MULTIPLE FRAME SAMPLING THEORY AND APPLICATIONS

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

OF

MIDDLE EAST TECHNICAL UNIVERSITY

BY

AYLİN DALÇIK

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR

THE DEGREE OF MASTER OF SCIENCE

IN

STATISTICS

FEBRUARY 2010

Approval of the thesis:

MULTIPLE FRAME SAMPLING THEORY AND APPLICATIONS

submitted by **AYLİN DALÇIK** in partial fulfillment of the requirements for the degree of **Master of Science in Statistics Department, Middle East Technical University** by**,**

Prof. Dr. Canan Özgen ____________
Director, Graduate School of **Natural and Applied Sciences**

Prof. Dr. H. Öztaş Ayhan ____________
Head of Department, **Statistics**

Prof. Dr. H. Öztaş Ayhan ____________
Supervisor, **Statistics, METU**

Examining Committee Members:

Prof. Dr. Ayşen Akkaya ________
Statistics Department, METU

Prof. Dr. H. Öztaş Ayhan ________
Statistics Department, METU

Assist. Prof. Dr. A. Sinan Türkyılmaz ________
Dept. of Technical Demography, Hacettepe Institute of Population Studies

Dr. Ceylan Yozgatlıgil ________
Statistics Department, METU

Dr. Vilda Purutçuoğlu ________
Statistics Department, METU

Date: 03.02.2010

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : Aylin Dalçık

Signature :

ABSTRACT

MULTIPLE FRAME SAMPLING THEORY AND APPLICATIONS

Dalçık, Aylin

Master of Science, Department of Statistics
Supervisor: Prof. Dr. H. Öztaş Ayhan

February 2010, 125 pages

One of the most important practical problems in conducting sample surveys is the list that can be used for selecting the sample is generally incomplete or out of date. Therefore, sample surveys can produce seriously biased estimates of the population parameters. On the other hand updating a list is a difficult and very expensive operation.

Multiple-frame sampling refers to surveys where two or more frames are used and independent samples are taken respectively from each of the frames. It is assumed that the union of the different frames covers the whole population. There are two major reasons for the use of multiple-frame sampling method. One is that, using two or more frames can cover most of the target population and therefore reduces biases due to coverage error. The second is that multiple-frame sampling design may result in considerable cost savings over a single frame design.

Key words: Multiple frames, overlapping frames, coverage, cost-efficiency

ÖZ

ÇOKLU ÇERÇEVE ÖRNEKLEME TEORİSİ VE UYGULAMALARI

Dalçık, Aylin

Yüksek Lisans, İstatistik Bölümü

Tez Yöneticisi: Prof. Dr. H. Öztaş Ayhan

Şubat 2010, 125 sayfa

Örnekleme araştırmalarındaki en önemli uygulama problemlerinden biri seçilen örnekler için kullanılan listenin genellikle tamamlanmamış olması veya verinin güncel olmamasıdır. Bunun sonucunda, örnekleme araştırmaları ciddi şekilde kitle parametrelerine ilişkin yanlı tahminler üretebilir. Listenin güncellenmesi zor ve çok pahalı bir işlemdir.

Çoklu çerçeve örneklemesi iki veya daha fazla çerçevenin kullanıldığı ve her bir çerçeveden bağımsız örneklerin çekildiği araştırmaları ifade eder. Farklı çerçevelerin birleşiminin tüm kitleyi kapsadığı varsayılır. Çoklu çerçeve örnekleme yönteminin kullanımının iki önemli sebebi vardır. Bir tanesi, iki veya daha fazla çerçevenin kullanımı, hedef kitleyi kapsayabilir ve neticesinde kapsama hatası yüzünden oluşabilecek yanlılığı azaltır. İkincisi, çoklu çerçeve örnekleme planı tek çerçeveli plana gore önemli maliyet tasarrufuna sebep olabilir.

Anahtar kelimeler: Çoklu çerçeveler, örtüşen çerçeveler, kapsam, uygun maliyet

ACKNOWLEDGEMENTS

I would like to thank my supervisor Prof. Dr. Öztaş Ayhan who gave me the opportunity to complete my thesis. His comments and encouragement were really encouraging for me.

I am grateful to my friend Onur Pekcan, my sister Assoc. Prof. Dr. Ayşegül Erençin and my brother in law Assoc. Prof. Dr. Arif Erençin for helping to find articles related with my thesis.

I owe my heartful thanks to my husband Hakan Dalçık. His great understanding and continuous support have helped me to put forth my full effort and to overcome all the difficulties that I came across during the study.

Very special thank is for my daughter İlsu Dalçık for her patience and understanding during the study.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

The list of population elements is called a sampling frame. To do sampling, list of population elements must be available or to be prepared. The sampling frame or list is the keystone around which a sampling process constructed. The frame can be phonebook, student record index, population index of record, voter lists, maps, etc.

The frame is perfect, if it covers all elements of the target population. In other words, sampling frame and selected sample elements must have one to one correspondence. But perfect frames are rare, and we must often use frames with serious deficiencies that must be detected and remedied.

There are four basic frame problems discussed by Kish (1965).

1. Missing elements, noncoverage, incomplete frame
2. Cluster of elements together in one listing
3. Blanks or foreign elements
4. Duplicate listing

Kish (1965) mentioned three general ways of avoiding the problems.

a. Ignore and disregard the problem if it is known to be small compared to other errors, and if correcting it would be too costly.

b. Redefine the population to fit the frame. It can be used if the result of the redefinition is trivial or preferred. For example, a firm`s payroll list may exclude recently hired employees; but this may be few and the researcher may prefer to exclude them.

c. Correct the entire population list. This means finding all missing elements, splitting each cluster, and eliminating all blanks, foreign elements and duplicate listings.

Kish (1965) proposed specific treatments for the four basic problems:

1. Missing elements

It is also called noncoverage and incomplete frame. There are two solutions for this problem.

- A supplement listing in a separate stratum for the missed elements may be formed to provide for their separate selection.
- Linking procedures append uniquely the selection of the missed elements to specified listings. Linking procedure is a convenient device when a separate stratum is too costly and the missed elements are scattered individually or in small clusters.

2. Cluster of elements can appear together, associated with single listings. In this case, three solutions can be used.

   - If the clusters are rare and small, include all the elements that occur with each selected listing.
   - Select one element from the cluster, at random, and weight it up with the number of elements in the cluster.
   - Relist a larger sample and then select from it an epsem of elements.

3. Blanks or foreign elements occur when some listings contain no elements of the target population. In this case the selected blanks must be rejected and omitted, because they contribute no element to the sample.

4. Duplicate listings can be treated as follows:

  - Problem can be solved by weighting each case by the inverse of its number of listings. If the element has P_i listings, assign the weight $1/P_i$ to its selection and include duplicate selections.
  - Unique identification of a single listing for each element may be determined at the time of the selection.
  - When the population consists of more than one frame, in which some units may be duplicated. For each sample case a search of all preceding frames must be made, and the selection thrown if a preceding duplicate is found.

Multiple frame surveys generally used to decrease cost of sampling or reduce undercoverage. A common example is a dual frame survey where one frame can be sampled cheaply but does not cover the whole target population, however the other frame complete coverage but it is expensive to sample. Supposing that complete frame such as a household address list is available, it is more efficient to take sample reduced size from the complete frame and supplement the sample by additional data taken from other frame such as telephone directories might be incomplete but less expensive to sample from. Figure 1 displays overlapping two frames, frame A that covers the whole population is very expensive to sample and frame B that does not cover the whole population but is cheaper to sample.

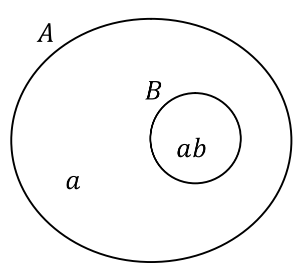


Figure 1: Overlapping Frames A and B and Two Domains, Because Frame A is Complete and Frame B is Incomplete.

Figure 2 displays overlapping three frame in which frame A is complete but frame B and C are incomplete. Hartley (1962) proposed the general methodology of multiple frame sampling. His applications concentrated on the circumstances where one frame was complete but expensive to sample, other frames were cheap to sample but incomplete. Hartley (1974) used multiple frame estimators for estimation of Crop and Livestock items in Texas. Frame A was the list frame of "farm operators" compiled from the country ASCS (Agricultural Stabilization and Conservation service of the U.S.) offices eliminating duplications of operators within the county (operators with several ASCS contracts) and between counties (operators operating in several counties). Frame B was the list of "tract farmers"; that is, operating a tract contained within the segment boundaries with regard to all operations conducted on the tracts. In this example frame B has 100% coverage but frame A only cover about 70% to 95% of the farm operators.

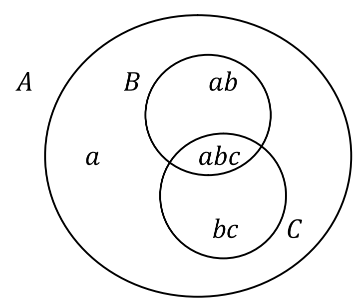


Figure 2: Overlapping Three Frames and Four Domains, Because Frame A is Complete and Frames B and C are Incomplete.

Under coverage is one of the most common problems of sampling frames. To reduce the effect of coverage error on survey estimates several frames can be combined in order to get a complete (or nearly complete) coverage of the target population. Multiple frame estimators have been developed to be used in the circumstances of multiple frame surveys (Maia and Vicenta 2009). Sampling frames may overlap; it means a single unit of the sample frame is related with more than one element of the target population. Figure 3 displays a two overlapping frames in which each frame is incomplete and Figure 4 displays a three overlapping frames in which all frames are incomplete. Kalton and Anderson (1986) described a multiple frame survey of persons with AIDS, one sample was taken from the frame for a general population health survey, meantime independent samples were taken of clients of sexually transmitted disease clinics, drug treatment centers, and ospitals. The multiple frame approach would be expected more accurate estimates of persons with AIDS, because the last three frames would provide larger numbers of persons with AIDS for the sample. Iachan and Dennis (1993) described a multiple frame survey of the homeless population where the frames were homeless shelters, soup kitchens, and street areas. Not all homeless persons visit the homeless shelters, so using more frames probably reduces the undercoverage bias.

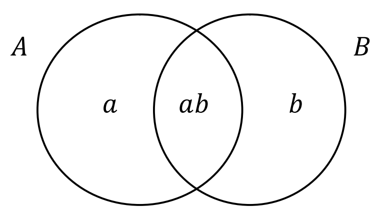


Figure 3: Overlapping Frames A and B and Three Domains. All Frames are Incomplete.

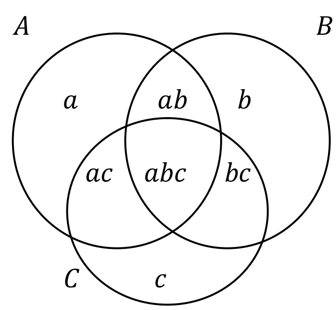


Figure 4: Overlapping Frames A, B and C and Seven Domains. All Frames are Incomplete.

In this study, making estimates without making a single list will be discussed when the population consists of more than one frame in which some elements may be duplicated. Because combining lists can not be possible. For example, when it is made a research about internet banking users in Ankara, combining lists of internet banking users taken each of different banks to provide perfect frame can not be possible. People having same name and surname can be seen in records of different banks. If those people have different address, it is possible to identify as different people. In such cases, it is difficult to determine duplications. In such circumstances multiple frame approach can be good solution.

In chapter 2, general description of multiple frame survey was made and coverage and matching errors were explained in chapter 3. In chapter 4 and 5, multiple frame estimators were discussed. In chapter 6, application with artificial data was made in order to compare and discuss these estimators. In chapter 7, conclusions and recommendations were made about use of multiple frame estimators.

CHAPTER 2

MULTIPLE FRAME SURVEY

When multiple frames are used, it is likely that some members of the target population will be included on more than one frame. This situation is called “matching frames” or “overlapping frames” in the literature (Kish 1965). The representation of some population members on more than one frame gives rise to the frame problem of duplicates.

There are two basic approaches for handling duplicates (Kalton and Anderson 1986).

- To redefine the frames so that they are non-overlapping
- To make compensations in the analysis.

The replications can be ignored if they are known to be a small henough proportion, or if the target population is redefined to include them. If it is not too costly, we can eliminate replications from all lists. Kish (1965) proposed a procedure for eliminating overlaps. These procedures can be represented in the following Table 1 for four frames.

Table 1: Procedure of Unique Identification

Frames	1	2	3	4
Number of Unique Elements	M_1	M_2	M_3	M_4
Number of Replicated Elements	0	M_{12}	M_{23} + $M_{13.2}$	M_{34} + $M_{24.3}$ + $M_{14.23}$
Number of Population Elements	N_1	N_2	N_3	N_4

Source: Kish (1965)

From the first frame, we accept all N_1 listings as the M_1 unique elements of the first frame. From the N_2 listings of the second frame, we eliminate the M_{12} listings which also appear on the first frame, and obtain the M_2 unique elements of the second frame. From the third frame of N_3, we eliminate M_{23} which appear on the second frame, and $M_{13.2}$ which appear on the first, though not on the second. From the fourth frame, we eliminate those found on the third, those on the second but not on the third, and those on the first but neither on the third nor on the second, and so on.

To save effort, this procedure of unique identification can also be carried out on a sample basis. From the first frame, all $n_1 = N_1 f_1$ selections are accepted. where $f_h = n_h / N_h$, h = 1,2,3,4. The $n_2 = N_2 f_2$ selections from the second frame are searched for duplications among the population N_1 of the first frame, and any found are eliminated; but this number m_{12} will be variant around its expected value $M_{12} f_2$. Hence $m_2 = n_2 - m_{12}$ is also a variate which can be regarded as the subclass that appears as the proportion $\bar{m}_2 = m_2 / n_2$ from the second frame. Similarly from the $n_3 = N_3 f_3$ selections of the third frame, a search of the second frame eliminates m_{23}, and a search of the first frame a further $m_{13.2}$; this leaves the random subclass proportion of $\bar{m}_3 = m_3 / n_3$ from the third frame. And so on to the fourth and later frames. The unique listings $m_1 + m_2 + m_3 + \cdots = \sum m_h$ denote a stratified sample of subclasses, except for the first stratum in which $\bar{m}_1 = m_1 / n_1 = 1$. Now the formulas of stratified subclasses can be applied.

Within the strata, the simple random samples of n_h elements result in simple random samples of the m_h of the subclass members. Thus, when the weights $W_h^* = M_h / M$ are known constants, the general stratified random formulas for the variance can be applied to the mean (Kish 1965).

$\bar{y}_w = \sum_h^{M_h} (M_h / M) \sum_i^{m_h} y_{hi} / m_h$ and its variance is

$$var(\bar{y}_w) = \sum W_h^{*2} (1 - f_h) \frac{s_h^2}{m_h}$$

Frequently M_h is not known. But we just know the random variable m_h of subclass members that are found among the n_h elements of the random sample selected in the h.th stratum. The proportion of subclass members, $\bar{m}_h = m_h / n_h$, is the binomial mean of the count variable M_{hi} that takes on the values

$m_{hi} = n_{hi} = 1$ for the m_h members of the subclass,

$m_{hi} = 0$ for the $(n_h - m_h)$ non members.

The computations and derivations for subclasses involve a similar auxiliary variable Y_{hi} for the variable X_{hi}

$y_{hi} = x_{hi}$ for the members of the subclass,

$y_{hi} = 0$ for the $(n_h - m_h)$ nonmembers.

The stratum total is

$y_h = \sum_i^{n_h} y_{hi} = \sum_i^{m_h} y_{hi}$ and

The stratum mean is

$$\bar{y}_h = \frac{y_h}{\bar{m}_h n_h} = \frac{y_h}{m_h} = \frac{\sum_i^{m_h} y_{hi}}{\sum_i^{m_h} m_{hi}}.$$

We want to estimate the subclass total $Y = \sum_h^H \sum_i^{M_h} Y_{hi}$ for the M subclass members in the population and the subclass mean $\bar{Y} = Y/\text{M}$. The subclass total can be estimated by

$$\tilde{Y}_w = \sum N_h \frac{y_h}{n_h} = \sum F_h y_h \qquad (2.1)$$

The subclass sample mean is

$$\bar{y}_w = \frac{\sum N_h y_h / n_h}{\sum N_h m_h / n_h} = \frac{\sum F_h y_h}{\sum F_h m_h} = \sum w_h \bar{y}_h, \qquad (2.2)$$

where $w_h = F_h m_h / \sum F_h m_h$ and $F_h = N_h / n_h$.

