Population Total and Variance Estimation in Multi-Stage Cluster Sampling: A Simple Random Sampling Approach

E. I. Enang [1] and I. L. Onyishi [1,2]
[1]Department of Statistics,
University of Calabar, Calabar, Nigeria.
[2]Department of Mathematics and Statistics,
Federal Polytechnic, Nasarawa, Nasarawa State, Nigeria.
Email: ifyifyedo2@gmail.com

Abstract

An estimator (a statistic) cannot be better than the parameter it estimates. Multi-stage sampling thrives because of its cost-effectiveness and not because its estimator is better. This study emphasis the need to develop higher-order estimators of multi-stage sampling and their use for large-scale sample survey. In this study, the researchers developed estimators of population total and variance of four-stage cluster sampling and illustrated their use with numerical example. Researchers should not sacrifice the development and application of higher-order estimators of multi-stage sampling that are robust statistical techniques on the altar of the complexity of design nature of large population. To reap the actual fruit of sampling that culminates in the high level of predication and estimation accuracy in sample survey, higher-order multi-stage estimators should be used in large-scale sample surveys.

Keywords: *Estimators, multi-stage sampling, population total, population variance, cost-effectiveness, simple random sampling*

1.0 Introduction

Although a lot of estimation procedures of population parameters have been in existence, more still need to be done in order to make sampling schemes and their results depict the real life situations which they attempt to explain. Meteorological predictions of weather conditions have high level of acceptability, because almost all predictions of meteorologists at local, national and international levels are successful. It would not be odd if sample survey estimates yield even more generally acceptable results. The level of success and acceptance of estimation is dependent on the level of efficiency and adequacy of estimator.

Multistage sampling is a sampling scheme which deepens the real essence of sampling. Its cost-effectiveness attests to this fact. The scheme has attracted the attention of many renowned scholars who have developed estimation procedures for it. Some of the procedures include those of Cochran (1977), Kelton (1983), Okafor (2002) and Nafiu et al (2012) that derived a generalized form of multistage cluster sampling design based on lower-order stage of the Nafius (2012) procedure.

Most of the procedures aforementioned focus on lower – order multistage, such as two or three-stage and their generalizations are often based on the lower – order stage. This limits the ability of other researchers who might not be in a position to develop, derive or extend the generalized procedure for use in large scale sampling. In this study, the researchers are of the view that higher – order multistage sampling procedures should be developed and used to improve the reliability of sample surveys, especially in large scale surveys. Moreso, every stage of a multistage sampling is a sampling technique itself, Lukman (2012). In order to make available higher – order multistage sampling procedure, the researchers' aim in this study is to develop estimators of the parameters of four-stage cluster sampling based on Okafor (2002) procedure. The objectives of the study are to (i) develop estimators of the population total and variance of four-stage cluster sampling, (ii) apply the developed estimators to the estimation of population total of farmers in a vicinity within Nasarawa and Keffi local government areas of Nasarawa State and (iii) determine the standard error and coefficient of variation of the estimated population total.

Wolter, (2007) noted that naturally, representative samples of large populations often have complex design features because of their cost efficiency and size, among other reasons. For purposes of estimating sampling variances based on complex multi-stage sample design involving cluster sampling, sampling error codes are provided by survey organizations in public use survey files, in some developed countries of the world. The codes, where they exist, are used for "ultimate cluster selection" for individuals, Rendtel and Harms (2009).

In some developed countries and developing countries like Nigeria where those codes do not exist, the normal procedure for estimating variance holds. This procedure is explained in the next section of this study. It is however necessary to note that if the first stage units (fsu's) were actually sampled without replacement in multi-stage design, which is typical of this study, standard estimates of population totals are estimated and Kerry and Bland, (2006) asserted that the variance estimator, in this case, only have a slight positive bias with the selection of ultimate cluster. Lavallee (2010) argue that since unbiased estimators for estimating totals in multi-stage cluster samples of various stages are primarily driven by the variance estimated totals within a stage, at least two first stage units are needed within a first stage sampling cluster for design based variance estimation purposes.

