Modification of Hansen-Hurwitz's Estimators for Negatively Correlated Variates

Jonathan A. Ikughur[1]

Most estimators in probability proportional to size (PPS) with replacement (WR) sampling design are defined for positively correlated variables. The literature on estimators involving negatively correlated variates is underdeveloped. This study utilized the law of inverse proportion to derive selection probabilities to modify the Hansen-Hurwitz's estimator for use when negatively correlated variables are encountered in surveys. The suggested estimator is unbiased and could perform better than other estimators when information conveyed by correlation coefficient is used.

Keywords: Negative correlation, multi-character survey, inverse transformation, PPS.

JEL Classification:

1.0 Introduction

Large scale surveys often utilize multiple auxiliary characteristics at both sampling design and estimation stages. In developing an estimator for survey data, certain statistical properties are usually exploited in order to increase precision and one of these characteristics is the population correlation coefficient which is usually unknown but is estimated from the sampled data.

It is well known that population correlation coefficient could assume any one of the three dimensions namely, positive, negative and zero correlation and at varying proportions in the case of positive and negative correlations. Thus, when survey data involving study variable and auxiliary characteristics are negatively correlated, and when it is desirable to utilize the normed-size measure in defining an estimator, existing conventional[2] estimators fail to perform effectively. In this case, new estimators are devised or existing estimators are modified to cater for such scenario.

Basically, there exist estimators in PPS with replacement sampling scheme to cater for situation when positive correlations are encountered in surveys. First of its kind is the one developed by Hansen and Hurwitz (1943) called the Hansen and Hurwitz's Estimator (HHE). Although this estimator was defined

[1] Department of Statistics, Federal University, Wukari.

[2] The conventional estimator usually specified when x and y are negatively correlated is the product estimator

by using the i^{th} selection probability, p_i of the i^{th} study variable y_i, it has been established in the works of Singh *et al.* (2004), Ikughur and Amahia (2011a,b) and Enang and Amahia (2012) among others that the HHE is a member in the generalized class of alternative estimators in PPS sampling. Specifically, under the transformed selection probability $p_i^* = \frac{1-\rho^c}{N} + \rho^c p_i$, c = 0,1,2,3,4 and $0 \leq \rho \leq 1$, HHE is defined for $\rho = 1$ and, or for c = 0. Similarly, when $\rho = 0$, then $p_i^* = \frac{1}{N}$ which defines Rao's (1966a,b) estimator. This transformation probability gives rise to a class of alternative estimators in PPS with replacement sampling scheme for varying degree of positive correlation including zero correlation.

However, there are scanty literatures concerning estimators in PPS sampling with replacement scheme for cases of negatively correlated variables. Importantly, these literatures do not take cognizance of selection probabilities (or transformed selection probabilities) and even when they do, they are complex in nature and applications. This study is set to provide alternative estimators for use in PPS sampling with replacement scheme when negative correlation is encountered in survey. To accomplish this, the Hansen and Hurwitz's estimator is modified by using the law of inverse proportions. The study further compare the proposed estimator with the estimator s developed by Sahoo *et al.* (1994) and Bedi and Rao (1997) for negatively correlated variables and also with other alternative estimators in the generalized class.

2.0 Literature Review

Hansen and Hurwitz (1943) proposed the idea of sampling with probability proportional to size with replacement for positively correlated characteristics. Under this scheme, one unit is selected at each of the n-draws. For each i^{th} unit selected from the population, a selection probability is given as

$$p_i = \frac{x_i}{X}, \qquad (1)$$

where x_i is the measure for i^{th} population unit and $X = \sum_i^N x_i$.