Kish (1965) proposed the variance for a subclass total $\tilde{Y}_w$. It is

$$var\ (\tilde{Y}_w) = \sum(1 - f_h) F_h^2 \frac{n_h}{n_{h-1}} \left[\sum_i^{m_h} y_{hi}^2 - \frac{y_h^2}{n_h}\right]$$

$$= \sum(1 - f_h) \frac{N_h^2}{n_h - 1} \bar{m}_h \left[v_h^2 + (1 - \bar{m}_h)\bar{y}_h^2\right] \qquad (2.3)$$

where $v_h^2 = \frac{1}{m_h} \sum_i^{m_h} (y_{hi} - \bar{y}_h)^2 \qquad (2.4)$

In a related example, procedure of Kish (1965) can be shown below. This example was made with small frames on Table 2 to show how the procedure worked. Samples s_A, s_B and s_C were selected from the population frames A, B and C respectively.

Table 2: Artificial Population Frames A, B and C

FRAME A			FRAME B			FRAME C		
Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element y_i	Pop. ID i	ID Label	Pop. Element Value y_i
1	Aylin	1	1	Suheyla	1	1	Suheyla	1
2	Suheyla	1	2	Vural	1	2	Aylin	1
3	Arif	2	3	Aysegul	0	3	Berrin	1
4	Aynur	3	4	Berrin	1	4	Hakan	2
5	Ayse	5	5	Mahmut	3	5	Ozgur	5
6	Hakan	2	6	Murat	0	6	Necla	2
7	Ebru	2	7	Hacer	6	7	Haydar	6
8	Ozcan	3	8	Ayse	5	8	Zehra	3
9	Birsen	4	9	Aylin	1	9	Murat	0
10	Kamil	2	10	Ebru	2	10	Kaya	6
11	Aysegul	0	11	Didem	1	11	Nurettin	2
12	Deniz	5	12	Tuna	7	12	Aysenaz	0
13	Nilgun	2	13	Hakan	2	13	Nurcan	2
14	Zehra	3	14	Zehra	3	14	Hacer	6
15	Sule	4	15	Ozcan	3	15	Umit	5
16	Ahmet	2	16	Necla	2	16	Tuna	7
17	Bora	1	17	Ismail	4	17	Osman	3
18	Hacer	6	18	Eda	3	18	Gizem	3
19	Nurettin	2	19	Ferit	2	19	Vural	1
20	Asli	1	20	Bora	1	20	Ilsu	0
21	Vural	1	21	Nurettin	2	21	Banu	2
22	Haydar	6	22	Gizem	3			
23	Tuna	7	23	Nurcan	2			
24	Ozlem	4	24	Doga	4			
25	Berrin	1	25	Haydar	6			
26	Umit	5						
27	Ilsu	0						
28	Aysenaz	0						
29	Osman	3						
30	Mehmet	1						

An application to this method was made with samples given in Table 3. These independent simple random samples s_A, s_B and s_C selected from the population frames A, B and C respectively. The sizes of samples were n_A = 10, n_B = 8 and n_C = 7.

Table 3: Independent simple random samples s_A, s_B and s_C

Sample A s_A				Sample B s_B				Sample C s_C			
Sample ID j	**Pop. ID i**	**ID Label**	**Sample Element Value y_i**	**Sample ID j**	**Pop. ID i**	**ID Label**	**Sample Element Value y_i**	**Sample ID j**	**Pop. ID i**	**ID Label**	**Sample Element Value y_i**
1	1	Aylin	1	1	11	Didem	1	1	10	Kaya	6
2	30	Mehmet	1	2	1	Suheyla	1	2	2	Aylin	1
3	22	Haydar	6	3	23	Nurcan	2	3	9	Murat	0
4	10	Kamil	2	4	14	Zehra	3	4	5	Ozgur	5
5	28	Aysenaz	0	5	4	Berrin	1	5	3	Berrin	1
6	17	Bora	1	6	9	Aylin	1	6	12	Aysenaz	0
7	7	Ebru	2	7	6	Murat	0	7	18	Gizem	3
8	26	Umit	5	8	10	Ebru	2				
9	14	Zehra	3								
10	4	Aynur	3								

Procedure of unique identification was made on a sample basis and replications were eliminated from all listings s_A^*, s_B^* and s_C^* were found. These unique elements can be seen in Table 4 and replicated elements can be seen in Table 5.

Table 4: Unique Elements Based on Samples s_A, s_B and s_C

Unique elements in Sample A s_A^*				Unique elements in Sample B s_B^*			
Sample ID j	**Pop. ID i**	**ID Label**	**Sample Element y_i**	**Sample ID j**	**Pop. ID i**	**ID Label**	**Sample Element y_i**
1	1	Aylin	1	1	11	Didem	1
2	30	Mehmet	1	3	23	Nurcan	2
3	22	Haydar	6	7	6	Murat	0
4	10	Kamil	2	**Unique elements in Sample C s_C^***			
5	28	Aysenaz	0				
6	17	Bora	1	**Sample ID j**	**Pop. ID i**	**ID Label**	**Sample Element y_i**
7	7	Ebru	2				
8	26	Umit	5				
9	14	Zehra	3	1	10	Kaya	6
10	4	Aynur	3	4	5	Ozgur	5

Table 5: Replications Based on Samples s_A, s_B and s_C

Sampled elements from frame B .It is also in frame A S_{AB}				
Sample ID j	**Pop. ID i (A)**	**i (B)**	**ID Label**	**Sample Element y_i**
2	2	1	Suheyla	1
5	25	4	Berrin	1
6	1	9	Aylin	1
8	7	10	Ebru	2
4	14	14	Zehra	3

Sampled elements from frame C. It is also in frame B. S_{BC}				
Sample ID j	**Pop. ID i (B)**	**i (C)**	**ID Label**	**Sample Element y_i**
2	9	2	Aylin	1
3	6	9	Murat	0
5	4	3	Berrin	1
7	22	18	Gizem	3

Sampled elements from frame C .It is also in frame A but not in frame B $s_{AC.B}$				
Sample ID j	**Pop. ID i (A)**	**i (C)**	**ID Label**	**Sample Element y_i**
6	28	12	Aysenaz	0

After eliminate the replications from all lists the number of unique elements, replications summarized in the Table 6.

Table 6: Number of replications and unique elements based on samples s_A, s_B and s_C

Samples	***A***	***B***	***C***
Number of UniqueElements in Sample	10	3	2
Number of Replicated Elements	0	5	4 + 1 = 5
Number of Sample Elements	n_A=10	n_B =8	n_C = 7
$\bar{m}_h$	$\bar{m}_A$=10/10=1	$\bar{m}_B$=3/8=0,375	$\bar{m}_C$=2/7=0,285
y_h	24	3	11

$$\hat{Y}_w = \sum N_h \frac{y_h}{n_h} = \sum F_h y_h = \frac{30}{10}\,24 + \frac{25}{8}\,3 + \frac{21}{7}\,11 = 114.36$$

Making a single list requires the matching of listings across frames. This process is error prone generally. Problems such as misspelt names and alternative addresses cause failures to match and mismatches. Biases arising from these problems need to be taken into account.

Multiple frames in sample surveys has been examined by several author such as Hartley (1962 and 1974), Lund (1968), Fuller and Burmeister (1972), Bankier (1986), Skinner (1991), Skinner and Rao (1996) and Lohr and Rao (2000). Different estimators were proposed whether domain and frame sizes are known or unknown. In most cases, the domain sizes are not known or only approximately known, due to the use of out of date information and lists, which makes difficult to determine whether a unit belongs to any other frame. In such a case, the estimator of the population total is biased and the bias remains constant as the sample size increases.

A general approach for overlapping frames introduced by Hartley (1962 and 1974) and followed by the others. Although literature has mostly dealt with the dual frame case, a theoretical framework for multiple frame surveys has been given by Lohr and Rao (2006) as follows:

F denotes the index of frames, $F = \{1,2,\ldots\ldots,Q\}$. They assumed that the union of the frames covers the target population. There are possible 2^{Q-1} domains with Q frames. For $K \sqsubseteq F$, the domain defined by K is

$D_K = (\bigcap_{i\in K} A_i) \cap \left(\bigcap_{i\notin K} A_j^c\right)$ where c denotes complementation.

For example, $Q = 3$ and there are 7 domains D_K denoted by index sets $K = \{(a),(b),(c),(ab),(ac).(bc),(abc)\}$ (see in Figure 3).

$N^{(q)}$ is the number of population units in frame A_q, and N_K is the number of population units in domain K. The frame sizes $N^{(q)}$ are often known but the domain sizes N_K are often unknown. In this context, the population total Y can be written as the sum of the population totals in distinct domains,

$$Y = \sum_{K\subseteq F} Y_K \tag{2.5}$$

where $$Y_K = \sum_{i=1}^{N} \delta_i(K) y_i \tag{2.6}$$

and $\delta_i(K)$ is an indicator of domain variable:

$$\delta_i(K) = \begin{cases} 1, & i \in D_k \\ 0, & otherwise \end{cases}$$

The number of population units in domain D_k is a special case of Y_K when y_i = 1 for all i.

$$N_K = \sum_{i=1}^{N} \delta_i(K) \tag{2.7}$$

s_q denote the probability sample from frame A_q,q = 1,2,…….Q. The multiple frame survey has K independent estimators of Y_K. Estimation of Y_K using the sample is given by

$$\hat{Y}_K^{(q)} = \sum_{i \in s_q} w_i^{(q)} \delta_i(K) y_i \tag{2.8}$$

The weights $w_i^{(q)}$ are the appropriate weights used for estimation from frame A_q. Similarly the estimators of domain sizes N_K is given by

$$N_K^{(q)} = \sum_{i \in s_q} w_i^{(q)} \delta_i(K). \tag{2.9}$$

In literature there are three main approaches to estimate of population total from setting weights $w_i^{(q)}$ in (2.8) which is given by Mecatti (2007).

(a) The Optimum Estimator $w_{i,opt}^{(q)}$ proposed by Hartley (1962), Lund (1968) and Fuller and Burmeister (1972) has minimal variance but, in operationally is very complex.

(b) The Single Frame Estimator $w_{i,SF}^{(q)}$ proposed by Bankier (1986), Kalton and Anderson (1986) and Skinner (1991) uses fixed weights guaranteeing unbiased estimates. However, they are less efficient than the optimal estimator.

(c) The Pseudo Maximum Likelihood Estimator $w_{i,PML}^{(q)}$ proposed by Skinner and Rao (1996) and Lohr and Rao (2000) extends the applicability of the optimal estimator. It increases its efficiency when compared with the single based estimator.

CHAPTER 3

COVERAGE AND MATCHING ERRORS

In surveys, the sample is selected from a sampling frame of all population members. An inadequate frame leads to coverage errors. An error in coverage occurs when there is an exclusion or duplication of the elements in the population or sample. Exclusions are referred to as under-coverage, while duplication and wrongful inclusions are called over- coverage. These errors are caused by defects in the sampling frame: inaccuracy, incompleteness, duplication, inadequacy and obsolescence. Many researchers use these kind not up-to-date sampling frames. This can lead to serious bias.

Coverage errors may also occur in field procedures (e.g., a survey is conducted, but the interviewer misses several households or persons due to poor transportation facilities or severe weather conditions).

Matching error is another source of error which may be also associated with the coverage error when mismatching occurs. In the following sections recent literature on coverage error and matching error are examined.

3.1 Coverage Errors

The population coverage errors were defined by Ayhan (2000) and Ayhan and Ekni (2003) as below:

N is the size of the target population,

N_i is the size of available frame population for source i ,

$\Delta_i = N - N_i$, i = 1,2, … ..

where $N > N_i$ refers to under-coverage error, $N < N_i$ refers to over-coverage error and $N = N_i$ refers to no coverage error.

Several methods have been proposed to solve the problem of coverage errors. The most popular method of these is dual-record system estimation. Dual-records system (DSR) estimator was first proposed by Chandra Sekar and Deming 1949 and was developed by many researchers such as Casady, Nathan, Sirken (1985), Hogan (1990, 1993a, and 1993b) Isaki (1992) and Ayhan (2000),

This method is based on collecting and matching data for the same units of observation, from two independent data collection sources. These two data sources may be any combination of registration, census or sample surveys (Ayhan 2000). The main idea is to compare the results of two sources, generally the census or household survey and a generated source and try to estimate the total population. Application of dual record system estimation is presented by a $2x2$ contingency table as in Table 7.

Table 7: Data Layout Structure of a Dual-Record System

	Data Source 2		
Data Source 1	Reported	Not reported	Total
Reported	n_{11}	n_{12}	n_{1+}
Not reported	n_{21}	n_{22}	n_{2+}
Total	n_{+1}	n_{+2}	n

Source: Ayhan (2000)

where

n = the total number of events

n_{11} = the total number of events reported by both data sources 1 and 2

n_{12} = the total number of events reported by only data source 1

n_{22} = the total number of events not reported by both data sources 1 and 2

n_{21} = the total number of events reported by only data source 2.

The Chandra Sekar - Deming (CD) estimator is given by Ayhan (2000) as below:

$$\hat{n}^{(CD)} = n_{11} + n_{12} + n_{21} + \hat{n}_{22}^{(CD)} \qquad (3.1.1)$$

where $\hat{n}_{22}^{(CD)} = n_{12}n_{21}/n_{11}$ (3.1.2)

By using the information on marjinal total, the same estimator can be presented in the following form (Ayhan 2000).

$$\hat{n}^{(CD)} = (n_{11} + n_{12})(n_{11} + n_{21})/n_{11} = n_{1+}n_{+1}/n_{11} \qquad (3.1.3)$$

The structure of the layout of data for the DRS can be altered to accommodate the new proposed estimators as shown in Table 8 and Table 9. The unmatched cases from both data sources for the cells of the original layout were further divided as missed as a result of undercoverage (UC) and not reported as a result of non-response (NR).

Table 8: Layout of the Breakdown of Error Components for the Proposed Estimator When $n_1 = n_2$

Data Source 2				
		Not Matched		
Data Source 1	Matched (reported)	Not reported NR ($\boldsymbol{n_2}$)	Missed UC ($\boldsymbol{n_2^*}$)	Total
Matched (reported	$\boldsymbol{n_{11}}$	$\boldsymbol{n_{12}}$		$\boldsymbol{n_{1+}}$
Not reported NR ($\boldsymbol{n_1}$)	$\boldsymbol{n_{21}}$	$(\boldsymbol{n_1 - n_{1+}})\Lambda(\boldsymbol{n_2 - n_{+1}})$	$(\boldsymbol{n_1 - n_{1+}})\Lambda(\boldsymbol{n_2^* - n_2})$	$\boldsymbol{n_1 - n_{1+}}$
Missed UC ($\boldsymbol{n_1^*}$)	$\boldsymbol{n_{21}}$	$(\boldsymbol{n_1^* - n_1})\Lambda(\boldsymbol{n_2 - n_{+1}})$	$(\boldsymbol{n_1^* - n_1})\Lambda(\boldsymbol{n_2^* - n_2})$	$\boldsymbol{n_1^* - n_1}$
Total	$\boldsymbol{n_{+1}}$	$\boldsymbol{n_2 - n_{+1}}$	$\boldsymbol{n_2^* - n_2}$	$\boldsymbol{\hat{n}}$

Notes: n_1 and n_2 relate to the matched and not reported cases from the data sources 1 and 2 respectively. n_1^* and n_2^* relate to the matched and not matched (both not reported and missed) cases from data sources 1 and 2 respectively.