It is a common mathematical knowledge that variance estimators for sample totals based on complex designs featuring without replacement selection of first stage units at the sampling stage, Canette (2010) affirm, are function of finite population correlations (FPC's) which account for the proportion of the finite population that was not included in a sample selected without replacement and joint probabilities of selection for sampled units. Carle (2009) and Rabe-Hesheth and Skrondal (2006) stress that generally, large sample sizes lead to FPC.

2.0 Materials and Methods

2.1 Notations for Sample Selection

The following notations are used in this study

N = the number of clusters in the first stage unit (fsu's)

n = the number of clusters selected from N fsu's by simple random sampling without replacement (srswor)

M_i = the number of clusters in the second stage unit (ssu)

m_i= the number of clusters selected by srswor from $M_i(i = 1,2,....... n)$ ssu's

M_{ij} = the number of clusters in the third stage unit (tsu)

m_{ij} = the number of clusters selected by srswor from M_{ij} (j = 1,2,......, mi) tsu's

Finally

M_{ijk} = the number of clusters (elements) in the fourth stage unit (4th su)

m_{ijk} = the number of clusters (elements) selected by srswor from M_{ijk} (K = 1,2, ...mij) 4thsu's in each of the selected tsu's

Y_{ijkl} = the value of the characteristic, y, from lth4th su in the kthtsu of the jthssu in the ith fsu.

Y = the population total

2.2 Sample Selection in Multi-Stage Sampling

(*equal probability sampling with first stage units of unequal sizes*)

From a population of N first stage units (fsu's), n fsu's are selected by simple random sampling without replacement (srswor). Next, a srswor of size mi second stage units (ssu's) is selected from M_i (i=1,2,......,n) within each of the selected fsu's. Then, a srswor of size m_{ij} third stage units (tsu's) is selected from M_{ij} (j = 1,2.......m_i) within each of the selected ssu's. A srswor of size m_{ijk} fourth stage units (4^{th}su's) is lastly selected from M_{ijk} (K = 1,2,.....m_{ij}) within each of the selected tsu's

Now, we assume that the value of the characteristic, y, of interest, y_{ijkl}, in view of the notations defined above, is

y_{ijkl}= the value of the characteristic, y, from lth 4^{th}su in the kth tsu of the jth ssu in the ith fsu.

2.3 List of Population Means and Totals of Interest

$$\bar{Y}_{ijk.} = \frac{1}{M_{ijk}} \sum_{1=1}^{Mijk} Y_{ijkl} = \frac{Y_{ijk.}}{M_{ijk}} =$$ the population mean of the kth tsu in the jth ssu of the ith fsu.

$Y_{ijk.}$ = the population total of the kth tsu in the jth ssu of the ith fsu

$$\bar{Y}_{ij..} = \frac{1}{M_{ij}} \sum_{K=1}^{Mij} Y_{ijk.} = \frac{Y_{ij..}}{M_{ij}}$$ = the population mean of the jth ssu in the ith fsu

$Y_{ij..}$ = the population total of the jth ssu in the ith fsu

$$\bar{Y}_{i...} = \frac{1}{M_i} \sum_{j=1}^{M_i} Y_{ij..} = \frac{Y_{i...}}{M_i} =$$ = the population mean of the ith fsu

$Y_{i..}$ = the population total of the ith fsu

2.4 List of Sample Means and Totals of Interest

$$\bar{y}_{ijk.} = \frac{1}{m_{ijk}} \sum_{i=1}^{mijk} y_{i\,jkl} = \frac{y_{ijk.}}{m_{ijk}}$$ = the sample mean of the kth tsu in the jth ssu of the ith fsu