Using the notations defined above, Hansen and Hurwitz (1943) gave the estimators of population total Y, as

$$\hat{\tau}_{HH} = \frac{1}{n}\sum_{i=1}^{N}\frac{y_i}{p_i} \qquad (2)$$

whose design based variance of the estimator, $V_p(\hat{\tau}_{HH})$ is

$$V_p(\hat{\tau}_{HH}) = \frac{1}{n}\sum_{i=1}^{N}\frac{1}{p_i}(y_i - p_i y_i)^2 \qquad (3)$$

and the possible unbiased estimators of population variance are given as

$$V_p(\hat{\tau}_{HH}) = \frac{1}{n(n-1)}\sum_{i=1}^{n}\left(\frac{y_i}{p_i} - \hat{\tau}_{HH}\right)^2 \qquad (4)$$

and

$$V_p(\hat{\tau}_{HH}) = \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\left(\frac{y_i}{p_i} - \frac{y_j}{p_j}\right)^2 \qquad (5)$$

To enhance the efficiency and wider usage of the Hansen and Hurwitz estimator, Rao (1966a,1966b) introduced the idea of estimators under multiple characteristics utilizing zero correlation coefficient in defining the estimator of population total by modifying the Hansen and Hurwitz's estimator for use when the variables encountered have zero correlation. Works ascertaining the validity of this estimator was carried out by Pathak (1966), Rao (1993a, 1993b) among others who introduced a new dimension in the study of PPS sampling schemes leading to other specific estimators by Bansal and Singh (1985), Amahia *et al.* (1989), Grewal (1997) and also Enang and Amahia(2012) whose work basically covered the area of positive correlation between the study variable and measure of size variable. Specifically, the estimators are generalized by the transformed probability defined as

$$g_k(x) = p_{i,k}^{*} = \frac{1-\rho^c}{N} + \rho^c p_i,\ c = 0,1,...,4\ ;\ 0 \leq \rho \leq 1,\ k = 1,2,...,7 \qquad (6)$$

Where $g_k(x)$ and $p_{i,k}^{*}$ is the density function of the auxiliary variable x and generalized transformed probability for the k^{th} estimator respectively. Specifically, the following estimators can be defined from (6) above as:

$$\hat{\tau}_k = \frac{1}{n}\sum_{i=1}^{N}\frac{y_i}{g_k(x)},\ k = 1,2,3,4,5,6,7 \qquad (7)$$

i) $g_1(x) = p_{i,1}$ known as Hansen and Hurwitz estimator (HHE) when $\rho = 1$ or c = 0;

ii) $g_2(x) = p_{i,2}^{*} = \frac{1}{N}$ known as Rao's(1966a,b) estimator when $\rho = 0$;

iii) $g_3(x) = p^*_{i,3} = \frac{1-\rho^1}{N} + \rho^1 p_i$, known as Amahia, Chaubey and Rao(1989) Estimator for $\rho > 0$, c=1;

iv) $g_4(x) = p^*_{i,4} = \frac{1-\rho^c}{N} + \rho^c p_i$ known as Enang and Amahia(2011) estimator, $\rho > 0$ and $0 \leq c \leq 1$[3];

v) $g_5(x) = p^*_{i,5} = \frac{1-\rho^c}{N} + \rho^c p_i$ known as Ikughur and Amahia(2011) estimator for $\rho > 0$ and c =2;

vi) $g_6(x) = p^*_{i,6} = \frac{1-\rho^c}{N} + \rho^c p_i$ known as Ikughur and Amahia(2011) estimator for $\rho > 0$ and c =3;

vii) $g_7(x) = p^*_{i,7} = \frac{1-\rho^c}{N} + \rho^c p_i$ known as Ikughur and Amahia(2011) estimator for $\rho > 0$ and c =4;

Apart from the estimator defined by $g_1(x) = p_{i,1}$, all other estimators are biased with the design based bias defined as $B_p(\hat{\tau}_k) = \sum_{i=1}^{N}(\frac{p_i}{g_k(x)} - 1)y_i$, where $B_p(\hat{\tau}_k)$ is an expression of the design based bias of the estimator $\hat{\tau}_k$

2.1 Estimators in Negative Correlation

Following postulation by Sahoo *et al.* (1994) and also Bedi (1995), the work of Bedi and Rao (1997) gave a new direction in determining estimators of population total under the PPSWR sampling scheme especially when negatively correlated variates are encountered in surveys. Singh and Horn (1998) also proposed an alternative estimator for estimating population totals in multi-character survey sampling when certain variables have poor positive correlation and others have poor negative correlation with selection probabilities. The estimators have same definition as the estimator in (7) above except in the transformation probabilities that combines two forms of selection probabilities namely;
$p_i^+ = \frac{x_i}{X}; X = \sum_{i=1}^{N} x_i$ for cases of positively correlated variates and $p_i^- = \frac{z_i}{X}; X = \sum_{i=1}^{N} x_i$ for negatively correlated variates, with $z_i = \frac{X - nx_i}{N-n}; X = \sum_{i=1}^{N} x_i$. Specifically, Sahoo *et al.* (1994) suggested the transformation probability for the estimator given as

