Dr. J.N.K. Rao for his valuable suggestions.

## References

[1] Chakrabarty, R.P. and : The Bias and Stability of Jak-Knife Variance Rao, J.N.K. (1967)
[2] Chakrabarty, R.P. (1968) : Contributions to the Theory of Ratio-Type Estimators, Ph.D. Thesis, Texas A \& M Univ.
[3] Chakrabarty, R.P. (1973) : A note on the Small Sample Theory of the Ratio Estimator in Certain Specified Populations. Jour. Ind. Soc. Agr. Stat., 25, 49-57.
[4] Cochran, W.G. (1963) : Sampling Techniques, John Wiley \& Sons, Inc., New York.
[5] Durbin, J. (1959) : A note on the application of Quenouille's Method of Bias Reduction in Estimation of Ratios. Biometrika, 46, 477-80.
[6] Hartley, H.O. and : Unbiased Ratio Ẹstimates, Nature, 174, $270-71$. Ross, A. (1954)
[7] Rao, J.N.K. and Webster, : On two Methods of Bias Reduction in EstiJ.T. (1966) mation of Ratios. Biometrika, 53, 571-77.
[8] Srivastava, S.K. (1967) : An Estimator Using Auxiliary Information in Sample Surveys. Calcutta. Stat. Assoc. Bulletin,

TABLE 1
Efficiencies, $E_{1}$ and $E_{2}$, of $t_{1}$, $\left(t_{2}\right.$ and $\left.t_{3}\right)$ over $\bar{y}$ and $\bar{y}_{r}$ for selected values of $\rho$ and $K$ and $W=1 / 4$

| $p$ | $K=0.5$ |  | $K=1.0$ |  | $K=1.5$ |  | $K=2.0$ |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $E_{1}$ | $E_{2}$ | $E_{1}$ | $E_{2}$ | $E_{1}$ | $E_{2}$ | $E_{3}$ | $E_{2}$ |
| . 1 | 101 | 116 | 99 | 178 | 94 | 277 | 89 | 400 |
| . 2 | 104 | 109 | 104 | 166 | 101 | 268 | 95 | 400 |
| . 3 | 106 | 101 | 110 | 153 | 109 | 257 | 105 | 400 |
| . 4 | 109 | 93 | 116 | 139 | 119 | 244 | 118 | 400 |
| . 5 | 112 | 84 | 123 | 123 | 131 | 229 | 133 | 400 |
| . 6 | 116 | 75 | 131 | 105 | 145 | 210 | 154 | 400 |
| . 7 | 119 | 65 | 140 | 84 | 162 | 187 | 182 | 400 |
| . 8 | 123 | 55 | 150 | 63 | 185 | 157 | 222 | 400 |
| . 9 | 126 | 44 | 163 | 33 | 215 | 118 | 285 | 400 |

TABLE 2
Efficiencies, $E_{1}$ and $E_{2}$, of $t_{1},\left(t_{2}\right.$ and $\left.t_{s}\right)$ over $\bar{y}$ and $\bar{y}_{r}$ for selected values of $\rho$ and $K$ and $W=1 / 2$

| $\rho$ | $K=0.5$ |  | $K=1.0$ |  | $K=1.5$ |  | $K=2.0$ |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $E_{1}$ | $E_{2}$. | $E_{1}$ | $E_{2}$ | ${ }^{\prime} E_{1}$ | $E_{2}$ | $E_{1}$ | $E_{2}$ |
| . 1 | 99 | 114 | 87 | 157 | 71 | 209 | 56 | 256 |
| . 2 | 104 | 109 | 95 | 152 | 79 | 210 | 62 | 262 |
| . 3 | 110 | 104 | 105 | 147 | 90 | 211 | 71 | 271 |
| . 4 | 116 | 99 | 117 | 141 | 104 | 213 | 83 | 283 |
| . 5 | 123 | 92 | . 133 | 133 | 123 | 215 | 100 | 300 |
| . 6 | 131 | 85 | 153. | 123 | 151 | 219 | 125 | 325 |
| . 7 | 140 | 77 | 182 | 109 | 195 | 224 | 167 | 367 |
| . 8 | 150 | 68 | 222 | 89 | 276 | 234 | 250 | 450 |
| . 9 | 163 | 57 | 286 | 57 | 471 | 259 | 500 | 700 |