$y_{ijk.}$ = the sample total of the kth tsu in the jth ssu of the ith fsu

$$\bar{y}_{ij.} = \frac{1}{m_{ij}} \sum_{k=1}^{mij} y_{i\,jk.} = \frac{y_{ij..}}{m_{ij}}$$ = the sample mean of the jth ssu in the ith fsu

y_{ij} = The sample total of the jth ssu in the ith fsu

$$\bar{y}_{i...} = \frac{1}{m_i} \sum_{j=1}^{mi} y_{i\,j..} = \frac{y_{i...}}{m_i}$$ = the sample mean of the ith fsu

$y_{i...}$ = the sample total of the ith fsu

2.5 Estimation of Population Total

We start from the inclusion probability as follows:

The inclusion probability of Horvitz-Thompson Estimator in the jth ssu within the ith fsu according to Okafor (2002) and Wolter (2007) is given by

$$\pi_{(ij)} = (\frac{n}{N})(\frac{m_i}{M_i}) \quad (1)$$

Where

$$\frac{n}{N} = f_1 \quad , \quad (2)$$

is the sampling fraction of fsu,

$$\frac{m_i}{M_i} = f_2 \quad , \quad (3)$$

is the sampling fraction of ssu

In view of equation (1), we add that

the inclusion probability of the kth tsu within the jth ssu in the ith fsu is

$$\pi_{(ijk)} = (\frac{n}{N})\,(\frac{m_i}{M_i})\,(\frac{m_{ij}}{M_{ij}}) \quad (4)$$

Where

$$\frac{m_{ij}}{M_{ij}} = f_3 \quad , \qquad (5)$$

is the sampling fraction of the tsu

Similarly, we include that;

the inclusion probability of the lth 4th su within the kth tsu in the jth ssu of the ith fsu is

$$\pi_{(ijk)} = (\frac{n}{N})\,(\frac{m_i}{M_i})\,(\frac{m_{ij}}{M_{ij}})(\frac{m_{ijk}}{M_{ijk}}) \qquad (6)$$

Where

$$\frac{m_{ijk}}{M_{ijk}} = f_4 \quad , \qquad (7)$$

is the sampling fraction of the 4th su.

$$\frac{N}{n} = \frac{1}{f_1};\ \frac{M_i}{m_i} = \frac{1}{f_2};\ \frac{M_{ij}}{m_{ij}} = \frac{1}{f_3} \qquad (8)$$

are the raising factors of fsu's, ssu's, tsu's, respectively.

Using the inclusion probability, in this study, as shown in (1) from which (4) and (6) were derived, we further state that the unbiased estimator of the population total, y, of four-stage cluster sampling, $\hat{Y}$, given by

$$\hat{Y} = \frac{N}{n}\sum_{i=1}^{n}\frac{M_i}{m_i}\sum_{j=1}^{mi}\frac{M_{ij}}{m_{ij}}\sum_{k=1}^{mij}\frac{M_{ijk}}{m_{ijk}}\sum_{i=1}^{mijk}Y_{ijkl} = \frac{N}{n}\sum_{i=1}^{n}\frac{M_i}{m_i}\sum_{j=1}^{mi}\frac{M_{ij}}{m_{ij}}\sum_{k=1}^{mij}M_{ijk}\,\bar{y}_{ijk} \qquad (9)$$

Where

$$= \frac{M_{ijk}}{m_{ijk}} = \frac{1}{f_4} \quad , \qquad (10)$$

is the raising factor of the 4th su

2.6 Notations for variance estimation

E_4 = the conditional expectation over the selection of the 4th su's selected from the tsu's that was obtained from the ssu's in each fsu.