Source: Ayhan (2000)

The total number of events obtained from data source 1 is n_1, with

$$\hat{n}_1 = n_{1+} + \text{NR}\ (n_1) \quad (3.1.4)$$

The total number of events obtained from data source 2 is n_2, with

$$\hat{n}_2 = n_{2+} + \text{NR}\ (n_2) \quad (3.1.5)$$

The total number of events that should have been obtained from data source 1 is n_1^*, with

$$\hat{n}_1^* = n_{1+} + \text{NR}\ (n_1) + (n_1^* - n_1) \quad (3.1.6)$$

The total number of events that should have been obtained from data source 2 is n_2^*, with

$$\hat{n}_2^* = n_{2+} + \text{NR}\ (n_2) + (n_2^* - n_2) \quad (3.1.7)$$

Ayhan (2000) proposed the following estimators.

(i) When the selected sample sizes are equal ($n_1 = n_2$), estimators based on data source 1 is given by

$$\hat{n}^{(A)} = (n_{11} + n_{21}) + n_{12} + \{[n_1 + (n_1 + n_{1+}) + (n_1^* - n_1)] - n_{1+}\} - n_{21} \quad (3.1.8)$$

$$\hat{n}^{(A)} = n_{1+} + n_{12} + [(n_1^* - n_{1+}) - n_{21}] \quad (3.1.9)$$

where $\hat{n}_{22}^{(A)} = [(n_1^* - n_{1+}) - n_{21}]$

The same estimators can also be determined on the basis of data source 2, as follows:

$$\hat{n}^{(A)} = (n_{11} + n_{12}) + n_{21} + \{[n_{+1} + (n_2 + n_{+1}) + (n_2^* - n_2)] - n_{+1}\} - n_{12} \quad (3.1.10)$$

$$\hat{n}^{(A)} = n_{1+} + n_{21} + [(n_2^* - n_{+1}) - n_{12}] \quad (3.1.11)$$

where $\hat{n}_{22}^{(A)} = [(n_2^* - n_{+1}) - n_{12}]$ (3.1.12)

(ii) When the selected sample sizes are equal ($n_1 \neq n_2$)

Table 9: Layout of the Breakdown of Error Components for the Proposed Estimator When $n_1 \neq n_2$

Data Source 2				
	Not Matched			
Data Source 1	Matched (reported)	Not reported NR ($\boldsymbol{n_2}$)	Missed UC ($\boldsymbol{n_2^*}$)	Total
Matched (reported	$\boldsymbol{n_{11}}$	$\boldsymbol{n_{12}}$	0	$\boldsymbol{n_{1+}}$
Not reported NR ($\boldsymbol{n_1}$)	$\boldsymbol{n_{21}}$	$(\boldsymbol{n_1 - n_{1+}}) \wedge (\boldsymbol{n_2 - n_{+1}})$	$(\boldsymbol{n_1 - n_{1+}}) \wedge (\boldsymbol{n_2^* - n_2})$	$\boldsymbol{n_1 - n_{1+}}$
Missed UC ($\boldsymbol{n_1^*}$)	$\boldsymbol{0}$	$(\boldsymbol{n_1^* - n_1}) \wedge (\boldsymbol{n_2 - n_{+1}})$	$(\boldsymbol{n_1^* - n_1}) \wedge (\boldsymbol{n_2^* - n_2})$	$\boldsymbol{n_1^* - n_1}$
Total	$\boldsymbol{n_{+1}}$	$\boldsymbol{n_2 - n_{+1}}$	$\boldsymbol{n_2^* - n_2}$	$\boldsymbol{\hat{n}}$

Notes: n_1 and n_2 relate to the matched and not reported cases from the data sources 1 and 2 respectively.n_1^* and n_2^* relate to the matched and not matched (both not reported and missed) cases from data sources 1 and 2 respectively.

Source: Ayhan (2000)

The number of events determined from the data source 1 is

$$\hat{n}_1^* = n_{1+} + (n_1 - n_{1+}) + (n_1^* - n_1) \qquad (3.1.13)$$

$$= n_{1+} + \text{NR}\ (n_1) + \text{UC}\ (n_1^*)$$

The number of events determined from the data source 2 is

$$\hat{n}_2^* = n_{+1} + (n_2 - n_{+1}) + (n_2^* - n_2) \qquad (3.1.14)$$

$$= n_{+1} + \text{NR}\ (n_2) + \text{UC}\ (n_2^*)$$

Thus, the alternative estimator was proposed by Ayhan (2000) as follows:

$$\hat{n}^{(A)} = Max\ (\hat{n}_1^*, \hat{n}_2^*) \qquad (3.1.15)$$

There are two approaches to overcome coverage error. One is to use up-to-date frames. But updating a frame is a difficult and very expensive operation. The second is to adjust the data after they are collected. Several methods were proposed to improve the coverage errors by now. In this thesis, multiple frame methodology was used to improve the population estimates.

3.2 Matching Errors

There are two kinds of matching errors, erroneous matches and erroneous nonmatches. An "erroneous match" means a report that is classified as a match but does not correspond to any report in the other system and, an `erroneous nonmatch` means a report that is classified as a nonmatch and does correspond to a report in the other system. Erroneous matches predominate within some sampling units because of similarities of names and duplications or other inadequacies of the address identification. Within other sampling units, erroneous nonmatches may predominate because of a poor interviewer or a careless registrar who makes errors in recording names, addresses, and other identification data. The definitions of "erroneous match" and "erroneous nonmatch" are important primarily during the process of trying to determine matching rules that will minimize bias and variance.

To minimize the matching bias we need only be concerned with the net matching error, that is

Matching error = (the number of erroneous matches) – (the number of nonerroneous matches)

On the basis of the definitions of matching error, following notations were established by Marks, Seltzer and Krotki (1974)

n_{1a} = number of erroneous matches among source 1 reports using some specific set of matching rules; that is, the number of source 1 reports classified as matched for which no true counterpart exists among any of the source 2 reports;

n_{1b} = number of erroneous nonmatches among source 1 reports under the same matching rules; that is, the number of source 1 reports classified as nonmatched for which a true counterpart does exist among the source 2 reports;

n_{2a} = number of erroneous matches among source 2 reports under same matching rules;

n_{2b} = number of erroneous nonmatches among source 2 reports;

$n_c = n_{1a} - n_{1b} = n_{2a} - n_{2b}$ = estimated net matching error ;

$n_{1d} = n_{1a} + n_{1b}$ = gross matching error for source 1; and

$n_{2d} = n_{2a} + n_{2b}$ = gross matching error for source 2.

The probabilities $P_{1a}, P_{2a}, P_{1b},$ and P_{2b} and the expected net and gross error rates $P_c, P_{1d},$ and P_{2d} are proposed by Marks, Seltzer and Krotki (1974) as follows:

$$P_{1a} = \frac{E(n_{1a})}{N}, \quad P_{2a} = \frac{E(n_{2a})}{N},$$

$$P_{1b} = \frac{E(n_{1b})}{N}, \quad P_{2b} = \frac{E(n_{2b})}{N},$$

$$P_c = \frac{E(n_c)}{N}, \qquad P_{1d} = \frac{E(n_{1d})}{N}, \text{ and } \quad P_{2d} = \frac{E(n_{2d})}{N}.$$

In addition,

m_t = Estimated number of events that should match (that is, the number of " true matches") and

$$P_{12} = \frac{E(m_t)}{N}.$$

n_1 = the number of events reported by source 1

n_2 = the number of events reported by source 2

m = the number of matches obtained in a particular study

So the number of matched reports is,

$$m = m_t + n_{1a} - n_{1b} = m_t + n_{2a} - n_{2b} = m_t + n_c, \qquad (3.2.1)$$

where $m_t, n_{1a}, n_{1b}, n_{2a}, n_{2b}$, and n_c are the sample estimates.

Marks, Seltzer and Krotki (1974) proposed the following estimator,

$$\hat{n} = \frac{n_1 n_2}{m} = \frac{n_1 n_2}{m_t + n_c}. \qquad (3.2.2)$$

The relative bias due to matching error B_m is

$$B_m = -\frac{E(n_c)}{E(m)} = -\frac{E(n_c)}{E(m_t)+E(n_c)} = -\frac{P_c}{P_{12}+P_c} \; . \qquad (3.2.3)$$

B_m can be estimated by

$$b_m = -\frac{n_c}{m} = \frac{m_t - m}{m}. \qquad (3.2.4)$$

As we have seen this formula, the relative matching bias is estimated by ratio of net matching error to the observed number of matches. In general, m_t and n_c are unknown. However it is possible to make a reasonably good approximation of m_t for a (small) sample. This procedure would be too expensive to use as a regular operating procedure, but the expenditure would be m_t and n_c Marks, Seltzer and Krotki (1974).

When multiple frame approach is used with more than two frames, some of them may be out of date, in this case record matching is very difficult and error prone. Errors in records matching are another sources of bias.

CHAPTER 4

SIMPLE RANDOM SAMPLING FROM EACH OF TWO FRAMES

The subject of multiple frames in sample surveys has been examined by Hartley (1962 and 1974), Cochran (1967), Fuller and Burmeister (1972), Lund (1968), Bankier (1986), Kalton and Anderson (1986), Skinner (1991), Skinner and Rao (1996), Lohr and Rao (2000 and 2006) and Mecatti (2005 and 2007). Different estimators were anticipated under several assumptions whether domain and frame sizes are known or unknown.

We follow the notation of Hartley (1962). Notations for two-frame surveys are shown in Table 10.

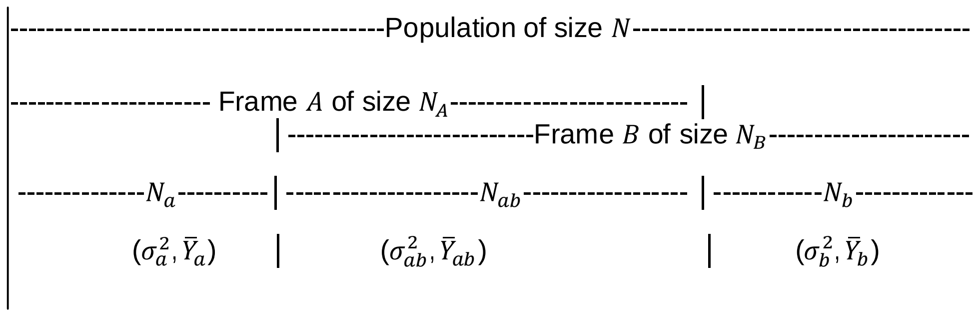
Table 10: Notation for Two Frame Designs and Estimates

	Frame		Domain		
	A	B	a	b	ab
Number of population elements	N_A	N_B	N_a	N_b	N_{ab}
Number of selected sample elements *	n_A	n_B	n_a	n_b	n'_{ab}, n''_{ab}
Population total for Y	Y_A	Y_B	Y_a	Y_b	Y_{ab}
Population mean for Y	$\bar{Y}_A$	$\bar{Y}_B$	$\bar{Y}_a$	$\bar{Y}_b$	$\bar{Y}_{ab}$
Sample total for Y*	y_A	y_B	y_a	y_b	y'_{ab}, y''_{ab}
Sample mean for Y*	$\bar{y}_A$	$\bar{y}_B$	$\bar{y}_a$	$\bar{y}_b$	$\bar{y}'_{ab}, \bar{y}''_{ab}$
Cost of enumeration of the sampling unit*	c_A	c_B			

*Applies to case of drawing random samples from both frames

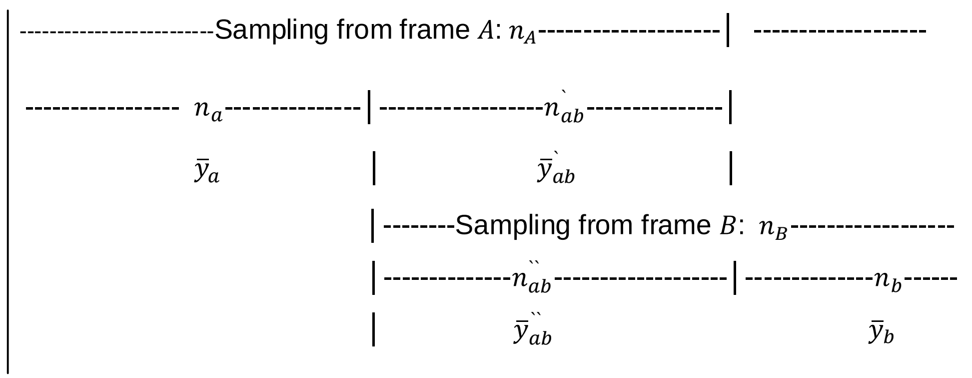
Source: Hartley (1962)

According to Lund (1968), the foundation for dual frames can be presented in Figure 5 and Figure 6 as:



Source: Lund (1968)

Figure 5: Division of the Population Elements for any Frame Between Two Domains



Source: Lund (1968)

Figure 6: Division of the Selected Sample Elements for any Frame Between Two Domains.

U is population, consisting of the union of two overlapping frames A and B. We can write

$U = A \cup B$, $a = A \cap B^c$,

$b = A^c \cap B$, $ab = A \cap B$,

where c denotes complement of a set. For example, B^c is the complement of B.

The sizes of U, A, B, a, b and ab are denoted N, N_A, N_B, N_a, N_b, N_{ab}, respectively, then

$$N = N_a + N_b + N_{ab} = N_a + N_B = N_A + N_b \tag{4.1}$$

$$N_A = N_a + N_{ab} \tag{4.2}$$

$$N_B = N_b + N_{ab} \tag{4.3}$$

Independent simple random samples s_A and s_B are selected from A and B respectively, then we can write

$s_a = s_A \cap a$, $s'_{ab} = s_A \cap ab$

$s_b = s_B \cap b$, $s''_{ab} = s_B \cap ab$

and denote the sizes of s_A, s_B, s_a, s_b, s'_{ab} and s''_{ab} by n_A, n_B, n_a, n_b, n'_{ab} and n''_{ab}, respectively.

An important point here is how we will classify sampled units into domains. This classification can be made by using algorithm in Figure 7. Note that, simple random samples s_A and s_B were drawn without replacement from frames A and B respectively.

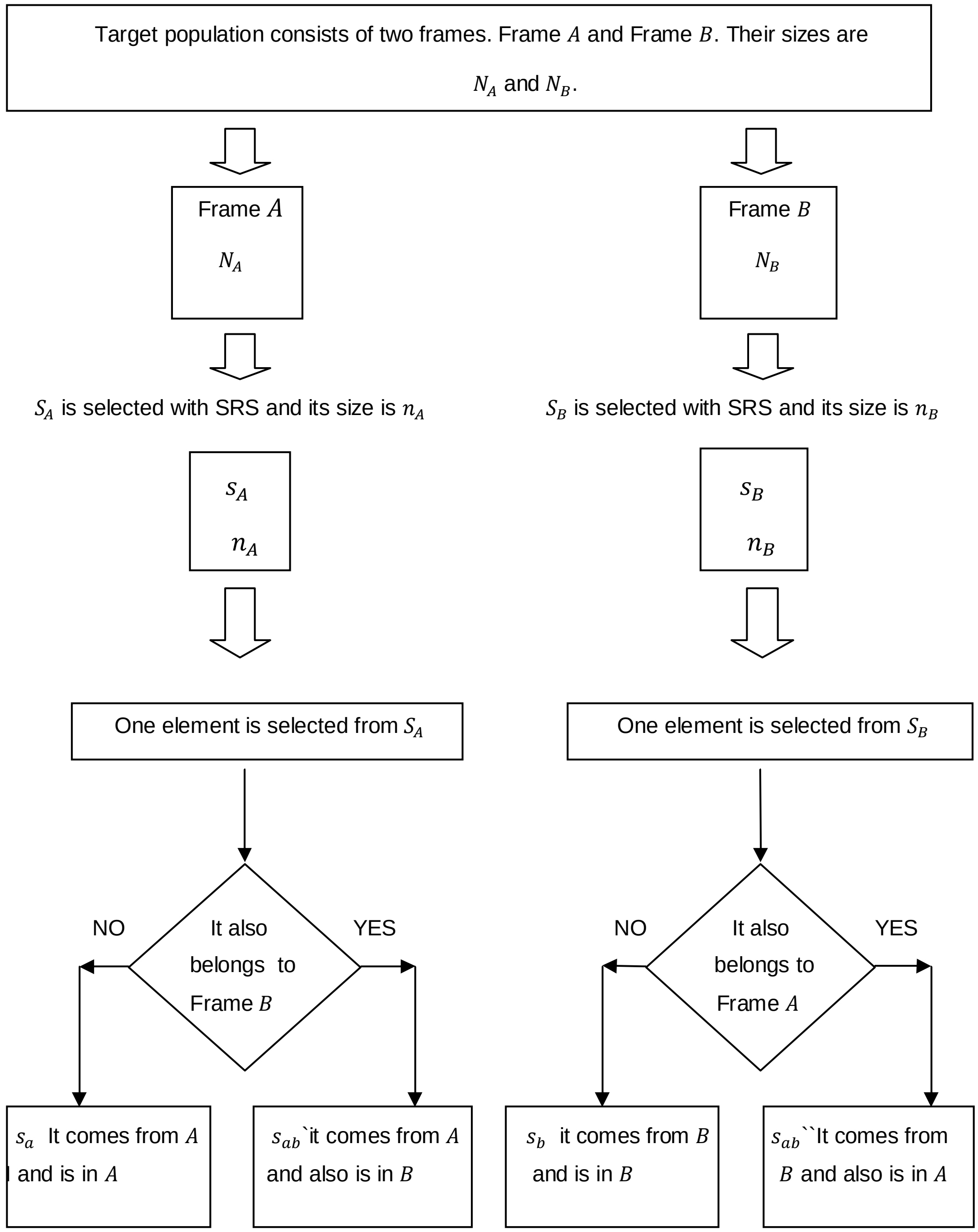


Figure 7: Algorithm for Classification of Sampled Elements to Domains for Dual Frame

A small application concerning classification can be shown as below. Independent simple random samples s_A, s_B shown in Table 12 selected from population frames A, B shown in Table 11 respectively.