[3] The estimator defined by $g_4(x) = p^*_{i,4}$ include that of Grewal, Bansal and Singh(1989) for which c = 1/3

$$p_{i,8}^* = \frac{z_i}{X}, \quad z_i = \frac{X - nx_i}{N - n} \qquad (8)$$

Singh and Tailor (2003, 2005) suggested series of estimators of population totals under the following complex transformations of selection probabilities:

$$p_{i,9}^* = \left(1 + \frac{1}{N}\right)^{(1-\rho)(1+\rho)} (1 + p_i^+)^{\frac{\rho(1+\rho)}{2}} (1 + p_i^-)^{-\frac{\rho(1-\rho)}{2}}$$

$$\left(\frac{1}{N}\right)^{(1-\rho)(1+\rho)} - 1 \qquad (9)$$

$$p_{i,10}^* = \frac{(1-\rho)(1+\rho)}{N} + \frac{1}{2}[\rho(1+\rho)p_i^+ - \rho(1-\rho)p_i^-] \qquad (10)$$

Other estimators with such transformations include that suggested by Singh and Horn(1998) while Bedi and Rao(1997) suggested a transformation of the form

$$p_{i,11}^* = \frac{1 - p_i}{N - 1}, i = 1,2, \ldots, N \qquad (11)$$

Recently, Sahoo *et al.* (2012) suggested an estimator for PPSWR sampling using harmonic mean of the auxiliary variable without clearly defining the normed-size measure and hence, there was no clear application to PPS sampling scheme. In this study, a simple estimator is suggested for negatively correlated variates. This estimator utilizes a transformation that changes the correlation coefficient from negative to positive so that under this transformation, an estimator in the form of HHE can be utilized.

3.0 Suggested Estimator

Definition 1: Consider a finite population Ω of N identifiable units uniquely labeled 1,2, …, N on which are defined two real value variables y and x assuming y_i(>0) and x_i(>0). Let a sample of size n be drawn with replacement from Ω and we suppose that *y* and *x* are negatively correlated, then the following theorems are applicable:

Theorem 1: Let $y \propto 1/x$ *(or p)* such that y and *x* (or *p*) are negatively correlated, then the transformation for the selection probabilities required is

$$p_i' = \frac{z_i}{Z} \qquad (12)$$

where

$z_i = \frac{1}{x_i}$ and $Z = \sum_{i=1}^{N} z_i$.

Proof: y $\propto$ 1/x $\Rightarrow$ y $\propto$ z, where z = 1/x. Then, $p_i' = \frac{1/x_i}{\sum_{i=1}^{N} 1/x_i} \Rightarrow p_i' = \frac{z_i}{Z}$

Proposition 1: Under the transformation above, the nomenclature $p_i' = \frac{z_i}{Z}$ with $z_i = \frac{1}{x_i}$ is the selection probabilities realized for variables that are inversely proportional to each other. This transformation has the properties of harmonic mean. As a consequence of this transformation, the realized selection probabilities are directly proportional to Z, hence the corresponding estimator is called Modified Hansen and Hurwitz (MHH) Estimator.

Theorem 2: Let y and x be negatively correlated. If $y \propto z$ *(or* p_i'*)* can be defined as in theorem 1 above, then the estimator of population total which is the Modified Hansen and Hurwitz's estimator denoted by $\hat{\tau}_{MHH}$ is

$$\hat{\tau}_{MHH} = \frac{1}{n}\sum_{i=1}^{n} \frac{y_i}{p_i'} \tag{13}$$

Proof: Let $y \propto z$ *(or* p_i'*)* then $y = \tau z (or\ p_i')$

and $\tau = y_i / z(or\ p_i')$

Taking summation on both sides over the sample, we have

$$n\tau = \sum_{i=1}^{n} \frac{y_i}{p_i'}$$

so that

$$\tau = \frac{1}{n}\sum_{i=1}^{n} \frac{y_i}{p_i'} = \hat{\tau}_{MHH}$$

Proposition 2: The generalized transformation of the selection probabilities given as