V_4 = its corresponding variance

E_3 = the conditional expectation over the selection of tsu's from the ssu's in each fsu

V_3 = its corresponding variance

E_2 = the conditional expectation over the selection of the ssu's from each fsu

V_2 = its corresponding variance

E_1 = the unconditional expectation over all possible samples of n fsu

V_1 = its corresponding variance

In the same vein, we state that the variance of the estimator of the population total of four-stage cluster sampling, $\hat{Y}$, which is $V\left(\hat{Y}\right)$ is derived by the model

$$V\left(\hat{Y}\right) = E_1E_2E_3V_4\left(\hat{Y}\right) + E_1E_2V_3E_4\left(\hat{Y}\right) + E_1V_2E_3E_4\left(\hat{Y}\right) + V_1E_2E_3E_4\left(\hat{Y}\right) \qquad (11)$$

Applying the general theorem for obtaining the sampling variance of the sample mean $\left(\bar{y}\right)$ in element sampling which states that

$$V\left(\bar{y}\right) = \frac{\sigma^2}{n} \quad \textit{for srswr} \qquad (12)$$

Or $$V\left(\bar{y}\right) = \frac{N-n}{N-1}\frac{\sigma^2}{n} \quad \textit{for srswor} \qquad (13)$$

Where $\sigma^2 = \sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2 / N =$ the population (element) variance, (14)

using the analysis of variance approach which gives the population variance as

$$S^2 = \sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2 / (N-1), \qquad (15)$$

and by substituting equation (15) in (14); then, (14) in (13), we obtain that the sampling variance of the fsu is

$$V^1\left(\hat{y}\right) = N^2\frac{1-f_1}{n}S_1^2 \qquad (16)$$

Similarly, the sampling variance of the ssu and the tsu are obtained. The sampling variance of the 4^{th}su is therefore derived as

$$V^{1111}\left(\hat{y}\right) = \frac{N}{n}\sum_{i=1}^{n}\frac{M_i}{m_i}\sum_{j=1}^{M_i}\frac{M_{ij}}{m_{ij}}\sum_{k=1}^{m_{ij}}M_{ijk}\frac{1-f_{4ijk}}{m_{ijk}}S_{4ijk}^2 \qquad (17)$$

Hence, we further state that the unbiased sample estimator of the sampling variance, $\hat{V}\left(\hat{Y}\right)$ of a four-stage cluster sampling is given by

$$\hat{V}\left(\hat{y}\right) = N^2 \frac{1-f_1}{n} s_1^2 + \frac{N}{n}\sum_{i=1}^{n} M^2{}_i \frac{1-f_{2i}}{m_i} s_{2i}^2 + \frac{N}{n}\sum_{i=1} \frac{M_i}{m_i}\sum_{j=1}^{mi} M_i{}^2 \frac{1-f_{3ij}}{m_{ij}} s_{3ij}^2$$

$$+\frac{N}{n}\sum_{i=1}^{n} \frac{M_i}{m_i}\sum_{j=1}^{mi} \frac{M_{ij}}{m_{ij}}\sum_{k=1}^{mij} M^2{}_{ijk} \frac{1-f_{4ij}}{m_{ijk}} s_{4ijk}^2 \qquad (18)$$

Where

$$s_1^2 = \sum_{j=1}^{n}\left(\hat{y}_{i\ldots} - \hat{y}_{\ldots}\right)^2 /(n-1) \quad , \qquad (19)$$

is the sample variability among cluster totals in the fsu

$$s_{2i}^2 = \sum_{j=1}^{mi}\left(\hat{y}_{ij..} - \hat{y}_{i\ldots}\right)^2 /(m_i - 1) \qquad (20)$$

$$s_{3ij}^2 = \sum_{k=1}^{mij}\left(\hat{y}_{ijk.} - \hat{y}_{ij..}\right)^2 /(m_{ij} - 1) \qquad (21)$$

$$s_{4ijk}^2 = \sum_{i=1}^{mijk}\left(Y_{ijkl} - \hat{y}_{ijk.}\right)^2 /(m_{ijk} - 1) \qquad (22)$$

Equations (20), (21) and (22) are the sample variabilities among cluster totals in the 2^{nd}, 3^{rd} and 4^{th} stage units, respectively.

f_1, f_2, f_3, and f_4 are as defined in equations 2, 3, 5 and 7, respectively.

3.0 Illustrative Examples

Application of the developed estimators of population total and population variance of four-stage cluster sampling.