Table 11: Two Population Frames A and B for Classification

FRAME A			FRAME B		
Pop. ID i	ID Label	Pop. Element y_i	Pop. ID i	ID Label	Pop. Element y_i
1	Aylin	1	1	Suheyla	1
2	Suheyla	1	2	Vural	1
3	Arif	2	3	Aysegul	0
4	Aynur	3	4	Berrin	1
5	Ayse	5	5	Mahmut	3
6	Hakan	2	6	Murat	0
7	Ebru	2	7	Hacer	6
8	Ozcan	3	8	Ayse	5
9	Birsen	4	9	Aylin	1
10	Kamil	2	10	Ebru	2
:	:	:	:	:	:

Table 12: Two Samples Selected from Frame A and Frame B for Classification

s_A				s_B			
Sample ID j	Pop. ID i	ID Label	Sample Element y_i	Sample ID j	Pop. ID i	ID Label	Sample Element y_i
1	4	Aynur	3	1	2	Vural	1
2	1	Aylin	1	2	6	Murat	0
3	10	Kamil	2	3	8	Ayse	5
4	7	Ebru	2	4	10	Ebru	2
5	8	Ozcan	3	5	4	Berrin	1
:	:	:	:	:	:	:	:

In this example, Aynur, Kamil and Ozcan are unduplicated elements because they are sampled from frame A and not in frame B. Aylin and Ebru are duplicated elements because they are sampled from frame A but also in frame B. In the same way Vural, Murat and Berrin are unduplicated elements because they are sampled from frame B and not in frame A. But Ayse and Ebru are duplicated elements. Since they are sampled from B and also in frame A. Hence, classification of the samples between two domains was shown in Table 13 and 14.

Table 13: Sample A Elements Classification to Domains a and ab

SAMPLE A s_A								
Sample element in domain a s_a				**Sample element in domain ab** s'_{ab}				
Sample ID j	**Pop. ID i**	**ID Label**	**Sample Element y_i**	**Sample ID j**	**Pop. ID i (A)**	**i (B)**	**ID Label**	**Sample Element y_i**
1	4	Aynur	3	2	1	9	Aylin	1
3	10	Kamil	2	4	7	10	Ebru	2
5	8	Ozcan	3					
:	:	:	:	:	:	:	:	:

Table 14: Sample B Elements Classification to Domains a and ab

SAMPLE B s_B								
sample element in domain b s_b				**Sample element in domain ab** s''_{ab}				
Sample ID j	**Pop. ID i**	**ID Label**	**Sample Element y_i**	**Sample ID j**	**Pop. ID i (A)**	**i (B)**	**ID Label**	**Sample Element y_i**
1	2	Vural	1	3	5	8	Ayse	5
2	6	Murat	0	4	7	10	Ebru	2
5	4	Berrin	1					
:	:	:	:	:	:	:	:	:

One of the aim analyzing data from a multiple frame survey is often to estimate population total Y, using information from independent samples taken from the frames. In dual frame survey, we can write

$$Y = Y_a + Y_{ab} + Y_b, \tag{4.4}$$

where Y_a is the total of the population elements in domain a (unduplicated), Y_{ab} is the total of the population elements in domain ab, Y_b is the total of the population elements in domain b. Let s_A and s_B samples drawn independently from A and B. Suppose that y_i observed for each element in s_A and s_B.The estimation problem using this data to construct a suitable estimator $\hat{Y}$ of Y and estimator of its variance. This problem depends on what is known about $N_A, N_B,$ and N_{ab}. There are three different situations Skinner and Rao (1996).

- **Situation 1:** N_A, N_B, N_{ab} known
- **Situation 2**: N_A, N_B known but N_{ab} unknown
- **Situation 3:** N_A, N_B, N_{ab} unknown, it may arise if, for example, both A and B consist of clusters, and lists are only available of clusters with measures of size.

In this thesis, situation 1 and situation 2 were examined.

For the purpose of estimation of population total we assume that we know:

(a) n_A, n_B, n_a, n'_{ab}, n''_{ab}, n_b;

(b) y_A, y_B, y_a, y'_{ab}, y''_{ab} and y_b the totals of y_i across units *i* in s_A, s_B, s_a, s'_{ab}, s''_{ab} and s_b, respectively;

(c) N_A and N_B.

Finally we let Y_A, Y_B, Y_a, Y_{ab}, and Y_b denote the totals of y_i across units i in , B, a, ab, and b, respectively, and write

$$\bar{Y}_A = Y_A / N_A, \quad \bar{Y}_B = Y_B / N_B, \quad \bar{Y}_a = Y_a / N_a, \quad \bar{Y}_b = Y_b / N_b$$

$$\bar{y}_A = {}^{y_A}/_{n_A}, \quad \bar{y}_B = {}^{y_B}/_{n_B}, \quad \bar{y}_a = y_a / n_a, \quad \bar{y}_b = y_b / n_b,$$

$$\bar{y}_{ab} = y_{ab} / n_{ab}, \quad \bar{y}'_{ab} = y'_{ab} / n'_{ab}, \quad \bar{y}''_{ab} = y''_{ab} / n''_{ab}$$

where $y_{ab} = y'_{ab} + y''_{ab}$ and $n_{ab} = n'_{ab} + n''_{ab}$.

4.1 Estimators of Population Total When N_A, N_B, N_{ab} are Known

In literature Hartley (1962) and Lund (1968) proposed dual frame estimator of population total for situation 1. This situation may arise if frames A and B are available as known length and if the overlap size N_{ab} can either be determined from these lists or is known; e.g., if frame A is complete, then $N_{ab} = N_B$.

4.1.1 Hartley's Estimator

Hartley (1962) suggested the unbiased estimator for the population total

$$\hat{Y}_H = N_a\bar{y}_a + N_{ab}p\bar{y}_{ab}^{'} + N_{ab}q\bar{y}_{ab}^{''} + N_b\bar{y}_b \quad (4.1.1.1)$$

where the means $\bar{y}_a$, $\bar{y}_{ab}^{'}$ are replaced by $\bar{y}_A$ if either n_a = 0 or $n_{ab}^{`}$ = 0, and similarly the means $\bar{y}_b$, $\bar{y}_{ab}^{``}$ are replaced by $\bar{y}_B$ if either $n_b$ = 0 or $n_{ab}^{``}$ = 0.

The variance of this estimator is approximately equal to that in proportional allocation stratified sampling so that

$$Var\,(\hat{Y}_H) = \frac{N_A^2}{n_A}\left[(1-\alpha)\sigma_a^2 + \alpha p^2\sigma_{ab}^2\right] + \frac{N_B^2}{n_B}\left[(1-\beta)\sigma_b^2 + \beta q^2\sigma_{ab}^2\right] \quad (4.1.1.2)$$

where $= {}^{N_{ab}}/_{N_A}$, $\beta = {}^{N_{ab}}/_{N_B}$,

finite population correction have been ignored and σ_a^2, σ_b^2 and σ_{ab}^2 are the "within domain" population variances.

Assuming a linear function of cost can be expressed by

$$\text{Total Cost } C = n_A c_A + n_B c_B \quad (4.1.1.3)$$

where c_A and c_B define costs of enumeration of the sampling elements from each frame respectively .The problem of optimizing the sample allocation among the two frames and finding the optimum value for p consists of minimizing $Var\,(\hat{Y}_H)$ as a function of p, n_A and n_B subject to restriction cost function.

The optimum value for p is given by a solution of the bi-quadratic

$$c_A p^2/c_B q^2 = \frac{(1-\alpha)\sigma_a^2 + \alpha p^2\sigma_{ab}^2}{(1-\beta)\sigma_b^2 + \beta q^2\sigma_{ab}^2} \quad (4.1.1.4)$$

The optimum sampling fractions are given by

$$\frac{n_A}{N_A} = c\,\left\{\left[(1-\alpha)\sigma_a^2 + \alpha p^2 \sigma_{ab}^2\right] / \, c_A\right\}^{1/2} \tag{4.1.1.5}$$

$$\frac{n_B}{N_B} = c\,\left\{\left[(1-\beta)\sigma_b^2 + \beta q^2 \sigma_{ab}^2\right] / \, c_B\right\}^{1/2} \tag{4.1.1.6}$$

with c shows budget (4.1.1.3).

4.1.2 Lund's Estimator

Lund (1968) showed Hartley`s bi-quadratic p value, with a simple expression as below:

$$p_{oH} = \frac{\alpha n_A}{\alpha n_A + \beta n_B} \tag{4.1.2.1}$$

Lund considered that Hartley`s procedure does not consider the actual division achieved (at random) of $n_A$ and $n_B$ elements among the domains. He discussed whether a gain is achieved when the $p$ value is a function of $n_{ab}^{`}$ and $n_{ab}^{``}$.

The variance of $\hat{Y}_H$ as given by (4.1.1.1) can be taken in two steps

$$Var(\hat{Y}) = E\left[Var(\hat{Y} \mid n_{ab}^{`}, n_{ab}^{``})\right] + Var\left[E(\hat{Y} \mid n_{ab}^{`}, n_{ab}^{``})\right] \tag{4.1.2.2}$$

where $E(\hat{Y} \mid n_{ab}^{`}, n_{ab}^{``}) = Y$ and $Var\left[E(\hat{Y} \mid n_{ab}^{`}, n_{ab}^{``})\right] = 0$

Ignoring finite population corrections

$$Var(\hat{Y} \mid n_{ab}^{`}, n_{ab}^{``}) = \frac{N_a^2}{n_a}\sigma_a^2 + p^2 \frac{N_{ab}^2}{n_{ab}^{`}}\sigma_{ab}^2 + (1-p)^2 \frac{N_{ab}^2}{n_{ab}^{``}}\sigma_{ab}^2 + \frac{N_b^2}{n_b}\sigma_b^2 \tag{4.1.2.3}$$

The values of p which minimizes the variances in (4.1.2.3) is

$$\frac{\partial Var(\hat{Y} \mid n^{`}_{ab}, n^{``}_{ab})}{\partial p} = \frac{Var(\bar{y}^{``}_{ab})}{Var(\bar{y}^{`}_{ab}) + Var(\bar{y}^{``}_{ab})}$$

$$p_o = \frac{n^{`}_{ab}}{n^{`}_{ab} + n^{``}_{ab}} \qquad q = 1 - p \tag{4.1.2.4}$$

Substituting (4.1.2.4) into (4.1.2.3) and obtaining the expected value, the variance for the estimator with this value for p can be found.

The estimator of population total and it's approximately variance are

$$\hat{Y}_L = N_a \bar{y}_a + N_{ab} \bar{y}_{ab} + N_b \bar{y}_b \tag{4.1.2.5}$$

where $$\bar{y}_{ab} = \frac{n^{'}_{ab}\bar{y}^{'}_{ab} + n^{''}_{ab}\bar{y}^{''}_{ab}}{n^{'}_{ab} + n^{''}_{ab}} \tag{4.1.2.6}$$

$$Var\,(\hat{Y}_L) = \frac{N_A^2}{n_A}(1-\alpha)\,\sigma_a^2 + \frac{N_A N_B \alpha\beta}{\alpha n_A + \beta n_B}\,\sigma_{ab}^2 + \frac{N_B^2}{n_B}(1-\beta)\,\sigma_b^2. \tag{4.1.2.7}$$

4.2 Estimators of Population Total When N_A, N_B are Known But N_{ab} is Unknown

In literature different estimators were proposed by Hartley (1962, 1974), Lund (1968), Fuller and Burmeister (1972), Bankier (1986), Skinner (1991) Skinner and Rao (1996) and Mecatti (2005). This situation may arise if again frame A and B are lists of known length but reliable determination of overlap size are not possible because of practical difficulties in matching.

We assume two frames A and B, containing N_A and N_B elements respectively are available. Thus

$$N_A = N_a + N_{ab} \quad \text{and} \quad N_B = N_b + N_{ab}$$

and the total number of elements in the population, N, was given by

$$N = N_a + N_b + N_{ab} = N_a + N_B = N_A + N_b$$

From this information the estimates of N_a, N_b and N_{ab} are to be computed. When N_A and N_B are known it is only essential to estimate the number of elements in N_{ab}. After estimate N_{ab} the estimates of N_a and N_b can be easily found by subtraction as

$$\widehat{N}_a = N_A - \widehat{N}_{ab} \qquad \text{and} \qquad \widehat{N}_b = N_B - \widehat{N}_{ab}.$$

The number of distinct elements in the population can be estimated as

$$\widehat{N} = N_A + N_B - \widehat{N}_{ab}.$$

4.2.1 Hartley's Estimator

Hartley (1962) considered if domain sizes N_a, N_b, N_{ab} are not known the ordinary stratified sampling formulas applied to the u-variables in the two strata (frames) of sizes N_A and N_B must be applied. Thus the estimator of $U = Y$ is given by

$$\hat{Y}_H = \frac{N_A}{n_A}\left\{y_a + py'_{ab}\right\} + \frac{N_B}{n_B}\left\{y_b + qy''_{ab}\right\} \tag{4.2.1.1}$$

The variance of this estimator is

$$Var(\hat{Y}_H) = \frac{N_A^2}{n_A}\left\{(1-\alpha)\sigma_a^2 + \alpha p^2\sigma_{ab}^2 + \alpha(1-\alpha)(\bar{Y}_a - p\bar{Y}_{ab})^2\right\}$$

$$+ \frac{N_B^2}{n_B}\left\{(1-\beta)\sigma_b^2 + \beta q^2\sigma_{ab}^2 + \beta(1-\beta)(\bar{Y}_b - q\bar{Y}_{ab})^2\right\} \tag{4.2.1.2}$$

The optimum allocation formulas

$$\frac{{n_A}^2}{{N_A}^2} = c\left\{(1-\alpha)\sigma_a^2 + \alpha p^2\sigma_{ab}^2 + \alpha(1-\alpha)(\bar{Y}_a - p\bar{Y}_{ab})^2\right\}/c_A \tag{4.2.1.3}$$

$$\frac{{n_B}^2}{{N_B}^2} = c\left\{(1-\beta)\sigma_b^2 + \beta q^2\sigma_{ab}^2 + \beta(1-\beta)(\bar{Y}_a - \bar{Y}_{ab})^2\right\}/c_B. \quad (4.2.1.4)$$