table 3
The exact efficiencies，$E_{r}^{\prime}$ and $E_{i}^{\prime}$ ，of $\bar{y}_{r}$ and $t_{i}(i \square 1,2 \& 3)$ with $W=1 / 4$ ，for selected values of $m, K \& \rho$

|  | $\mathrm{in}^{0}$ | $\stackrel{\sim}{\square} \rightrightarrows$ |  | $\stackrel{\rightharpoonup}{-} \pm$ | $\stackrel{\sim}{\square} \cong$ |
| :---: | :---: | :---: | :---: | :---: | :---: |
| ＋ | 卥 | 흘 을 | $\stackrel{\otimes}{\square} \pm \pm$ | $\stackrel{\square}{\square} \pm \underline{ }$ | $\stackrel{\square}{\square} \because$ |
|  |  | 흥 | $\stackrel{\text { ® }}{\text { ® }}$ | $\stackrel{\square}{\square} \cong$ | $\stackrel{\square}{\square} \pm$ |
|  | 施 | $\hat{\sim} \stackrel{\sim}{\sim}$ | \％\％is | 8 잉 | 「 |
| $\begin{aligned} & \text { mi } \\ & i \end{aligned}$ | 閶 | 응 흘 | 扣点 | 扣 | nัٌ |
|  | 玉゙ | 兌 응 | 등 | $\stackrel{\square}{\square} \stackrel{0}{\square}$ | 은 |
|  | 囪 | 毋 \％\％ |  | 음 응 | 흥 흥． |
|  | 交 | \％\％ | ®ロ～0 | Q in | そ ¢－ |
| N | isi | ภッ \％ | 훙 응 | 끙 응 | 흥 |
|  | －${ }^{\text {a }}$ |  | 웅 웅 | 응 응 | ๕\％ |
|  | 可 | ふু ¢ | 응 $\mathrm{g}^{\text {g }}$ | 응 용 | 응 ${ }_{\text {O }}^{\circ}$ |
|  | 㐫 | ठ हो | ह \％${ }^{\text {a }}$ | あ ज | $\infty$ in |
| $z$ |  | $\bigcirc 88$ | $\bigcirc \stackrel{8}{9}$ | 웅 | $\bigcirc$ |
| E |  | $\infty$ | $\stackrel{\square}{\square}$ | \％ | N |
|  |  |  |  |  |  |

TABLE 4
The exact efficiencies, $E_{r}^{\prime}$ and $E_{i}^{\prime}$, of $y_{r}$, and $t_{i}(i=1,2 \& 3)$ with $W=1 / 2$, for selected values of $m, K \& \rho$