Once a high-order stage estimator of multi-stage cluster sampling is developed, based on a particular procedure, every other lower stage of the procedure can be derived from it.

We shall apply the estimators of two-stage and four-stage cluster sampling to the estimation of the total number and the variance of the estimated population total of professional farmers.

3.1 Illustrative Example 1

Application of the Estimators of Population total and variance of two-stage cluster sampling

In this study, a sample survey of professional farmers from a vicinity that comprises 60 Enumeration Areas (EA's) in the boundary of Nasarawa and Keffi LGA's of Nasarawa State was conducted. 10 of the EA's were selected by simple random sampling without replacement. A list of households (HH) from each EA was also selected according to the size (the number of households) of the EA by simple random sampling. The data is shown in Table 1.

Table 1: ***Data of Professional Farmers at Nasarawa/Keffi Vicinity***

EA	Number of HH	Sample HH	Number of Professional Farmers (yi) in each sample household (HH)
1	52	7	3, 0, 4, 1, 2, 1, 3
2	70	9	1, 4, 3, 11, 4, 2, 1, 3
3	39	5	3, 0, 1, 2, 2
4	33	4	1, 0, 2, 4
5	22	3	1, 1, 3
6	31	4	5, 5, 3, 0
7	30	4	1, 2, 2, 3
8	54	7	0, 1, 3, 4, 2, 1, 2
9	18	2	2, 1
10	30	4	1, 3, 2, 0

Source: ***A Survey of Selected EA's at Nasarawa/Keffi Vicinity in September, 2015***

Table 2: ***Estimation of the Total Number of Professional Farmers as Obtained from Table 1***

EA	M_i	m_i	$\bar{y}_i$	$M_i\bar{y}_i$	S_{wi}^2	f_{2i}	$M_i^2(1-f_{2i})\frac{S_{wi}^2}{m_i}$
1	52	7	2.00	104.00	2.00	0.134	12.867
2	70	9	2.22	155.40	1.69	0.129	11.449
3	39	5	1.60	62.40	1.30	0.128	8.841
4	33	4	1.75	57.75	2.92	0.121	21.175
5	22	3	1.67	36.74	1.33	0.136	8.427
6	31	4	3.25	100.75	5.58	0.129	37.667
7	30	4	2.00	60.00	0.67	0.133	4.357
8	54	7	1.86	100.44	1.81	0.130	12.148
9	18	2	1.50	27.00	0.50	0.111	4.001
10	30	4	1.50	45.00	1.67	0.133	10.859
Total	379	49		749.48			131.788

Source: Computed from Table 1

The survey is considered to have given rise to a two stage sampling scheme.

The estimators of the population total and variance of a two-stage cluster sampling derived from equations (9) and (18) of this study are given by

$$\hat{Y} = \frac{N}{n}\sum_{i=1}^{n}\frac{M_i}{m_i}\sum_{j=1}^{mi} y_{ij} = \frac{N}{n}\sum_{i=1}^{n} M_i\bar{y}_i = \frac{N}{n}\sum_{i=1}^{n}\hat{y}_i \qquad (23)$$

and

$$\hat{V}(\hat{Y}) = N^2 \frac{1-f_i}{n} S_1^2 + \frac{N}{n} \sum_{i=1}^{n} M_1^2 \frac{1-f_{2i}}{m_i} S_{w_i}^2, \quad (24)$$

Where

$$S_1^2 = \sum_{i=1}^{n} \left(M_i \bar{y}_i - \frac{1}{n} \sum_{i=1}^{n} M_i \bar{y}_i\right)^2 \Big/ (n-1) \quad (25)$$

From equation (23) and column 5 of Table 2, the estimate of the total number of professional farmers in the vicinity under study is

$\hat{Y} = \frac{60}{10} \times 749.48 = 4496.88$

$\cong 4497$

From equation (24) and table 2, the estimate of the variance is

$$\hat{V}(\hat{Y}) = 60^2 \frac{1 - \frac{10}{60}}{10} \times 1552 + \frac{60}{10} \times 131.788$$