$$p_{i,k}^{*} = \frac{1-\rho^c}{N} + \rho^c p_i' \tag{14}$$

with $0 \leq \rho^c \leq 1$ can be utilized in defining a class of alternative estimators and by extension, can be used along with the modified Hansen-Hurwitz's estimator. Detailed proof can be found in Ikughur and Amahia (2011a) and also, Enang and Amahia (2012). However, It is clear from the expression in (14) that

i. when ρ = 0, then $p_{i,k}^{*}$= 1/N;
ii. when ρ = 1 and c ≥ 0, then $p_{i,k}^{*} = p_{i}^{'}$
iii. when 0 < ρ^c < 1 and k > 0, we have $p_{i,k}^{*} = \frac{1-\rho^c}{N} + \rho^c p_i^{'}$. This provides a wider class of transformation to used with the proposed estimator.

3.1 Bias and variance of the proposed Estimator

The Hansen and Hurwitz's estimator is known to be unbiased. However when we consider it along with the generalized transformation of selection probabilities and hence, a class of alternative estimators in PPSWR sampling schemes defined by $p_{i,k}^{*},\ k = 1,2, \dots ,7$, then biased estimators are realized. It is worth to note here that the condition for unbiasedness is that $\rho = 1$ and $c \geq 0$ or when $c = 0$, a situation when $p_i = p_{i,1}^{*}$, which is satisfied by the suggested estimator.

Theorem 3: The Modified Hansen Hurwitz (MHH) Estimator is an unbiased estimator of population total Y, for $p_i^{'} = \frac{z_i}{Z}, z_i = \frac{1}{x_i}$ and $Z = \sum_{i=1}^{N} z_i$.

Proof: $E(\hat{\tau}_{MHH}) = E\left(\frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{p_i^{'}}\right) = \frac{1}{n}\sum_{i=1}^{n}\frac{E(y_i)}{p_i^{'}} = \frac{1}{n}\sum_{i=1}^{n}\sum_{i=1}^{N}\frac{y_i}{p_i^{'}}p_i^{'} = Y$

Note that the Bedi and Rao's (BR) and also Sahoo et al (S) estimators are biased as $E\left(\frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{p_i^{*}}\right) \neq Y$. Similarly, the alternative estimators in the class of MHH estimators are biased.

Theorem 4: The bias of the BR and S estimator under the design with probability $g_k(x) = p_i^{*}$ is

$$B_{p_i^{'}}(\hat{\tau}_k) = \sum_{i=1}^{N}\left(\frac{p_i}{p_{i,k}^{*}} - 1\right)y_i \text{ , k =0, 1,2,3,4,5,6,7} \qquad (13)$$

Proof : the bias of the estimators is defined as

$$B_{p_i^{'}}(\hat{\tau}_k) = E_{p_i^{'}}(\hat{\tau}_k) - Y$$

$$= \sum_{i\epsilon\Omega} \frac{E(y_i)}{p_{i,k}^*} - Y$$

$$= \sum_{i\epsilon n} \left(\frac{p_i^{'}}{p_{i,k}^*} - 1\right) y_i$$

Theorem 5: The design based variance of the proposed estimator is

$$V_{p_i^{'}}(\hat{\tau}_k) = \frac{1}{n}\left[\sum_{i\epsilon\Omega} \frac{y_i^2 p_i^{'}}{p_{i,k}^{*2}} - \left(\sum_{i\epsilon\Omega} \frac{y_i p_i^{'}}{p_{i,k}^*}\right)^2\right] \quad (14)$$

Proof: $V_{p_i^{'}}(\hat{\tau}_k) = V_{p_i^{'}}(\frac{1}{n}\sum_{i\epsilon s} \frac{y_i}{p_{i,k}^*})$

$$= \frac{1}{n^2} V_p(\sum_{i\epsilon s} \frac{y_i}{p_{i,k}^*})$$

$$= \frac{1}{n} V_p(z); \quad z = \frac{y_i}{p_{i,k}^*}$$

Since $\mathrm{Var}(z) = \mathrm{E}(z^2) - \mathrm{E}^2(z)$, it follows that

$$V_{p_i^{'}}(\hat{\tau}_k) = \frac{1}{n}\left[\sum_{i\epsilon\Omega} \frac{y_i^2 p_i^{'}}{p_{i,k}^{*2}} - \left(\sum_{i\epsilon\Omega} \frac{y_i p_i^{'}}{p_{i,k}^*}\right)^2\right]$$