4.2.2 Lund's estimator

Lund (1968) proposed estimators for the case of unknown N_a, N_b and N_{ab}. Lund used estimates of unknown domain sizes by use of sample data and known N_A and N_B. The expressions N_A (n_a / n_A) and N_B (n_b / n_B) are unbiased estimators of N_a and N_b, respectively. N_A ($n_{ab}^{`}$ / n_A) and N_B ($n_{ab}^{``}$ / n_B) are two unbiased estimators of N_{ab}. Using p and $1-p$ as undetermined weights for two estimates of N_{ab}, and replacing these estimator in (4.1.2.5), an unbiased estimator of population total is

$$\hat{Y}_L = \frac{N_A}{n_A} n_a \bar{y}_a + \left\{\frac{N_A}{n_A} n_{ab}^{`} p + \frac{N_B}{n_B} n_{ab}^{``} (1-p)\right\} \bar{y}_{ab} + \frac{N_B}{n_B} n_b \bar{y}_b \quad (4.2.2.1)$$

where as before $\bar{y}_{ab}$ is sample mean of all elements selected from domain of duplicated elements.

$$\bar{y}_{ab} = \frac{n_{ab}^{`}\bar{y}_{ab}^{`} + n_{ab}^{``}\bar{y}_{ab}^{``}}{n_{ab}^{`} + n_{ab}^{``}}$$

The variance of Lund`s estimator is

$$Var(\hat{Y}_L) = \frac{N_A^2}{n_A}(1-\alpha)\sigma_a^2 + \frac{N_A N_B \alpha\beta}{\alpha n_A + \beta n_B}\sigma_{ab}^2 + \frac{N_B^2}{n_B}(1-\beta)\sigma_b^2 +$$

$$+ \frac{N_A^2(1-\alpha)\alpha}{n_A}(\bar{Y}_a - p\bar{Y}_{ab})^2 + \frac{N_B^2(1-\beta)\beta}{n_B}(\bar{Y}_b - q\bar{Y}_{ab})^2. \quad (4.2.2.2)$$

The final two terms represent the increase in variance due to not knowing N_a, N_b and N_{ab}. Lund showed the optimum value of p and estimator of optimum p from the sample data as below,

$$p_L = \frac{\frac{N_A(1-\alpha)}{n_A}\bar{Y}_a + \frac{N_B(1-\beta)}{n_B}(\bar{Y}_{ab} - \bar{Y}_b)}{\left(\frac{N_A n_a}{n_A} + \frac{N_B n_b}{n_B}\right)\bar{Y}_{ab}} \qquad (4.2.1.3)$$

$$\hat{p}_L = \frac{\frac{N_A n_a}{n_A^2}\bar{y}_a + \frac{N_B n_b}{n_B^2}(\bar{y}_{ab} - \bar{y}_b)}{\left(\frac{N_A n_a}{n_A^2} + \frac{N_B n_b}{n_B^2}\right)\bar{y}_{ab}} \qquad (4.2.1.4)$$

Lund (1968) indicated Hartley`s weight variable $p$ was calculated for two individual estimates of the total over the duplicated elements ($N_{ab}, \bar{Y}_{ab}$ ). The optimum value of $p$ includes $\sigma_a^2, \sigma_{ab}^2$ and $\sigma_b^2$. Thus Lund`s estimator is more simple than the Hartley`s estimator. Furthermore Skinner (1991) showed that $Var\,(\hat{Y}_L) \leq \mathrm{Var}(\hat{Y}_H)$ for any value of p .

4.2.3 Fuller and Burmeister's Estimator

Fuller and Burmeister (1972) proposed the estimator

$$\hat{Y}_{FB} = (N_A - \hat{N}_{ab,FB})\,\bar{y}_a + \hat{N}_{ab,FB}\bar{y}_{ab} + (N_B - \hat{N}_{ab,FB})\,\bar{y}_b, \qquad (4.2.3.1)$$

where $\bar{y}_{ab} = p_{FB}\bar{y}_{ab}^{`} + (1 - p_{FB})\bar{y}_{ab}^{``},$ (4.2.3.2)

$\bar{y}_{ab}^{`} = \frac{1}{n_{ab}^{`}} \sum_{i=1}^{n_{ab}^{`}} y_i$ = the mean of elements in domain ab selected from frame A

$\bar{y}_{ab}^{``} = \frac{1}{n_{ab}^{``}} \sum_{i=1}^{n_{ab}^{``}} y_i$ = the mean of elements in domain ab selected from frame B

$$p_{FB} = \frac{n_{ab}^{`}(1-f_B)}{n_{ab}^{`}(1-f_B) + n_{ab}^{``}(1-f_A)}. \qquad (4.2.3.3)$$

Here, p_{FB} has been chosen to minimize the variance of $\bar{y}_{ab}$. If the finite correction term can be ignored the estimator of $\bar{Y}_{ab}$ reduces to the mean of all $n_{ab}^{`} + n_{ab}^{``}$ elements selected from domain ab. The estimator $\hat{Y}_{FB}$ is linear in observations y_i. Since weights are not a function of the characteristic they apply equally well for all y-characteristics (Fuller and Burmeister 1972).

Further, $\hat{N}_{ab,FB}$ is the smallest root of the quadratic equation

$$[n_A g_B + n_B g_A]\ \hat{N}^2_{ab,FB} - [n_A N_B g_B + n_B N_A g_A + n^{`}_{ab} N_A g_B + n^{``}_{ab} N_B g_A]\hat{N}_{ab,FB} + [n^{`}_{ab} g_B + n^{``}_{ab} g_A]\ N_A N_B = 0$$

Considering Hartley`s estimator for p value given by Fuller and Burmeister (1972), we have derived the estimator of $N_{ab,FB}$ as stated below.

$$\hat{N}_{ab,H} = \frac{p n^{`}_{ab} N_A}{n_A} + \frac{q n^{``}_{ab} N_B}{n_B} \tag{4.2.3.4}$$

$$\hat{N}_{ab,H} = p\hat{N}^{`}_{ab} + q\hat{N}^{``}_{ab},$$

The variance of Hartley`s estimator is

$$Var(\hat{N}_{ab,H}) = p^2\, Var(\hat{N}^{`}_{ab}) + q^2\, Var(\hat{N}^{``}_{ab})$$

$$Var(\hat{N}^{`}_{ab}) = (1 - \frac{n_A}{N_A})\ \frac{N_A}{N_A-1}\ \frac{N_A^2}{n_A}\ (\frac{N_{ab}}{N_A})\ (1 - \frac{N_{ab}}{N_A})$$

$$= \left(\frac{N_A-n_A}{N_A-1}\right) \frac{N_A^2}{n_A} (\frac{N_{ab}}{N_A})(\frac{N_A-(N_A-N_a)}{N_A})$$

$$= \frac{g_{AN_{ab}N_a}}{n_A} \quad \text{where} \quad g_A = \frac{N_A-n_A}{N_A-1} \tag{4.2.3.5}$$

$$Var(\hat{N}^{``}_{ab}) = (1 - \frac{n_B}{N_B})\ \frac{N_B}{N_B-1}\ \frac{N_B^2}{n_B}\ (\frac{N_{ab}}{N_B})\ (1 - \frac{N_{ab}}{N_B})$$

$$= \left(\frac{N_B-n_B}{N_B-1}\right) \frac{N_B^2}{n_B} (\frac{N_{ab}}{N_B})(\frac{N_B-(N_B-N_a)}{N_B})$$

$$= \frac{g_{BN_{ab}N_b}}{n_B} \quad \text{where} \quad g_B = \frac{N_B-n_B}{N_B-1} \tag{4.2.3.6}$$

The values of p which minimizes the variance is

$$\frac{\partial V(\hat{N}_{ab})}{\partial p} = 2p\, Var(\hat{N}^{`}_{ab}) - 2\, Var(\hat{N}^{``}_{ab}) + 2p \mathrm{Var}(\hat{N}^{``}_{ab}) = 0$$

$$p = \frac{Var(\hat{N}^{``}_{ab})}{Var(\hat{N}^{`}_{ab}) + Var(\hat{N}^{``}_{ab})} \qquad q = 1 - p \tag{4.2.3.7}$$

Thus the value of p using (4.2.3.7) is

$$p_H = \frac{\frac{g_B N_{ab} N_b}{n_B}}{\frac{g_B N_{ab} N_b}{n_B} + \frac{g_A N_{ab} N_a}{n_A}} = \frac{n_A N_b g_B}{n_A N_b g_B + n_B N_a g_A} \tag{4.2.3.8}$$

To obtain an estimator of N_{ab} that is a function of the sample data only we put

$$\hat{N}_a = N_A - \hat{N}_{ab}$$

$$\hat{N}_b = N_B - \hat{N}_{ab}$$

and substitute p_H into formula of $\hat{N}_{ab,H}$.

$$\hat{N}_{ab,FB} = \frac{n_A N_b g_B n^{`}_{ab} N_A}{[n_A N_b g_B + n_B N_a g_A] n_A} + \frac{n_B N_a g_A n^{``}_{ab} N_B}{[n_A N_b g_B + n_B N_a g_A] n_B}$$

$$\hat{N}_{ab,FB} = \frac{n_A (N_B - \hat{N}_{ab,FB}) g_B n^{`}_{ab} N_A}{[n_A (N_B - \hat{N}_{ab,FB}) g_B + n_B (N_A - \hat{N}_{ab,FB}) g_A] n_A} + \frac{n_B (N_A - \hat{N}_{ab,FB}) g_A n^{``}_{ab} N_B}{[n_A (N_B - \hat{N}_{ab,FB}) g_B + n_B (N_A - \hat{N}_{ab,FB}) g_A] n_B}$$

$$\hat{N}_{ab,FB} = \frac{n_A n_B N_A N_B [g_B n^{`}_{ab} + g_A n^{``}_{ab}] - n_A n_B \hat{N}_{ab} [g_B n^{`}_{ab} N_A + g_A n^{``}_{ab} N_B]}{[n_A N_B g_B - n_A \hat{N}_{ab} g_B + n_B N_A g_A - n_B \hat{N}_{ab} g_A] n_A n_B}$$

$$N_A N_B [g_B n^{`}_{ab} + g_A n^{``}_{ab}] - \hat{N}_{ab,FB} [g_B n^{`}_{ab} N_A + g_A n^{``}_{ab} N_B] - \hat{N}_{ab,FB} [n_A N_B g_B + n_B N_A g_A] - \hat{N}^2_{ab,FB} [n_A g_B + n_B g_A] = 0$$

Thus the estimation of N_{ab} can be found by this quadratic,

$$[n_A g_B + n_B g_A]\,\widehat{N}_{ab,FB}^2 - [n_A N_B g_B + n_B N_A g_A + n_{ab}^{`} N_A g_B + n_{ab}^{``} N_B g_A]\widehat{N}_{ab,FB}$$
$$+ [n_{ab}^{`} g_B + n_{ab}^{``} g_A]\, N_A N_B = 0$$

It can be shown that the roots of this quadratic are always real and the largest root is always greater than or equal to the minimum of N_A and N_B. The smallest root is always contained in the interval zero to minimum (N_A , N_B), inclusive. Hence we take the left (smallest) root of this quadratic as the estimator of $N_{ab.}$ Note that Hartley`s p value does not always fall into the range of feasible values for N_{ab} Fuller and Burmeister (1972).

The variance of $\widehat{Y}_{FB}$ is

$$Var\,(\widehat{Y}_{FB}) = N_a(f_A^{-1} - 1)S_a^2 + [(1 - f_B)f_A + (1 - f_A)f_B]^{-1}$$
$$(1 - f_A)\,(1 - f_B)N_{ab}S_{ab}^2 + N_b(f_B^{-1} - 1)S_b^2$$
$$+ (\bar{Y}_{ab} - \bar{Y}_a - \bar{Y}_b)^2\,\frac{N_{ab}N_a N_b g_A g_B}{n_A N_b g_B + n_B N_a g_A} \qquad (4.2.3.9)$$

where

$$S_a^2 = \frac{1}{N_a - 1}\sum_{i=1}^{N_a}(y_{ai} - \bar{Y}_a)^2, \qquad (4.2.3.10)$$

$$S_{ab}^2 = \frac{1}{N_{ab} - 1}\sum_{i=1}^{N_{ab}}(y_{abi} - \bar{Y}_{ab})^2 \qquad (4.2.3.11)$$

and

$$S_b^2 = \frac{1}{N_b - 1}\sum_{i=1}^{N_b}(y_{bi} - \bar{Y}_b)^2. \qquad (4.2.3.12)$$

4.2.4 Single Frame Estimator

Bankier (1986) proposed an estimator based on Horvitz-Thompson estimator using the following simple form. It is assumed that units selected in more than one sample can be identified. But, the estimator of Kalton and Anderson (1986) and Skinner (1991) do not require identification of units found in both samples.

For SRS of each frame, the Single Frame estimator for dual frame is given by:

$$\hat{Y}_{SF} = f_A^{-1} y_a + (f_A + f_B)^{-1} y_{ab} + f_B^{-1} y_b \qquad (4.2.4.1)$$

Skinner (1991) indicated that $\hat{Y}_{SF}$ is a special case of $\hat{Y}_H$ or of $\hat{Y}_L$ with $p = f_A (f_A + f_B)^{-1}$ and $Var(\hat{Y}_{SF}) \geq Var(\hat{Y}_{FB})$.

$$\hat{Y}_{SF} = f_A^{-1}\left[y_a + p y_{ab}^{`}\right] + f_B^{-1}\left[y_b + (1-p) y_{ab}^{``}\right] \qquad (4.2.4.2)$$

In a single frame estimator data from two frames are combined by using fixed weights which guarantee design-unbiasedness. Fixed weights usually differ from optimum weights, the SF estimator is generally less efficient than optimum estimator (Lohr and Rao 2000). Finally, computationally and algebraically it is very easy to extend the method to any number of frames. However the correct classification of sampled units into domains still required.

4.2.5 Raking Ratio Estimator

Bankier (1986) the rth iteration raking ratio estimator is given by

$$\hat{Y}_{RR}^{(r)} = W_a^{(r)} y_a + W_b^{(r)} y_b + W_{ab}^{(r)} y_{ab}, \qquad (4.2.5.1)$$

where

$$W_a^{(0)} = f_A^{-1}, \quad W_b^{(0)} = f_B^{-1}, \quad W_{ab}^{(0)} = (f_A + f_B)^{-1},$$

$$W_a^{(r)} = (N_A/\widehat{N}_A^{(r-1)})W_a^{(r-1)}, \quad W_b^{(r)} = W_b^{(r-1)},$$

$$W_{ab}^{(r)} = (N_A/\widehat{N}_A^{(r-1)})W_{ab}^{(r-1)}, \quad r = 1,3,5,\ldots\ldots\ldots$$

$$W_a^{(r)} = W_a^{(r-1)}, \quad W_b^{(r)} = (N_B/\widehat{N}_B^{(r-1)})W_b^{(r-1)},$$

$$W_{ab}^{(r)} = \left(N_B/\widehat{N}_B^{(r-1)}\right)W_{ab}^{(r-1)}, \quad r = 2,4,6,\ldots\ldots\ldots$$

$$\widehat{N}_A^{(r)} = W_a^{(r)}n_a + W_{ab}^{(r)}n_{ab}, \quad \widehat{N}_B^{(r)} = W_b^{(r)}n_b + W_{ab}^{(r)}n_{ab}. \tag{4.2.5.2}$$

$\hat{Y}_{RR}^{(r)}$ was obtained by taking $\hat{Y}_{RR}^{(0)} = \hat{Y}_S$ and applying r successive ratio adjustment with respect to domains A and B alternately.

Theorem 1 of Skinner (1991) showed that the raking procedure in fact converges to the following raking ratio (RR) estimator for multiple frame surveys.

Theorem 1: As $r \to \infty$, $\hat{Y}_{RR}^{(r)} \to \hat{Y}_{RR}^{(\infty)}$,

where

$$\hat{Y}_{RR}^{(\infty)} = \left(N_A - \widehat{N}_{ab,RR}\right)\bar{y}_a + \widehat{N}_{ab,RR}\bar{y}_{ab} + (N_B - \widehat{N}_{ab,RR})\bar{y}_b, \tag{4.2.5.3}$$

and $\widetilde{N}_{ab,RR}$ is the smallest root of the quadratic equation

$$n_{ab}\widehat{N}_{ab,RR}{}^2 - [n_{ab}(N_A + N_B) + (f_A^{-1} + f_B^{-1})n_a n_b]\widehat{N}_{ab,RR} + n_{ab}N_A N_B = 0 \tag{4.2.5.4}$$

Proof of Theorem 1 was given by Skinner (1991) as follows.