= 465,600 + 790.728

= 466,390.728

The estimated standard error is

$Se(\hat{Y}) = \sqrt{466390.728} = 682.9280548$

The estimated coefficient of variation (relative standard error) is

$$Cv(\hat{Y}) = \frac{682.9280548}{4497} \times 100 = 15.19\%$$

The estimate of the variability among cluster (EA) totals is

$\text{Est}(S_1^2) = 1552 - 131.788/10 = 1538.8212$

3.2 Illustrative Example 2

Application of the Estimators of Population Total and Variance of Four-Stage Cluster Sampling

In this case, secondary data is used. From a densely populated community made up of 60 enumeration areas (EA), 10 enumeration areas were first selected. Due to the size, the number of households (HH), of each selected EA, subsequent sub-groups of households were further selected from each EA in order to obtain a feasible and sizable list of households from which professional farmers were finally sampled. Computation of the population total and variance of the professional farmers in the densely populated community, using equations (9) and (18) are shown in tables 3, 4, 5 and 6 which are given in appendices I, II, III and IV, respectively, in this study.

From equation (25) using column 5 of table 3;

S_1^2 = 59085.05689

From equation (9) and Table 6, the estimate of the total number of professional farmers in the vicinity with larger number of households is;

$$\frac{60}{10} \times 3634.276585 = 21805.65951$$

≅ 21806

From equation (18) using column 8 of Tables (3), Table (4a), Table (4b), column 4 of Table 5 and column 7 of Table 6, the estimate of the variance of $\hat{Y}$, is

$$\hat{V}(\hat{Y}) = 60^2 \frac{1-\frac{10}{60}}{10} \times 59085.05689 + \frac{60}{10} \times 4432.586 + \frac{60}{10} \times 3918.543571 + \frac{60}{10} \times 20878.55033$$

= 17,725,517.07 + 26,595.516 + 23,511.26143 + 125,271.302

= 17,900,895.15

The estimate of the standard error is

$Se(\hat{Y}) = \sqrt{17900895.15}$ = 4,230.944948

The estimated coefficient of variation (relative standard error) is

$$Cv(\hat{Y}) = \frac{4230.944948}{21806} \times 100\% = 19.40\%$$

4.0 Discussion of Results

In this study, the inclusion probability which according to Okafor (2002) and Wolter (2007) originated from horvit-Thompson Estimator is employed by the researchers to develop estimators of four-stage cluster sampling (for large-scale survey) for population total and variance. The development is initiated from equation (1) through equation (9) to equation (18), where the derivation procedure terminates.

Two numerical illustrative examples are given. The first is the use of the developed estimators of population total and variance for two-stage cluster sampling in equations (23) and (24) derived from equations (9) and (18), respectively. The results of illustrative example 1, from the data of table 1 and the computation in table 2 show that the population of professional farmers in the given Nasarawa/Keffi vicinity is estimated to be 4497 with a variance estimate of 466,391. The estimated standard error is 683 while the coefficient of variation is estimated to be 15.19%. The estimate of the variability among cluster totals is 1539.

The computation of illustrative example 2 shown in tables 3, 4a, 4b, 5 and 6 indicate that in the larger vicinity, the estimate of the population of professional farmers is 21806 with an estimated variance of 4231. Computation of the result of illustrative example 2 is based on the use of the derived estimators of population total and variance of four-stage cluster sampling of equations (9) and (18) which further give the estimate of coefficient of variation of the professional farmers in the larger vicinity as 19.40%.

5.0 Conclusion

The benefits of sampling are much more appreciated in multi-stage cluster sampling. Its elegance is more obvious in higher-order multistage sampling which is typical of large-scale surveys. The use of lower-order multi-stage estimators for the analysis of data arising from large population, more often than not, end up in misleading or unacceptable results. The researchers, in this study, used the estimators of two-stage cluster sampling to illustrate the use of lower-order multi-stage estimators while the derived four-stage estimator was used to illustrate the use of higher-order multi-stage estimator.