3.2 Percentage Variance Relative to the suggested Estimators

The measure of efficiency used in this study is the Percentage Variance (PV) relative to the suggested estimators. This is given by

$$PV = \frac{V_{p_i^{'}}(\hat{\tau}_k)}{V_{p_i^{'}}(\hat{\tau}_{MHH})}\%, \; k = BR, S, 1,2,3,4 \quad (15)$$

$\hat{\tau}_k$ is more efficient than $\hat{\tau}_{MHH}$ if and only if $V_{p_i^{'}}(\hat{\tau}_k) < V_{p_i^{'}}(\hat{\tau}_{MHH}) < 100\%$.

4.0 Numerical Illustration.

For illustration, two populations namely population I and II respectively with ρ = -0.32 and -0.775 are considered in order to make comparison between the MHH estimator, Sahoo's (S), Bedi and Rao's (BR) estimator in the first place and also, with alternative estimators in PPS with replacement sampling denoted by $\hat{\tau}_{MHH}$, $\hat{\tau}_{BR}$, $\hat{\tau}_S$, $\hat{\tau}_1$, $\hat{\tau}_2$, $\hat{\tau}_3$ and $\hat{\tau}_4$. These estimators are defined by the probabilities p_i', $p_{i,11}^*$, $p_{i,8}^*$, $p_{i,3}^*$, $p_{i,5}^*$, $p_{i,6}^*$ and $p_{i,7}^*$ respectively. Under the inverse transformation, the correlation coefficients are $\rho = 0.205$ and 0.091 for population I and II respectively. It is clear that under inverse transformation, the estimated correlation coefficients are transformed from negative to positive. Detailed results for estimates of population totals, bias, variance and the percentage variance relative to the suggested estimator are shown on Tables 1 to 8.

Table 1: Estimate of Population Total for Population I with ρ = -0.37

P	$\hat{\tau}_{MHH}$	$\hat{\tau}_1$	$\hat{\tau}_2$	$\hat{\tau}_3$	$\hat{\tau}_4$	$\hat{\tau}_{BR}$	$\hat{\tau}_S$
0.0	5247.3	5055.0	5055.0	5055.0	5055.0	5055.7	5053.374
0.1	5247.3	5061.3	5054.7	5055.0	5055.0	5055.7	5053.374
0.2	**5247.3**	**5074.2**	**5054.0**	**5054.7**	**5054.9**	**5055.7**	**5053.374**
0.5	5247.3	5153.9	5061.6	5054.5	5053.8	5055.7	5053.374
0.9	5247.3	5435.8	5177.0	5152.0	5131.9	5055.7	5053.374
1.0	5247.3	5567.0	5247.3	5247.3	5247.3	5055.7	5053.374

Table 2: Estimate of Bias of the Estimators for Population I with ρ= -0.37

ρ	$B(\hat{\tau}_{MHH})$	$B(\hat{\tau}_1)$	$B(\hat{\tau}_2)$	$B(\hat{\tau}_3)$	$B(\hat{\tau}_4)$	$B(\hat{\tau}_{BR})$	$B(\hat{\tau}_S)$
0.000	0.000	7.148	7.148	7.148	7.148	1.424	-1.626
0.100	0.000	13.874	6.554	7.088	7.142	1.424	-1.626
0.205	**0.000**	**22.624**	**4.741**	**6.635**	**7.042**	**1.424**	**-1.626**
0.500	0.000	59.333	-3.961	0.656	3.653	1.424	-1.626
0.900	0.000	160.788	-5.723	-7.211	-8.064	1.424	-1.626
1	0	204.7835	0	0	0	1.424	-1.626