| $m$. | $K$ | $p=.5$ |  |  |  | $p=.7$ |  |  |  | $p=.9$ |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | $E_{r}^{\prime}$ | $E_{1}^{\prime}$ | $E_{2}^{\prime}$ | $E_{3}^{\prime}$ | $E_{r}^{\prime}$ | $E_{1}^{\prime}$ | $E_{2}^{\prime}$ | $E_{3}^{\prime}$ | $E_{r}^{\prime}$ | $E_{1}^{\prime}$ | $E_{2}^{\prime}$ | $E_{3}^{\prime}$ |
| 8 | . 25 | 7987 | $\begin{array}{r} 94 \\ 104 \end{array}$ | 99109 | 99109 | 86117 | 99120 | $\begin{aligned} & 103 \\ & 126 \end{aligned}$ | $\begin{aligned} & 105 \\ & 126 \end{aligned}$ | 91168 | 103 | 105146 | 1.13149 |
|  | . 50 |  |  |  |  |  |  |  |  |  |  |  |  |
|  | 1.00 |  |  |  |  |  |  |  |  |  | 140 |  |  |
|  | 1.50 | 62 33 | 106 | 105 | 116 | 105 |  |  |  |  |  |  |  |
|  |  |  | 87 | 78 | 102 | - 50 | 139 | 122 | 162 163 | 324 | 260 | 262 | 269 |
| 16 |  | 100109 | $\begin{aligned} & 103 \\ & 114 \end{aligned}$ | $\begin{aligned} & 108 \\ & 119 \end{aligned}$ | $\begin{aligned} & 106 \\ & 116 \end{aligned}$ | 111 | $\begin{aligned} & 109 \\ & 130 \end{aligned}$ | 114136 | 112113 | $\begin{aligned} & 123 \\ & 222 \end{aligned}$ | $\begin{aligned} & 115 \\ & 152 \end{aligned}$ | 120157 |  |
|  | . 25 |  |  |  |  |  |  |  |  |  |  |  | 120156 |
|  | . 50 |  |  |  |  |  |  |  |  |  |  |  |  |
|  | 1.00 | $\begin{aligned} & 80 \\ & 44 \end{aligned}$ | $\begin{aligned} & 120 \\ & 105 \end{aligned}$ | $\begin{aligned} & 124 \\ & 107 \end{aligned}$ | $\begin{aligned} & 124 \\ & 112 \end{aligned}$ |  |  |  |  |  |  |  |  |
|  | 1.50 |  |  |  |  | $\begin{array}{r} 134 \\ 67 \end{array}$ | $\begin{aligned} & 168 \\ & 167 \end{aligned}$ | $\begin{aligned} & 173 \\ & 168 \end{aligned}$ | $\begin{aligned} & 172 \\ & 179 \end{aligned}$ | $\begin{aligned} & 408 \\ & 139 \end{aligned}$ | $\begin{aligned} & 274 \\ & 409 \end{aligned}$ | $\begin{aligned} & 279 \\ & 391 \end{aligned}$ | $\begin{array}{r} 278 \\ 444 \end{array}$ |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 20 | . 25 | 104114 | $\begin{aligned} & 105 \\ & 116 \end{aligned}$ | $\begin{aligned} & 109 \\ & 120 \end{aligned}$ | $\begin{aligned} & 107 \\ & 117 \end{aligned}$ | 117154 | 111133 | $\begin{aligned} & 115 \\ & 137 \end{aligned}$ | 114135 | $\begin{aligned} & 130 \\ & 234 \end{aligned}$ | 118155 | 121159 | 121 |
|  | . 50 |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | $\begin{aligned} & 83 \\ & 46 \end{aligned}$ | $\begin{aligned} & 123 \\ & 108 \end{aligned}$ |  |  |  |  |  |  |  |  |  |  |
|  | 1.00 1.50 |  |  | $\begin{aligned} & 127 \\ & 111 \end{aligned}$ | 126 | 14070 | $\begin{aligned} & 171 \\ & 173 \end{aligned}$ | $\begin{aligned} & 175 \\ & 175 \end{aligned}$ | $\begin{aligned} & 174 \\ & 182 \end{aligned}$ | 425146 | $\begin{array}{r} 277 \\ 422 \end{array}$ | $\begin{aligned} & 280 \\ & 410 \end{aligned}$ | 279450 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 32 | . 25 | $\begin{aligned} & 111 \\ & 121 \end{aligned}$ | 108119 | 110121 | 109120 | 125164 | $\begin{aligned} & 114 \\ & 136 \end{aligned}$ | 117138 | 116138 | 142 | 121158 | 124161 | 123160 |
|  | . 50 |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | 8950 |  |  | $\begin{aligned} & 129 \\ & 118 \end{aligned}$ |  |  |  |  |  |  |  |  |
|  | 1.50 |  | $\begin{aligned} & 127 \\ & 114 \end{aligned}$ | 129116 |  | 15076 | $\begin{aligned} & 175 \\ & 181 \end{aligned}$ | $\begin{aligned} & 177 \\ & 183 \end{aligned}$ | $\begin{aligned} & 177 \\ & 187 \end{aligned}$ | $\begin{aligned} & 453 \\ & 159 \end{aligned}$ | $\begin{array}{r} 280 \\ 441 \end{array}$ | $\begin{aligned} & 283 \\ & 435 \end{aligned}$ | $\begin{aligned} & 282 \\ & 458 \end{aligned}$ |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |

TABLE 5
The absolute values of $\% \mathrm{Bias} /(M S E)^{1 / 2}, B_{r}$ and $B_{i}$ of $\bar{y}_{r}$ and $t_{i}(i=1,2 \& 3)$ with $W=1 / 4$, for selected values of $m, K \& \rho$

| $m$ | $K$ | $p=.2$ |  |  |  | $p=.3$ |  |  |  | $\rho=.4$ |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | $\boldsymbol{B r}_{\boldsymbol{r}}$ | $B_{1}$ | $\boldsymbol{B}_{2}$ | $B_{3}$ | $\boldsymbol{B}_{r}$ | $B_{1}$ | $B_{2}$ | $B_{3}$ | $B_{r}$ | $B_{1}$ | $B_{2}$ | $B_{3}$ |
| 8 | . 50 | 9.48 | 2.96 | 1.01 | 1.07 | 6.68 | 2.00 | . 69 | . 18 | 3.55 | 1.02 | . 35 | . 74 |
|  | 1.00 | 19.74 | 7.80 | 2.64 | 3.92 | 18.57 | 7.05 | 2.38 | 3.12 | 17.29 | 6.24 | 2.11 | 2.26 |
|  | 1.50 | 25.02 | 12.31 | 4.06 | 6.65 | 24.57 | 11.88 | 3.91 | 6.01 | 24.15 | 11.42 | 3.78 | 5.33 |
| 16 | . 50 | 7.03 | 2.00 | . 29 | . 74 | 4.94 | 1.35 | . 20 | . 10 | 2.62 | . 68 | . 10 | . 55 |
|  | 1.00 | 14.90 | 5.31 | . 77 | 2.74 | 13.99 | 4.78 | . 69 | 2.16 | 12.99 | 4.22 | . 61 | 1.56 |
|  | 1.50 | 18.56 | 8.42 | 1.22 | 4.67 | 18.22 | 8.10 | 1.17 | 4.21 | 17.91 | 7.77 | 1.12 | 3.72 |
| 20 | . 50 | 6.34 | 1.77 | . 20 | . 65 | 4.46 | 1.20 | . 13 | . 09 | 2.40 | . 61 | . 07 | . 50 |
|  | 1.00 | 13.50 | 4.71 | . 53 | 2.45 | 12.66 | 4.42 | . $48{ }^{\circ}$ | 1.93 | 11.77 | 3.74 | . 42 | 1.39 |
|  | 1.50 | 16.85 | 7.48 | . 84 | 4.17 | 16.54 | 7.12 | . 81 | 3.76 | 16.25 | 6.90 | . 77 | 3.32 |
| 32 | . 50 | 5.08 | 1.38 | . 09 | . 51 | 3.56 | . 93 | . 06 | . 06 | 1.88 | . 47 | . 03 | . 40 |
|  | 100 | 10.86 | 3.68 | . 25 | 1.93 | 10.18 | 3.31 | . 22 | 1.52 | 943 | 2.92 | . 20 | 1.90 |
|  | 1.50 | 13.62 | 5.86 | . 39 | 3.29 | 13.36 | 5.64 | . 38 | 2.96 | 13.12 | 5.40 | . 36 | 261 |