Researchers should be encouraged to develop and use higher-order multi-stage estimators for the analysis of large-scale surveys to increase the level of accuracy and acceptability of sample survey results. Encouragement, in this regard, should be in the form of sponsorship of survey, commercialization and patent (where necessary).

References

Carnette, I. (2010): *Analysis of Complex Survey Data in Stata.* Paper Present at the 2010 Mexican Stata Users Group Meeting, April 29, 2010

Carle, A. C. (2009): *Fitting Multilevel Models in Complex Survey data with Design Weights: Recommendations*. BMC Medical Research Methodology, 1471-2288-9-49

Cochran, W. G. (1977): *Sampling Techniques.* Third Edition. John Wiley and Sons New York.

Kalton, G. (1983): *Introduction to Survey Sampling.* Sape Publications: Los Angeles, CA.

Kerry, I. and Bland J. (2006): *The Intra-cluster Correlation Coefficient in Cluster Recommendation*. British Medical Journal, 316, 1455-1460

Lavallee, P. (2010): *Cross Section Weighting of Longitudinal Surveys of Individuals and Households Using the Weight Share Method.* Survey Methodology, 32, 165-176

Lukman, A. N. (2012): *Comparison of One-Stage, Two-Stage and Three-Stage Estimators Using Finite Population.* The Pacific Journal of Science and Technology, Volume 13, Number 2, November 2012.

Nafiu, L. A, Oshungade, I. O. and Adewara, A. A. (2012): *Generalization (Unequal Probability Class) of Multi-Stage Cluster Sampling Using Finite Population.* International Journal of Engineering and Applied Sciences.

Okafor, F. C. (2002): *Sample Survey Theory with Applications.* Afro-Orbis Publications Ltd 33 Catering Rest House Road, P. O. Box 3160, University of Nigeria, Nsukka, Nigeria.

Rabe-Hesheth, S. and Skrondal, A. (2006): *Multilevel Modelling of Complex Survey Data.* Journal of the Royal Statistical Society-Series A, 169, 805-827

Rendtel, W. and Harms, T. (2009): *Weighting and Calibration for Household Panels in Methodology of Longitudinal Surveys*. Ed. P. Lynn, Chichester: John Wiley and Sons, 265-286

Wolter, K. M. (2007): *Introduction to Variance Estimation.* Second Edition Springer-Verlag.

Appendix 1

Table 3: Estimation of Professional Farmer in the 1st and 2nd Stages of a 4-Stage Cluster Sampling

S/No of EA	M_i	m_i	$\bar{y}_i$	$M_i\bar{y}_i$	S_{2i}^2	f_{2i}	$M_i^2 \frac{1-f_{2i}}{m_i} S_{2i}^2$
1.	250	115	2.34	585.00	2.15	0.460	630.978
2.	370	155	2.51	928.70	2.01	0.419	1031.440
3.	189	87	1.12	211.68	1.42	0.460	314.838
4.	159	73	1.75	278.25	3.02	0.459	515.816
5.	106	49	1.38	146.28	1.21	0.471	146.777
6.	150	69	1.29	193.50	4.34	0.460	764.217
7.	146	67	2.54	370.84	0.53	0.458	91.392
8.	261	120	1.36	354.96	1.77	0.460	542.584
9.	87	40	1.59	138.33	0.48	0.460	49.047
10.	146	67	1.86	271.56	1.72	0.460	295.497
	1864	842		3479.10			4432.586

Appendix 2

Table 4: Estimation of Total Number of Variance and Professional Farmers in the 3rd and 4th Stages of a four – stage Cluster Sampling