Table 3: Estimate of Variance of the Estimators for Population I with ρ= -0.37

ρ	$V(\hat{\tau}_{MHH})$	$V(\hat{\tau}_1)$	$V(\hat{\tau}_2)$	$V(\hat{\tau}_3)$	$V(\hat{\tau}_4)$	$V(\hat{\tau}_{BR})$	$V(\hat{\tau}_S)$
0	26386.7	9193.9	9193.9	9193.9	9193.9	10670.3	10259.2
0.1	26386.7	9531.6	9191.7	9193.4	9193.9	10670.3	10259.2
0.205	**26386.7**	**10633.3**	**9219.1**	**9191.7**	**9193.1**	**10670.3**	**10259.2**
0.5	26386.7	20685.7	10478.9	9513.3	9262.8	10670.3	10259.2
0.9	26386.7	83653.1	20722.7	18676.4	17004.2	10670.3	10259.2
1	26386.7	126435.7	26386.7	26386.7	26386.7	10670.3	10259.2

The important feature of MHH estimator is that it is unbiased (see tables 2 and 6) despite the fact that it has higher estimate of the population total as shown on tables 1 and 5 for the two study populations. Results of bias of the estimators as shown on tables 2 and 6 indicate that apart from MHH estimator, all other estimators are bias. The bias is constant throughout the parameter space defined by the correlation coefficient, ρ which makes the estimate a fixed value. Thus, the bias of the MHH estimator $\mathcal{B}_{p_i'}(\hat{\tau}_{MHH}) = 0$ at all values of ρ while $\mathcal{B}_{p_i'}(\hat{\tau}_S) > \mathcal{B}_{p_i'}(\hat{\tau}_{MHH})$ and also, $\mathcal{B}_{p_i'}(\hat{\tau}_{BR}) > \mathcal{B}_{p_i'}(\hat{\tau}_{MHH})$. On the other hand, $\mathcal{B}_{p_i'}(\hat{\tau}_S) > \mathcal{B}_{p_i'}(\hat{\tau}_{BR})$.

Table 4: Percentage Variance Relative to the suggested Estimators for Population I with ρ= -0.37

ρ	$\hat{\tau}_{MHH}$	$\hat{\tau}_1$	$\hat{\tau}_2$	$\hat{\tau}_3$	$\hat{\tau}_4$	$\hat{\tau}_{BR}$	$\hat{\tau}_S$
0.1	100.0	36.1	34.8	34.8	34.8	40.4	38.9
0.215	**100.0**	**40.3**	**34.9**	**34.8**	**34.8**	**40.4**	**38.9**
0.5	100.0	78.4	39.7	36.1	35.1	40.4	38.9
0.9	100.0	317.0	78.5	70.8	64.4	40.4	38.9
1	100.0	479.2	100.0	100.0	100.0	40.4	38.9

Table 5: Estimate of Population Total for Population II with ρ= -0.775

P	$\hat{\tau}_{MHH}$	$\hat{\tau}_1$	$\hat{\tau}_2$	$\hat{\tau}_3$	$\hat{\tau}_4$	$\hat{\tau}_{BR}$	$\hat{\tau}_S$
0.0	44137.5	11680.0	11680.0	11680.0	11680.0	11691.7	11320.57
0.1	44137.5	11772.7	11674.4	11679.2	11679.9	11691.7	11320.57
0.5	44137.5	14399.8	12353.5	11837.5	11703.3	11691.7	11320.57
0.9	44137.5	26872.5	21379.8	18555.4	16815.9	11691.7	11320.57
1.0	44137.5	44137.5	44137.5	44137.5	44137.5	11691.7	11320.57

Table 6: Estimate of Bias of the Estimators for Population II with ρ= -0.775

ρ	$B(\hat{\tau}_{MHH})$	B($\hat{\tau}_1$)	$B(\hat{\tau}_2)$	B($\hat{\tau}_3$)	$B(\hat{\tau}_4)$	$B(\hat{\tau}_{BR})$	$B(\hat{\tau}_S)$
0.000	0.000	77.395	77.395	77.395	77.395	23.384	−359.434
0.100	0.000	-83.404	55.835	75.163	77.171	23.384	−359.434
0.500	0.000	-271.979	−202.05	-110.258	−35.017	23.384	−359.434
0.900	0.000	-168.806	−227.52	-255.588	−269.20	23.384	−359.434
1.000	0.000	0.000	0.000	0.000	0.000	23.384	−359.434