Proof: $$\hat{Y}_{RR}^{(\infty)} = W_a^{(\infty)}y_a + W_b^{(\infty)}y_b + W_{ab}^{(\infty)}y_{ab} \tag{4.2.5.5}$$

where $W_a^{(\infty)}$, $W_b^{(\infty)}$, and $W_{ab}^{(\infty)}$ are the limit of sequence $\left\{W_a^{(r)}\right\}$, $\left\{W_b^{(r)}\right\}$, and $\left\{W_{ab}^{(r)}\right\}$, respectively.

Then

$$W_a^{(\infty)} = f_A^{-1}\prod_{r\ odd}(N_A/\widehat{N}_A^{(r-1)}) = f_A^{-1}k_A,$$

$$W_b^{(\infty)} = f_B^{-1} \prod_{r\,even} (N_B / \widehat{N}_B^{(r-1)}) = f_B^{-1} k_B,$$

$$W_{ab}^{(\infty)} = (f_A + f_B)^{-1} \prod_{r\,odd} (N_A / \widehat{N}_A^{(r-1)}) \prod_{r\,odd} (N_A / \widehat{N}_A^{(r-1)})$$

$$= (f_A + f_B)^{-1} k_A k_B. \tag{4.2.5.6}$$

We can write the frame sizes N_A and N_B as below,

$$N_A = W_a^{(\infty)} n_a + W_{ab}^{(\infty)} n_{ab}$$

$$= f_A^{-1} k_A n_a + (f_A + f_B)^{-1} k_A k_B n_{ab} \tag{4.2.5.7}$$

$$N_B = W_b^{(\infty)} n_b + W_{ab}^{(\infty)} n_{ab}$$

$$= f_B^{-1} k_A n_b + (f_A + f_B)^{-1} k_A k_B n_{ab} \tag{4.2.5.8}$$

Let $\widehat{N}_{ab,RR} = W_{ab}^{(\infty)} n_{ab} = (\, f_A + f_B)^{-1} k_A k_B n_{ab}$

From this equation we can write

$$k_A k_B = \frac{\widehat{N}_{ab,RR}(f_A + f_B)}{n_{ab}} \tag{4.2.5.9}$$

Then, from (4.2.5.7) and (4.2.5.8)

$$(N_A - \widehat{N}_{ab,RR})(\, N_B - \widehat{N}_{ab,RR}) = f_A^{-1} f_B^{-1} n_a n_b k_A k_B \tag{4.2.5.10}$$

the equation (4.2.5.9) was inserted in (4.2.5.10). Thus

$$N_A N_B - N_A \widehat{N}_{ab} - N_B \widehat{N}_{ab} + \widehat{N}_{ab}{}^2 = f_A^{-1} f_B^{-1} n_a n_b \frac{\widehat{N}_{ab}(f_A + f_B)}{n_{ab}}$$

$$n_{ab} \widehat{N}_{ab}{}^2 - [n_{ab}(N_A + N_B) + n_a n_b (f_A^{-1} + f_B^{-1})] \widehat{N}_{ab} + n_{ab} N_A N_B = 0$$

and $N_A - \widehat{N}_{ab} = W_a^{(\infty)} n_a, \quad N_B - \widehat{N}_{ab} = W_b^{(\infty)} n_b$. Hence from (4.2.5.7)

$$\widehat{Y}_{RR}^{(\infty)} = \left(N_A - \widehat{N}_{ab}\right) \bar{y}_a + \widehat{N}_{ab} \bar{y}_{ab} + (N_B - \widehat{N}_{ab}) \bar{y}_b,$$

The variance of $\hat{Y}_{RR}^{(\infty)}$ is

As n→∞

$$Var(\hat{Y}_{RR}^{(\infty)}) \sim f_A^{-1} N_a S_a^2 + f_B^{-1} N_b S_b^2 + (f_A + f_B)^{-1} N_{ab} S_{ab}^2$$

$$+ (\bar{Y}_a + \bar{Y}_b - \bar{Y}_{ab})^2 \; \frac{N_{ab} N_a N_b}{n_A N_b + n_B N_a} (1 + \lambda^2), \qquad (4.2.5.11)$$

where

$$\lambda^2 = \frac{N_a N_b N_{ab}^2 (f_A^2 N_A - f_B^2 N_B)^2}{N_A N_B f_A f_B (f_A + f_B)^2 (N_{ab}^2 - N_A N_B)^2} \; . \qquad (4.2.5.12)$$

If λ = 0 asymptotically $Var\ (\hat{Y}_{FB}) \leq Var\ (\hat{Y}_{RR}^{(\infty)})$.

4.2.6 Pseudo-Maximum Likelihood Estimator

Fuller and Burmeister`s estimator for simple random samples was

$$\hat{Y}_{FB} = (N_A - \hat{N}_{ab,FB})\ \bar{y}_a + \hat{N}_{ab,,FB} \bar{y}_{ab} + (N_B - \hat{N}_{ab,FB})\ \bar{y}_b$$

and ignoring finite-population correction, $\hat{N}_{ab}$ was the smallest root of the quadratic equation

$$[n_A + n_B]\ \hat{N}_{ab,FB}^2 - [n_A N_B + n_B N_A + n_{ab}^{`} N_A + n_{ab}^{``} N_B] \hat{N}_{ab,FB}$$

$$+ [n_{ab}^{`} + n_{ab}^{``}]\ N_A N_B = 0$$

Fuller and Burmeister (1972) showed that $\hat{N}_{ab}$ is a MLE of N_{ab}. Skinner and Rao (1996) indicated that $\hat{Y}_{FB}$ can not be directly applied in the case of complex designs, because it will generally be design inconsistent for Y. They proposed modifying the maximum likelihood estimator for a simple random sample to obtain pseudo-maximum likelihood (PML) estimator for a complex design. It is given by

$$\hat{Y}_{PML} = \left(N_A - \hat{N}_{ab,PML}\right)\bar{y}_a + (N_B - \hat{N}_{ab,PML})\bar{y}_b + \hat{N}_{ab,PML}\bar{y}_{ab}, \qquad (4.2.6.1)$$

where

$$\bar{y}_{ab} = \left[\frac{n_A}{N_A}\hat{N}'_{ab}\bar{y}'_{ab} + \frac{n_B}{N_B}\hat{N}''_{ab}\bar{y}''_{ab}\right] / \left[\frac{n_A}{N_A}\hat{N}'_{ab} + \frac{n_B}{N_B}\hat{N}''_{ab}\right] \qquad (4.2.6.2)$$

Further, $\hat{N}_{ab,PML}$ is the smallest root of the quadratic equation:

$$(n_A + n_B)N^2_{ab,PML} - (n_A N_B + n_B N_A + n_A\hat{N}'_{ab} + n_B\hat{N}''_{ab})\hat{N}_{ab,PML}$$

$$+(n_A\hat{N}'_{ab}N_B + n_B\hat{N}''_{ab}N_A) = 0 \qquad (4.2.6.3)$$

The PML estimator has a simple form and uses the same set of weights for each response variable.

4.2.7 The Single Frame Multiplicity Estimator

A multiple frame estimator is on the basis of single frame multiplicity approach which does not require domain classification. Consequently they are insensitive to misclassification. In addition the proposed estimator is analytically simple so that it is easy to apply in practical applications.

Multiplicity is a characteristic of every unit $m_i = \sum_q \delta_i^{(A_q)}$ where $\delta_i^{(A_q)}$ denotes the frame membership indicator,

$$\delta_i^{(A_q)} = \begin{cases} 1 & , i \in A_q \\ 0 & , otherwise \end{cases}$$

Unit multiplicity;

m_i = the number of frames in which unit i is included.

Thereby unit multiplicity can be collected simply by asking sampled units “how many frames they belong to” (see in Figure 8).

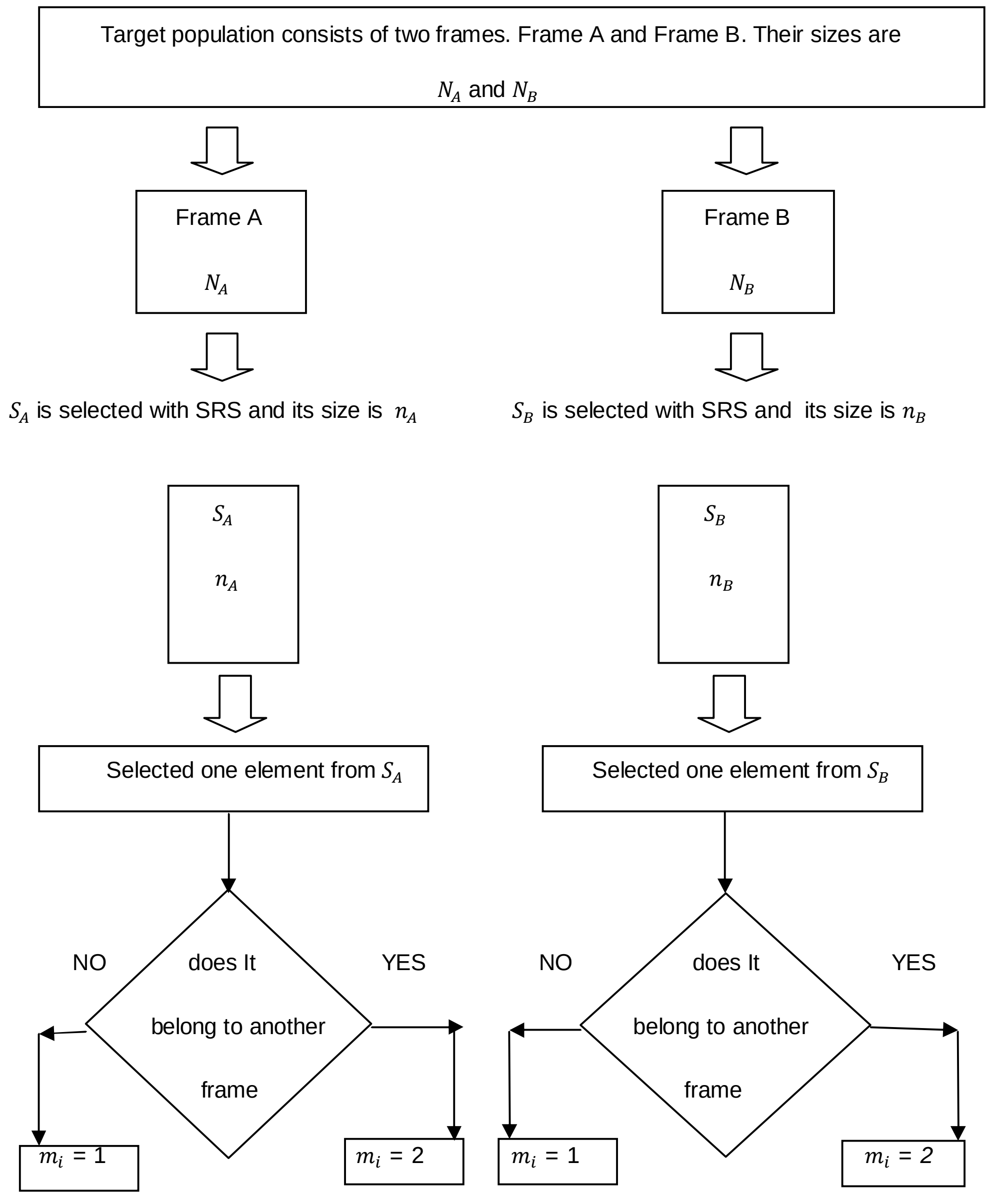


Figure 8: Algorithm for Multiplicity Units

Fulvia Mecatti (2007) proposed estimator by using multiplicity the population total to be estimated is expressed as a sum over frames instead of a sum over domains:

$$Y = \sum_{q=1}^{Q} \sum_{i \in A_q} y_i m_i^{-1}$$

As a consequence a SF multiplicity estimator given by:

$$\hat{Y}_{SFMulti} = \sum_{q=1}^{Q} \sum_{i \in S_q} w_i^{(q)} \, y_i m_i^{-1} \qquad (4.2.7.1)$$

For SRS we have fixed weight $w_i^{(q)} = f_q^{-1}, \forall i \in s_q$.

The estimator variance is given by

$$Var(\hat{Y}_{SFMulti}) = \sum_{q=1}^{Q} \frac{N_q - n_q}{n_q(N_q - 1)} \left[N_q \sum_{i \in A_q} y_i^2 m_i^{-2} - \left(\sum_{i \in A_q} y_i \, m_i^{-1} \right)^2 \right] \qquad (4.2.7.2)$$

An unbiased variance estimator for SRS of every frame is then

$$\hat{v}(\hat{Y}_{SFMulti}) = \sum_{q=1}^{Q} \frac{N_q(N_q - n_q)}{n_q^2(N_q - 1)} \left[N_q \sum_{i \in s_q} y_i^2 m_i^{-2} - f_q^{-1} \left(\sum_{i \in s_q} y_i \, m_i^{-1} \right)^2 \right] \qquad (4.2.7.3)$$

CHAPTER 5

SIMPLE RANDOM SAMPLING FROM EACH OF THREE FRAMES

In literature, most of the study about multiple frame surveys was about dual frame. Cochran (1967) proposed estimators of population sizes for three overlapping frames. Lohr and Rao (2006) presented methods that may be used combine information from more than two frames. These methods were sensitive to misclassification of units into domains, and estimates can be biased because of misclassification. Mecatti (2005 and 2007) proposed multiple frame estimator by using multiplicity approach. This estimator is insensitive to misclassification.

In three overlapping frames survey, we can write population total and size as below:

$$Y = Y_a + Y_b + Y_c + Y_{ab} + Y_{ac} + Y_{bc} + Y_{abc} \tag{5.1}$$

$$N = N_a + N_b + N_c + N_{ab} + N_{ac} + N_{bc} + N_{abc} \tag{5.2}$$

where Y_a, Y_b and Y_c are the total of the population elements in domain a, b and c (unduplicated) respectively. Y_{ab}, Y_{bc}, Y_{ac} and Y_{abc} are the total of the population elements in domain ab, bc, ac and abc (duplicated) respectively. Similarly N_a, N_b and N_c are the number of the population elements in domain a, b and c (unduplicated), respectively. N_{ab}, N_{bc}, N_{ac} and N_{abc} are the number of the population elements in domain ab, bc, ac and abc (duplicated) respectively. Let s_A , s_B and s_C are the three independent samples taken from frames A, B and C respectively. The basic question is how to estimate population total Y using three frames.

For the purpose of estimation of population total we assume that we know:

(a) n_A, n_B, n_C, n_a, n'_{ab}, $n^{`}_{ac}$, $n^{`}_{abc}$, n_b, $n^{``}_{ab}$, $n^{`}_{bc}$, $n^{``}_{abc}$, n_c, $n^{``}_{ac}$, $n^{``}_{bc}$, $n^{```}_{abc}$;

(b) y_A, y_B, y_C, y_a, y'_{ab}, $y^{`}_{ac}$, $y^{`}_{abc}$, y_b, $y^{``}_{ab}$, $y^{`}_{bc}$, $y^{``}_{abc}$, y_c, $y^{``}_{ac}$, $y^{``}_{bc}$, $y^{```}_{abc}$ the totals of y_i across units i in s_A, s_B, s_C, s_a, s'_{ab}, $s^{`}_{ac}$, $s^{`}_{abc}$, s_b, $s^{``}_{ab}$, $s^{`}_{bc}$, $s^{``}_{abc}$, s_c, $s^{``}_{ac}$, $s^{``}_{bc}$, $s^{```}_{abc}$, respectively;

(c) N_A, N_B and N_C.

Classification of sampled units to domains for three frames can be made by using Figure 9, 10 and 11.