Table 4a

S/No	M_{ij}	m_{ij}	$\bar{y}_{ij}$	$M_{ij}\bar{y}_{ij}$	S_{3ij}^2	f_{3ij}	$M_{ij}^2 \frac{1-f_{3ij}}{m_{ij}} S_{3ij}^2$
1.	115	52	2.20	253.00	2.44	0.452	340.066
2.	155	70	2.31	358.05	1.47	0.451	276.984
3.	87	39	1.24	107.88	1.20	0.448	128.557
4.	73	33	1.57	114.61	2.04	0.452	180.527
5.	49	22	1.82	89.18	1.23	0.449	73.965
6.	69	31	3.00	200.79	4.38	0.492	339.383
7.	67	30	2.91	194.97	0.70	0.448	57.818
8.	120	54	1.12	134.40	1.66	0.450	243.467
9.	40	18	1.45	58.00	0.60	0.450	29.333
10.	67	30	1.60	107.20	1.27	0.448	104.899
	842	179		1618.08			1774.999

Table 4b

S/No of EA	M_{ijk}	m_{ijk}	$\bar{y}_{ijk}$	$M_{ijk}\bar{y}_{ijk}$	S_{4ijk}^2	f_{4ijk}	$M_{ijk}^2 \frac{1-f_{4ijk}}{m_{ijk}} S_{4ijk}^2$
1.	52	7	2.15	111.80	2.34	0.135	781.881
2.	70	9	2.41	168.70	1.43	0.129	678.122
3.	39	5	1.77	69.03	1.24	0.128	328.925
4.	33	4	1.54	50.82	2.55	0.121	610.235
5.	22	3	1.37	30.14	1.13	0.136	157.513

6.	31	4	3.01	93.31	2.91	0.129	608.940
7.	30	4	2.11	63.30	0.76	0.133	148.257
8.	54	7	1.36	73.44	1.44	0.129	522.481
9.	18	2	1.49	26.82	0.38	0.111	54.727
10.	30	4	1.62	48.60	1.88	0.133	366.741
	179	49		735.96			4257.822

Appendix 3

Table 5: Continuation of Table 4

S/N of EA	$\frac{M_i}{m_i}$	$M_{ij}^2 \frac{1-f_{3ij}}{m_{ij}} S_{3ij}^2$	$\frac{M_i}{m_i} \times M_{ij}^2 \frac{1-f_{3ij}}{m_{ij}} S_{3ij}^2$
1.	2.174	340.066	739.3035
2.	2.387	276.984	661.1608
3.	2.172	128.557	279.2258
4.	2.178	180.527	393.1878
5.	2.163	73.965	159.5307
6.	2.174	339.383	737.8186
7.	2.179	57.818	125.9854
8.	2.175	243.467	529.5407
9.	2.175	29.333	63.7993
10.	2.179	104.889	228.5553
			3918.543571

Appendix 4

Table 6: Continuation of Table 5

S/N of EA	$\frac{M_i}{m_i}$	$\frac{M_{ij}}{m_{ij}}$	$M_{ijk}\bar{y}_{ijk}$	$\frac{M_i}{m_i}\frac{M_{ij}}{m_{ij}} M_{ijk}\bar{y}_{ijk}$	$M_{ijk}^2 \frac{1-f_{4ijk}}{m_{ijk}} S_{4ijk}^2$	$\frac{M_i}{m_i}\frac{M_{ij}}{m_{ij}} M_{ijk}^2 \frac{1-f_{4ijk}}{m_{ijk}} S_{4ijk}^2$
1	2.174	2.212	111.8	537.63	781.88	3759.98
2	2.387	2.214	168.7	891.55	678.12	3583.75
3	2.172	2.231	69.03	334.5	328.93	1593.88
4	2.178	2.212	50.82	244.84	610.24	2939.95
5	2.163	2.227	30.14	145.18	157.51	758.74
6	2.174	2.226	93.31	451.56	608.94	2946.86
7	2.179	2.233	63.3	308	148.26	721.375
8	2.175	2.222	73.44	354.92	522.48	2525.07
9	2.175	2.222	26.82	129.62	54.727	264.487
10	2.179	2.233	48.6	236.47	366.74	1784.45
	21.929	22.222		3634.276585		20878.55033