In terms of variance, it is observed that $\hat{\tau}_S$ is more efficient with $V_{p_i'}(\hat{\tau}_S)$ < $V_{p_i'}(\hat{\tau}_{BR})$ < $V_{p_i'}(\hat{\tau}_{MHH})$. However, when the transformed probabilities are used along with MHH estimator, then $V_{p_i'}(\hat{\tau}_3)$ = $V_{p_i'}(\hat{\tau}_4) < V_{p_i'}(\hat{\tau}_S)$ < $V_{p_i'}(\hat{\tau}_{BR})$ < $V_{p_i'}(\hat{\tau}_{MHH})$ at ρ=0.0205 which is the estimated correlation coefficient of population I under inverse transformation. This is further shown by the Percentage Variance Relative to the suggested Estimators in which

$V_{p_i'}(\hat{\tau}_3)$ = $V_{p_i'}(\hat{\tau}_4) = 34.8\% < V_{p_i'}(\hat{\tau}_S) = 38.9\%$ < $V_{p_i'}(\hat{\tau}_{BR}) = 40.4\%$ < $V_{p_i'}(\hat{\tau}_{MHH}) = 100\%$.

Table 7: Estimate of Variance of the Estimators for Population II with ρ= -0.775

ρ	$V(\hat{\tau}_{MHH})$	$V(\hat{\tau}_1)$	$V(\hat{\tau}_2)$	$V(\hat{\tau}_3)$	$V(\hat{\tau}_4)$	$V(\hat{\tau}_{BR})$	$V(\hat{\tau}_S)$
0	1468770.7	177392.9	177392.9	177392.9	177392.9	141525.3	62159.05
0.1	1468770.7	171679.7	175628.9	177200.9	177373.5	141525.3	62159.05
0.5	1468770.7	257446.0	190726.9	173253.4	171163.9	141525.3	62159.05
0.9	1468770.7	653465.2	467124.2	380823.7	329402.4	141525.3	62159.05
1	1468770.7	1468770.7	1468770.7	1468770.7	1468770.7	141525.3	62159.05

Table 8: Percentage Variance Relative to the suggested Estimators for Population II with ρ= -0.775

ρ	$\hat{\tau}_{MHH}$	$\hat{\tau}_1$	$\hat{\tau}_2$	$\hat{\tau}_3$	$\hat{\tau}_4$	$\hat{\tau}_{BR}$	$\hat{\tau}_S$
0	100.0	12.1	12.1	12.1	12.1	9.6	4.2
0.1	100.0	11.7	12.0	12.1	12.1	9.6	4.2
0.5	100.0	17.5	13.0	11.8	11.7	9.6	4.2
0.9	100.0	44.5	31.8	25.9	22.4	9.6	4.2
1	100.0	100.0	100.0	100.0	100.0	9.6	4.2

Considering population II, MHH estimator appears to give a very high estimate of population total than $\hat{\tau}_S$, $\hat{\tau}_{BR}$ and all other estimators. However, $\hat{\tau}_S$ has a very high bias than all other estimators with $B(\hat{\tau}_S) = |359.434|$ yet, is more efficient than other estimators considered in terms of percentage variance relative to the suggested estimators

5.0 Concluding remark

The main features of the Modified Hansen and Hurwitz (MHH) estimator is that it gives an unbiased estimator of population total. Apart from its property of unbiasness, it provides an efficient estimate when used along with the transformation in (14) above for the class of alternative estimators suggested by Enang and Amahia (2012) especially, when there is need to exploit information contained in correlation coefficient to provide estimate of population total.

By the law of inverse transformation and variable substitution, correlation structure is changes to positive correlation so that the conventional Hansen and Hurwitz estimator is modified under this condition. The use of this estimator could be extended to multistage sampling, randomized responses, indirect responses and even when non-sampling error are encountered.

References

Amahia, G.N., Chaubey, Y.P. and Rao, T.J. (1989). Efficiency of a new PPS sampling for multiple characteristics. *Journal of Statistical Planning and Inference, (21):75-84.*

Bansal M.L. and Singh, R. (1985). An alternative estimator for multiple characteristics in PPS sampling. *Journal of Statistical Planning and Inference.* (21):75-84.