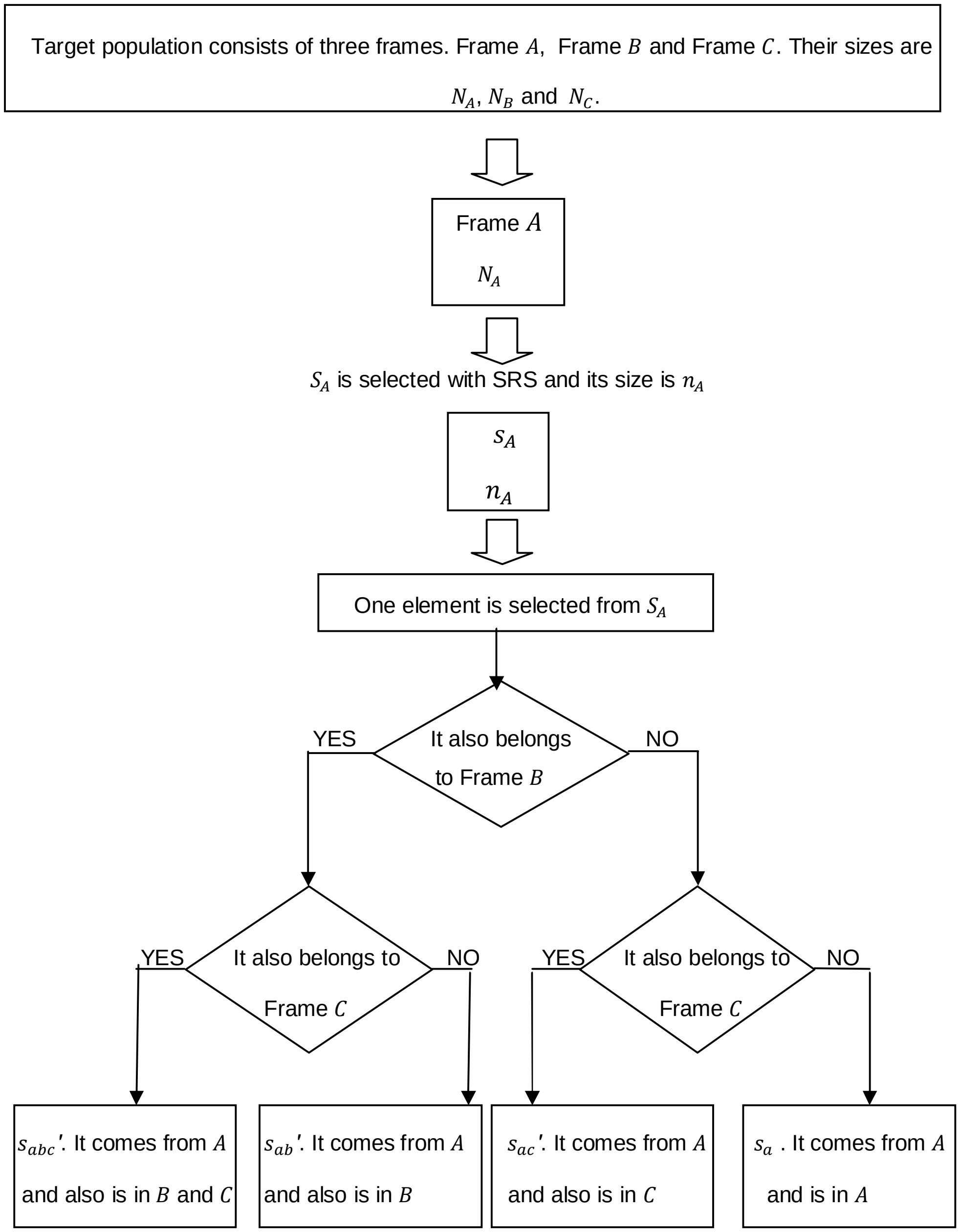


Figure 9: Algorithm for Classification of Sampled Elements Selected from Frame A to Domains for Three Frames

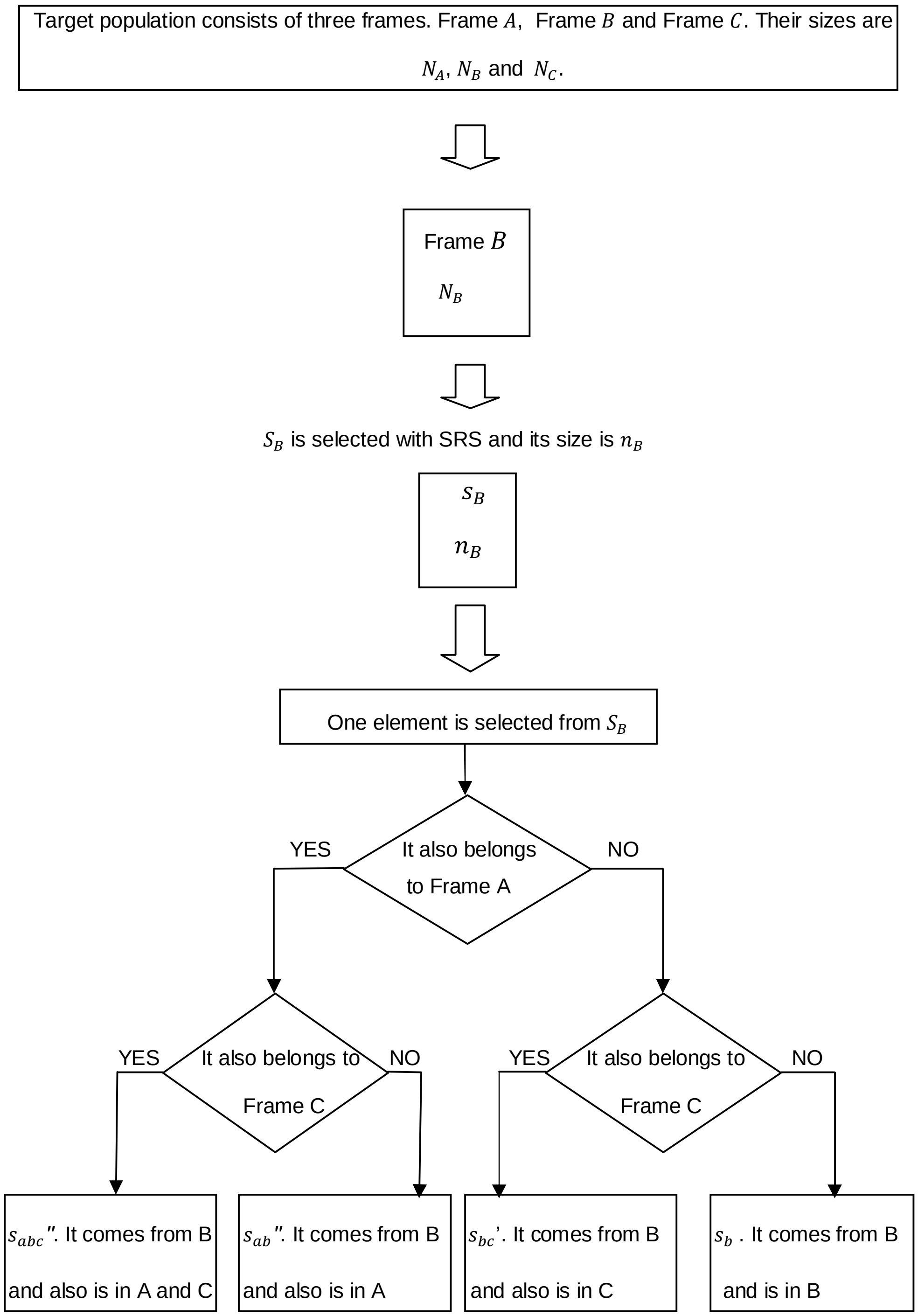


Figure 10: Algorithm for Classification of Sampled Elements Selected from Frame B to Domains for Three Frames

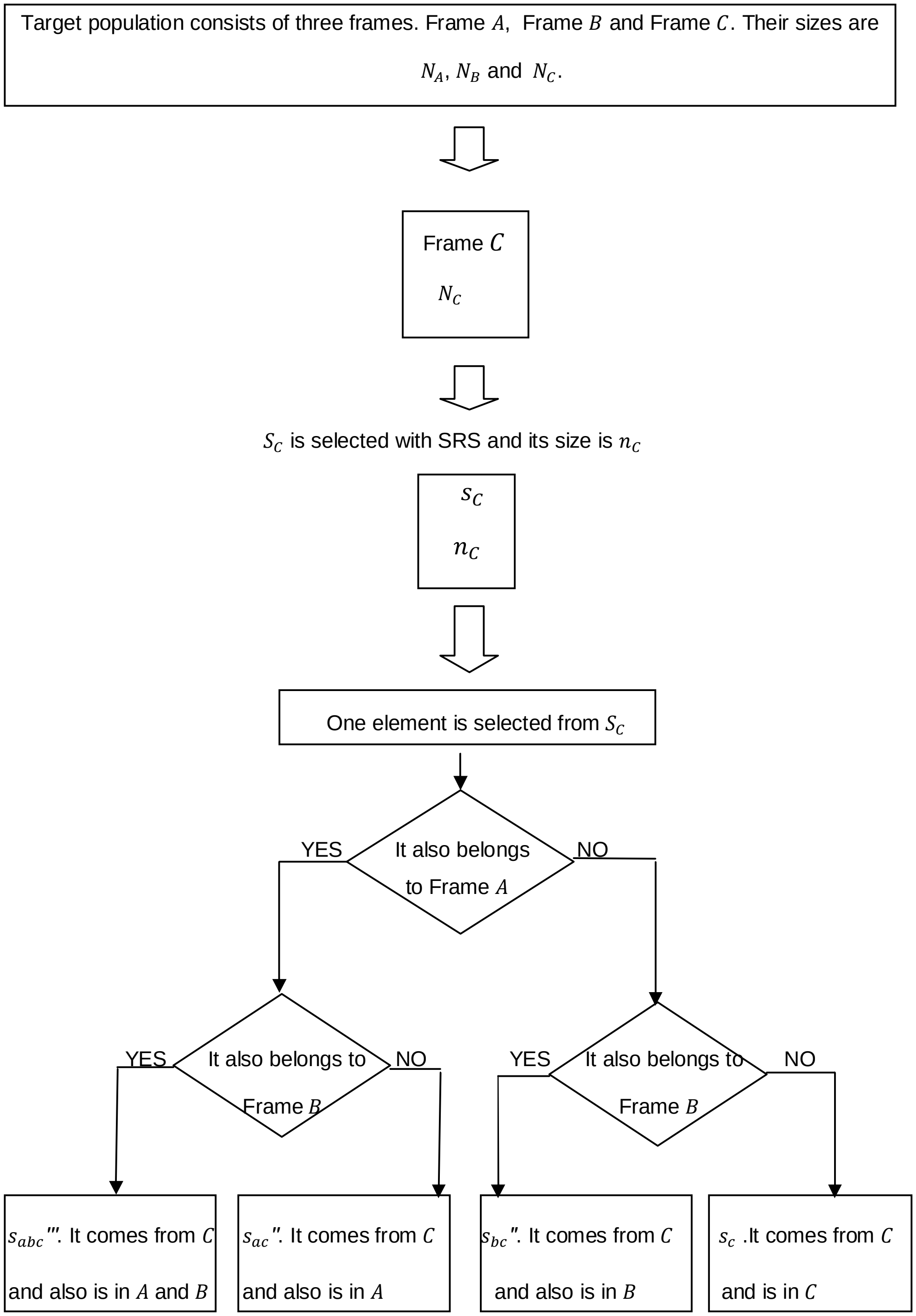


Figure 11: Algorithm for Classification of Sampled Elements Selected from Frame C to Domains for Three Frames

5.1 Cochran`s Estimator of Population Total When N_A, N_B, N_C are Known But $N_{ab}, N_{ac}, N_{bc}, N_{abc}$ are Unknown

Cochran (1967) considered the estimation of the number in the domains created by the overlapping of three frames. When sampling from three frames, A, B and C, there are $2^3 - 1 = 7$ domains. It is only necessary to directly estimate the number of units in ab, ac, bc and abc (see in Figure 4).

Using extension of the notation and procedures of the two-frame case the following estimates were obtained Cochran (1967):

$$\hat{N}_{ab,C} = p_{ab} \frac{N_A}{n_A} n_{ab}^{`} + q_{ab} \frac{N_B}{n_B} n_{ab}^{``}$$

$$\hat{N}_{ac,C} = p_{ac} \frac{N_A}{n_A} n_{ac}^{`} + q_{ac} \frac{N_C}{n_C} n_{ac}^{``}$$

$$\hat{N}_{bc,C} = p_{bc} \frac{N_B}{n_B} n_{bc}^{`} + q_{bc} \frac{N_C}{n_C} n_{bc}^{``}$$

$$\hat{N}_{abc,C} = p_A \frac{N_A}{n_A} n_{abc}^{`} + p_B \frac{N_B}{n_B} n_{abc}^{``} + p_C \frac{N_C}{n_C} n_{abc}^{```} \quad (5.1.1)$$

The variances of the quantities are

$$Var\,(\hat{N}_{ab,C}) = p_{ab}^2 \frac{N_A^2}{n_A} \alpha_1(1-\alpha_1) + q_{ab}^2 \frac{N_B^2}{n_B} \alpha_2(1-\alpha_2), \quad \alpha_1 = \frac{N_{ab}}{N_A}; \quad \alpha_2 = \frac{N_{ab}}{N_B}$$

$$Var\,(\hat{N}_{ac,C}) = p_{ac}^2 \frac{N_A^2}{n_A} \gamma_1(1-\gamma_1) + q_{ac}^2 \frac{N_C^2}{n_C} \gamma_2(1-\gamma_2), \quad \gamma_1 = \frac{N_{ac}}{N_A}; \quad \gamma_2 = \frac{N_{ac}}{N_C}$$

$$Var\,(\hat{N}_{bc,C}) = p_{bc}^2 \frac{N_B^2}{n_B} \beta_1(1-\beta_1) + q_{bc}^2 \frac{N_C^2}{n_C} \beta_2(1-\beta_2), \quad \beta_1 = \frac{N_{bc}}{N_B}; \quad \beta_2 = \frac{N_{bc}}{N_C}$$

$$Var\,(\hat{N}_{abc,C}) = p_A^2 \frac{N_A^2}{n_A} \delta_1(1-\delta_1) + P_B^2 \frac{N_B^2}{n_B} \delta_2(1-\delta_2) + P_C^2 \frac{N_C^2}{n_C} \delta_3(1-\delta_3)\,, \quad (5.1.2)$$

$$\delta_1 = \frac{N_{abc}}{N_A}; \quad \delta_2 = \frac{N_{abc}}{N_B}; \; \delta_3 = \frac{N_{abc}}{N_C}.$$

The values of the p`s that minimize these variances are:

$$Var\,(\hat{N}_{ab}) = {p_{ab}}^2 \, Var\,(\hat{N}'_{ab}) + {q_{ab}}^2 \, Var\,(\hat{N}''_{ab})$$

$$= {p_{ab}}^2 \, Var\,(\hat{N}'_{ab}) + Var\,(\hat{N}''_{ab}) - 2p_{ab}Var\,\left(\hat{N}''_{ab}\right) + {p_{ab}}^2 Var\,\left(\hat{N}''_{ab}\right)$$

$$\frac{\partial Var\,(\hat{N}_{ab})}{\partial p_{ab}} = 2p_{ab}Var\left(\hat{N}'_{ab}\right) - 2Var\left(\hat{N}''_{ab}\right) + 2p_{ab}\,Var\left(\hat{N}''_{ab}\right) = 0$$

$$p_{ab} = \frac{Var\,\left(\hat{N}''_{ab}\right)}{Var\,\left(\hat{N}'_{ab}\right) + Var\,\left(\hat{N}''_{ab}\right)}, \qquad q_{ab} = 1 - p_{ab}$$

Similarly;

$$p_{ac} = \frac{Var\,\left(\hat{N}''_{ac}\right)}{Var\,\left(\hat{N}'_{ac}\right) + Var\,\left(\hat{N}''_{ac}\right)}, \qquad q_{ac} = 1 - p_{ac}$$

$$p_{bc} = \frac{Var\,\left(\hat{N}''_{bc}\right)}{Var\,\left(\hat{N}'_{bc}\right) + Var\,\left(\hat{N}''_{bc}\right)}, \qquad q_{bc} = 1 - p_{bc}$$

$$p_A = \frac{\frac{1}{Var\,(\hat{N}'_{abc})}}{\frac{1}{Var\,(\hat{N}'_{abc})} + \frac{1}{Var\,(\hat{N}''_{abc})} + \frac{1}{Var\,(\hat{N}'''_{abc})}}$$

$$p_B = \frac{\frac{1}{Var\,(\hat{N}''_{abc})}}{\frac{1}{Var\,(\hat{N}'_{abc})} + \frac{1}{Var\,(\hat{N}''_{abc})} + \frac{1}{Var\,(\hat{N}'''_{abc})}}$$

$$p_C = \frac{\frac{1}{Var\,(\hat{N}'''_{abc})}}{\frac{1}{Var\,(\hat{N}'_{abc})} + \frac{1}{Var\,(\hat{N}''_{abc})} + \frac{1}{Var\,(\hat{N}'''_{abc})}} \tag{5.1.3}$$