SOMB RATIO-TYPE ESTIMATORṠ $\quad \dot{1}$
TABLE 6
The absolute values of $\% \operatorname{Bias} /(M S E)^{1 / 2}, B_{r}$ and $B_{i}$ of $\bar{y}_{r}$ and $t_{i}(i=1,2, \& 3) W ø 1 / 2$, for selected values of $m, K$ and $\rho$

| $m$ | $K$ | $p=.5$ |  |  |  | $p=.7$ |  |  |  | $p=.9$ |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | $B_{r}$ | $3_{1}$ | $B_{2}$ | $B_{3}$ | $B_{r}$ | $B_{1}$ | $B_{2}$ | $B_{3}$ | $B_{r}$ | $B_{1}$ | $B_{2}$ | $B_{3}$ |
| 8 | . 25 | 8.99 | 4.89 | 1.67 | 5.71 | 16.86 | 9.04 | 3.09 | 9.71 | 25.11 | 13.36 | 4.49 | 13.98 |
|  | . 50 | 0.00 | 0.00 | 0.00 | 2.29 | 8.75 | 4.42 | 1.51 | 6.62 | 20.97 | 9.57 | - 3.25 | 11.73 |
|  | 1.00 | 15.89 | 10.42 | 3.46 | 5.27 | 12.45 | 7.47 | 2.49 | 1.51 | 7.27 | 3.26 | 1.09 | 4.14 |
|  | 1.50 | 23.15 | 18.80 | 5.95 | 12.10 | 22.88 | 19.04 | 5.94 | 10.58 | 24.55 | 22.36 | . 6.56 | 9.29 |
| 16 | . 25 | 6.67 | 3.39 | . 50 | 4.09 | 12.65 | 6.27 | . 91 | 6.92 | 19.25 | 9.31 | 1.36 | 9.96 |
|  | . 50 | 0.00 | 0.00 | 0.00 | 1.68 | 6.48 | 3.05 | . 44 | 4.75 | 15.88 | 6.58 | . 96 | 8.34 |
|  | 1.00 | 11.89 | 7.31 | 1.06 | 3.67 | 9.26 | 5.18 | . 75 | . 96 | 5.38 | 2.21 | . 32 | 3.05 |
|  | 1.50 | 17.64 | 13.65 | 1.96 | 8.62 | -17.42 | 13.79 | 1.98 | 7.46 | 18.79 | 16.17 | 2.26 | 6.35 |
| 20 | . 25 | 6.01 | 3.02 | . 34 | 3.67 | 11.44 | 5.59 | . 63 | 6.21 | 17.48 | 8.30 | . 94 | 8.92 |
|  | . 50 | 0.00 | 0.00 | 0.00 | 1.51 | 5.85 | 2.71 | . 31 | 4.26 | 14.39 | 5.85 | . 66 | 7.47 |
|  | 1.00 | 10.75 | 6.52 | . 74 | 3.27 | 8.37 | 4.61 | . 52 | . 83 | 4.85 | . 1.96 | . 22 | 2.75 |
|  | 1.50 | 16.00 | 12.25 | 1.38 | 7.71 | 15.80 | 12.38 | 1.38 | 6.67 | 17.06 | 14.50 | 1.59 | 5.64 |
| 32 | . 25 | 4.81. | 2.37 | . 16 | 2.91 | 9.19 | 4.39 | . 30 | 4.92 | 14.14 | 6.53 | . 44 | 7.06 |
|  | . 50 | 0.00 | 0.00 | 0.00 | 1.21 | 4.68 | 2.12 | . 14 | 3.38 | 11.59 | 4.59 | . 31 | 5.92 |
|  | 1.00 | 8.63 | 5.14 | . 35 | 2.57 | 6.70 | 3.62 | . 24 | . 64 | 3.88 | 1.53 | . 10 | 2.20 |
|  | 1.50 | 12.92 | 9.74 | . 65 | 6.11 | 12.75 | 9.83 | . 66 | 5.26 | 13.79 | 11.49 | . 76 | 4.40 |


[^0]:    *Revised version of the paper presented at the annual meeting of the American Statistical Association hẹld at New Yoṛk ị Dẹcember, 1973