Bedi P.K (1995). An alternative estimator in Midzuno scheme for multiple characteristics. *Commun. Statist. Simula.: 17-30.*

Bedi, P.K and Rao, T.J. (1997). PPS method of estimation under a transformation. *Journal of the Indian Society of Agricultural Statistics, pp184-195*

Enang, E.I and Amahia, G.N. (2012). A Class of Alternative Estimators in Probability Proportional to Size Sampling with Replacement for Multiple Characteristics. *Journal of Mathematics Research;* 4(3):66-77

Grewal, I.S, Bansal M.L. and Singh. (1997): An alternative estimator for multiple characteristics using randomized response technique in PPS sampling. *Aligarh Jour. Statist,* (19):51-65.

Hansen, M. H. and Hurwitz, W.N. (1943). On the theory of sampling from a finite population. *Ann. Math. Stat*. (14):333 – 362.

Ikughur, A.J. and Amahia, G.N. (2011a). A generalized linear transformation for selection probabilities in unequal probability sampling schemes. *Jour. Scientific and Industrial Studies.* 9(1):58-63

Ikughur, A.J and Amahia, G.N. (2011b). On the direction of relationship: implication to finite population estimation. *Journal of Research in Physical Sciences,* 7(1):84-87

Pathak, P.K. (1966). An estimator in PPS sampling for multiple characteristics. *Sankhya, A,* 28(1):35-40.

Rao, J.N.K. (1966a) Alternative estimators in the PPS sampling for multiple characteristics. *Sankhya,* 28(A):47-60.

Rao, J.N.K. (1966b). On the relative efficiency of some estimators in PPS sampling for multiple characteristics. *Sankhya,* 28(A):61-70.

Rao, T.J (1993a). On certain alternative estimators for multiple characteristics in varying probability sampling. *Jour. Indian Soc. Agric. Statist.* 45(3):307-318

Rao, T.J (1993b). On certain problems of sampling design and estimation for multiple characteristics. *Sankhya,* 55(B):372-381.

Sahoo, L.N, Dalabehera, M and Mangaraj A.K. (2012). A *PPS* Sampling Scheme Using Harmonic Mean of an Auxiliary Variable. *ARPN Journal of Systems and Software*, 2(7):207-210.

Sahoo, J., Sahoo, L.N. and Mohanty, S. (1994).Unequal probability sampling using a transformed auxiliary variable. *Metron*, 52:71-83.

Singh, S. and Horn, S. (1998). An alternative estimator for multi-character surveys. *Metrika*, 48:99-107

Singh S, Grewal I.S and Joarder A. (2004). General class of estimators in multi-character surveys. *Statistical papers*, 45:571-582.

Singh, H.P and Tailor, R. (2003). Use of known correlation coefficient in estimating the finite population mean. *Statistics in Transition*, 6:555-560.

Singh, H.P and Taylor, R. (2005). Estimation of population mean with known coefficient of variation of auxiliary characters. *Statistica Anna* LXV: 301-313.

Appendix: Study Populations I and II

Population I

x	60	89	77	71	59	58	72
y	115	80	82	93	105	109	130
1/x	0.016667	0.011236	0.012987	0.014085	0.016949	0.017241	0.013889

x	70	52	39
y	93	109	95
1/x	0.014286	0.019231	0.025641

Population II

x	6.8	6.2	5.5	0.85	0.71	9	1.4	4.5	3.8
y	20	23	38	86	92	16	81	53	42
ρ	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	0.1470	0.1612	0.1818	1.1764	1.4084	0.1111	0.7142	0.2222	0.2631
1/x	59	9	18	71	51	11	86	22	58

x	2.1	4.85	3.197	0.443	0.468	0.59	0.339	0.161	0.787
y	62	39	35	87	91	84	75	54	64
ρ	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	0.4761	0.2061	0.3127	2.2573	2.1367	1.6949	2.9498	6.2111	1.2706
1/x	9	86	93	36	52	15	53	8	48

x	0.069	0.11
y	26	100
ρ	0.1	0.1
	14.492	9.0909
1/x	75	09