The estimation of N_a, N_b and N_c can be found by

$$\hat{N}_a = N_A - (\hat{N}_{ab} + \hat{N}_{ac} + \hat{N}_{abc}),$$

$$\widehat{N}_b = N_B - (\widehat{N}_{ab} + \widehat{N}_{bc} + \widehat{N}_{abc}),$$

$$\widehat{N}_c = N_C - (\widehat{N}_{bc} + \widehat{N}_{ac} + \widehat{N}_{abc}), \qquad (5.1.4)$$

Estimation of N can be found easily using these estimators by this equation

$$\widehat{N} = \widehat{N}_a + \widehat{N}_b + \widehat{N}_c + \widehat{N}_{ab} + \widehat{N}_{ac} + \widehat{N}_{bc} + \widehat{N}_{abc}$$

Cochran (1967) shows the variances of $\widehat{N}_a$ to be

$$Var\,(\widehat{N}_a) = \frac{N_A^2}{n_A}\{p_{ab}^2\,\alpha_1(1-\alpha_1) + p_{ac}^2\,\gamma_1(1-\gamma_1) + p_A^2\,\delta_1(1-\delta_1) - 2\,p_{ab}p_{ac}\alpha_1\gamma_1$$

$$-2\,p_{ab}p_A\alpha_1\delta_1 - 2p_{ac}p_A\gamma_1\delta_1\} + \frac{N_B^2}{n_B}\{q_{ab}^2\alpha_2(1-\alpha_2) + p_B^2\delta_2(1-\delta_2)$$

$$-2p_Bq_{ab}\alpha_2\delta_2\} + \frac{N_C^2}{n_C}\{q_{ac}^2\ \gamma_2(1-\gamma_2) + p_C^2\,\delta_3(1-\delta_3)$$

$$-2\,p_Cq_{ac}\gamma_2\delta_3\} \qquad (5.1.5)$$

5.2 Extension of Fuller-Burmeister`s Estimator of Population Total When N_A, N_B, N_C are Known But $N_{ab}, N_{ac}, N_{bc}, N_{abc}$ are Unknown

In this section, using Fuller and Burmeister approach for dual frame the following estimators were derived. As mentioned in 5.1, the estimators of number of units in the areas of overlap were

$$\widehat{N}_{ab} = p_{ab}\widehat{N}_{ab}^{`} + q_{ab}\ \widehat{N}_{ab}^{``}$$

$$\widehat{N}_{ac} = p_{ac}\ \widehat{N}_{ac}^{``} + q_{ac}\ \widehat{N}_{ac}^{``}$$

$$\widehat{N}_{bc} = p_{bc}\ \widehat{N}_{bc}^{`} + q_{bc}\ \widehat{N}_{bcb}^{``}$$

$$\widehat{N}_{abc} = p_A\ \widehat{N}_{abc}^{`} + p_B\ \widehat{N}_{abc}^{``} + p_C\ \widehat{N}_{abc}^{```}$$

The variance of these estimators

$$Var(\widehat{N}_{ab}) = p_{ab}^2 Var(\widehat{N}_{ab}^{`}) \; + \; q_{ab}^2 Var(\widehat{N}_{ab}^{``})$$

$$Var(\widehat{N}_{ab}^{`}) = (1-\frac{n_A}{N_A}) \frac{N_A}{N_A-1} \frac{N_A^2}{n_A} (\frac{N_{ab}}{N_A}) (1-\frac{N_{ab}}{N_A})$$

$$= \frac{g_{A N_{ab} N_a}}{n_A} \qquad \text{where} \qquad g_A = \frac{N_A - n_A}{N_A - 1}$$

$$Var(\widehat{N}_{ab}^{``}) = (1-\frac{n_B}{N_B}) \frac{N_B}{N_B-1} \frac{N_B^2}{n_B} (\frac{N_{ab}}{N_B}) (1-\frac{N_{ab}}{N_B})$$

$$= \frac{g_{B N_{ab} N_b}}{n_B} \qquad \text{where} \qquad g_B = \frac{N_B - n_B}{N_B - 1}$$

$$p_{ab} = \frac{V(\widehat{N}_{ab}^{``})}{V(\widehat{N}_{ab}^{`}) + V(\widehat{N}_{ab}^{``})} = \frac{n_A N_b g_B}{n_A N_b g_B + n_B N_a g_A} \qquad (5.2.1)$$

To obtain an estimator of N_{ab} that is a function of the sample data only we put

$$\widehat{N}_a = N_A - \widehat{N}_{ab} \qquad , \qquad \widehat{N}_b = N_B - \widehat{N}_{ab}$$

$$\widehat{N}_{ab} = \frac{n_A N_b g_B n_{ab}^{`} N_A}{[n_A N_b g_B + n_B N_a g_A] n_A} + \frac{n_B N_a g_A n_{ab}^{``} N_B}{[n_A N_b g_B + n_B N_a g_A] n_B}$$

Thus, the estimation of N_{ab} can be seen by this quadratic equation,

$$[n_A g_B + n_B g_A] \widehat{N}_{ab}^2 - [n_A N_B g_B + n_B N_A g_A + n_{ab}^{`} N_A g_B + n_{ab}^{``} N_B g_A] \widehat{N}_{ab}$$
$$+ [n_{ab}^{`} g_B + n_{ab}^{``} g_A] N_A N_B = 0 \qquad (5.2.2)$$

$$Var(\widehat{N}_{ac}) = p_{ac}^2 Var(\widehat{N}_{ac}^{`}) \; + \; q_{ac}^2 Var(\widehat{N}_{ac}^{``})$$

$$Var(\widehat{N}_{ac}^{`}) = (1-\frac{n_A}{N_A}) \frac{N_A}{N_A-1} \frac{N_A^2}{n_A} (\frac{N_{ab}}{N_A}) (1-\frac{N_{ab}}{N_A})$$

$$= \frac{g_{A N_{ab} N_a}}{n_A} \qquad \text{where} \qquad g_A = \frac{N_A - n_A}{N_A - 1}$$

$$Var(\hat{N}_{ac}^{``}) = (1 - \frac{n_C}{N_C}) \frac{N_C}{N_C - 1} \frac{N_C^2}{n_C} (\frac{N_{ab}}{N_C}) \left(1 - \frac{N_{ab}}{N_C}\right)$$

$$= \frac{g_C N_{ab} N_c}{n_C} \qquad \text{where} \qquad g_C = \frac{N_C - n_C}{N_C - 1}$$

$$p_{ac} = \frac{V(\hat{N}_{ac}^{``})}{V(\hat{N}_{ac}^{`}) + V(\hat{N}_{ac}^{``})} = \frac{n_A N_c g_c}{n_A N_c g_C + n_C N_a g_A} \qquad (5.2.3)$$

To obtain an estimator of N_{ac} that is a function of the sample data only we put

$$\hat{N}_a = N_A - \hat{N}_{ac} \quad , \quad \hat{N}_c = N_c - \hat{N}_{ac}$$

$$\hat{N}_{ac} = \frac{n_A N_c g_c N_A n_{ac}^{`}}{[n_A N_c g_C + n_C N_a g_A] n_A} + \frac{n_C N_a g_A N_C n_{ac}^{``}}{[n_A N_c g_C + n_C N_a g_A] n_C}$$

Thus, the estimation of N_{ac} can be seen by this quadratic equation,

$$[n_A g_C + n_C g_A]\, \hat{N}_{ac}^2 - [n_A N_C g_C + n_C N_A g_A + n_{ac}^{`} N_A g_C + n_{ac}^{``} N_C g_A] \hat{N}_{ac}$$

$$+ [n_{ac}^{`} g_C + n_{ac}^{``} g_A]\, N_A N_C = 0 \qquad (5.2.4)$$

$$Var(\hat{N}_{bc}) = p_{bc}^2 Var(\hat{N}_{bc}^{`}) + q_{bc}^2 Var(\hat{N}_{bc}^{``})$$

$$Var(\hat{N}_{bc}^{`}) = (1 - \frac{n_B}{N_B}) \frac{N_B}{N_B - 1} \frac{N_B^2}{n_B} (\frac{N_{ab}}{N_B}) \left(1 - \frac{N_{ab}}{N_B}\right)$$

$$= \frac{g_B N_{ab} N_b}{n_B} \qquad \text{where} \qquad g_B = \frac{N_B - n_B}{N_B - 1}$$

$$Var(\hat{N}_{bc}^{``}) = (1 - \frac{n_C}{N_C}) \frac{N_C}{N_C - 1} \frac{N_C^2}{n_C} (\frac{N_{ab}}{N_C}) \left(1 - \frac{N_{ab}}{N_C}\right)$$

$$= \frac{g_C N_{ab} N_c}{n_C} \qquad \text{where} \qquad g_C = \frac{N_C - n_C}{N_C - 1}$$

$$p_{bc} = \frac{V(\hat{N}_{bc}^{``})}{V(\hat{N}_{bc}^{`}) + V(\hat{N}_{bc}^{``})} = \frac{n_b N_c g_c}{n_b N_c g_C + n_C N_b g_B} \qquad (5.2.5)$$

To obtain an estimator of N_{bc} that is a function of the sample data only we put

$$\widehat{N}_b = N_B - \widehat{N}_{bc} \quad , \quad \widehat{N}_c = N_C - \widehat{N}_{bc}$$

$$\widehat{N}_{bc} = \frac{n_B N_c g_c N_B n_{bc}^{`}}{[n_B N_c g_C + n_C N_b g_B] n_B} + \frac{n_C N_b g_B N_C n_{ac}^{``}}{[n_B N_c g_C + n_C N_b g_B] n_C}$$

Thus the estimation of N_{bc} can be seen by this quadratic equation,

$$[n_B g_C + n_C g_B]\, \widehat{N}_{bc}^2 - \left[n_B N_C g_C + n_C N_B g_B + n_{bc}^{`} N_B g_C + n_{bc}^{``} N_C g_B\right] \widehat{N}_{bc}$$

$$+ \left[n_{bc}^{`} g_C + n_{bc}^{``} g_B\right] N_B N_C = 0 \qquad (5.2.6)$$

$$Var(\widehat{N}_{abc}) = p_A^2 Var(\widehat{N}_{abc}^{`}) + p_B^2 Var(N_{abc}^{``}) + p_C^2 Var(N_{abc}^{```})$$

$$Var(\widehat{N}_{abc}^{`}) = (1 - \frac{n_A}{N_A}) \frac{N_A}{N_A - 1} \frac{N_A^2}{n_A} \frac{N_{abc}}{N_A} (1 - \frac{N_{abc}}{N_A})$$

$$= \frac{g_A N_{abc}(N_A - N_{abc})}{n_A}, \quad \text{where} \quad g_A = \frac{N_A - n_A}{N_A - 1}$$

$$Var(\widehat{N}_{abc}^{``}) = (1 - \frac{n_B}{N_B}) \frac{N_B}{N_B - 1} \frac{N_B^2}{n_B} \frac{N_{abc}}{N_B} (1 - \frac{N_{abc}}{N_B})$$

$$= \frac{g_B N_{abc}(N_B - N_{abc})}{n_B}, \quad \text{where} \quad g_B = \frac{N_B - n_B}{N_B - 1}$$

$$Var(\widehat{N}_{abc}^{```}) = (1 - \frac{n_C}{N_C}) \frac{N_C}{N_C - 1} \frac{N_C^2}{n_C} \frac{N_{abc}}{N_C} (1 - \frac{N_{abc}}{N_C})$$

$$= \frac{g_C N_{abc}(N_C - N_{abc})}{n_C}, \quad \text{where} \quad g_C = \frac{N_C - n_C}{N_C - 1} \qquad (5.2.7)$$

Using equations in (5.1.3) the values of p_A, p_B and p_C were found as follows.

$$p_A = \frac{\frac{n_A}{g_A N_{abc}(N_A - N_{abc})}}{\frac{n_A}{g_A N_{abc}(N_A - N_{abc})} + \frac{n_B}{g_B N_{abc}(N_B - N_{abc})} + \frac{n_C}{g_C N_{abc}(N_C - N_{abc})}}$$

$$= \frac{n_A}{g_A(N_A - N_{abc})} \frac{g_A(N_A - N_{abc})g_B(N_B - N_{abc})g_C(N_C - N_{abc})}{A}$$

where

$$A = n_A g_B g_C(N_B - N_{abc})(N_C - N_{abc}) + n_B g_A g_C(N_A - N_{abc})(N_C - N_{abc}) + n_C g_A g_B(N_A - N_{abc})(N_B - N_{abc})$$

$$A = n_A g_B g_C\left[N_B N_C - N_B N_{abc} - N_C N_{abc} + N_{abc}^2\right] + n_B g_A g_C\left[N_A N_C - N_C N_{abc} - N_A N_{abc} + N_{abc}^2\right] + n_C g_A g_B\left[N_B N_A - N_A N_{abc} - N_B N_{abc} + N_{abc}^2\right]$$

$$A = n_A g_B g_C N_B N_C - n_A g_B g_C N_B N_{abc} - n_A g_B g_C N_C N_{abc} + n_A g_B g_C N_{abc}^2$$
$$+ n_B g_A g_C N_A N_C - n_B g_A g_C N_C N_{abc} - n_B g_A g_C N_A N_{abc} + n_B g_A g_C N_{abc}^2$$
$$+ n_C g_A g_B N_B N_A - n_C g_A g_B N_A N_{abc} - n_C g_A g_B N_B N_{abc} + n_C g_A g_B N_{abc}^2$$

$$A = [n_A g_B g_C + n_B g_A g_C + n_C g_A g_B]N_{abc}^2$$
$$- [\, g_B g_C(n_A N_B + n_A N_C) + g_A g_C(n_B N_C + n_B N_A) + g_A g_B(n_C N_A + n_C N_B)]N_{abc}$$
$$+ [n_A g_B g_C N_B N_C + n_B g_A g_C N_A N_C + n_C g_A g_B N_B N_A]$$

$$p_A = \frac{n_A g_B g_C(N_B - N_{abc})(N_C - N_{abc})}{A} \quad (5.2.8)$$

Similarly, p_B and p_C were as follows

$$p_B = \frac{n_B g_A g_C(N_A - N_{abc})(N_C - N_{abc})}{A} \quad (5.2.9)$$

$$p_C = \frac{n_C g_B g_A(N_B - N_{abc})(N_A - N_{abc})}{A} \quad (5.2.10)$$

To obtain an estimator of N_{abc} that is a function of the sample data, p_A, p_B and p_C were only put into formula of $\widehat{N}_{abc}$.

$$\hat{N}_{abc} = \frac{n_A g_B g_C (N_B-\hat{N}_{abc})(N_C-\hat{N}_{abc})}{A}\,\frac{N_A}{n_A}\,n^{`}_{abc}$$

$$+\frac{n_B g_A g_C (N_A-\hat{N}_{abc})(N_C-\hat{N}_{abc})}{A}\,\frac{N_B}{n_B}\,n^{``}_{abc}$$

$$+\frac{n_C g_B g_A\,(N_B-\hat{N}_{abc})(N_A-\hat{N}_{abc})}{A}\,\frac{N_C}{n_C}\,n^{```}_{abc}$$

$\hat{N}_{abc} = \frac{B}{A}$ where

$$B = g_B g_C\left[N_B N_C N_A n^{`}_{abc} - N_B N_A \hat{N}_{abc} n^{`}_{abc} - N_C N_A \hat{N}_{abc} n^{`}_{abc} + N_A \hat{N}^2_{abc} n^{`}_{abc}\right]$$
$$+ g_A g_C\left[N_A N_C N_B n^{``}_{abc} - N_C N_B \hat{N}_{abc} n^{``}_{abc} - N_A N_B \hat{N}_{abc} n^{``}_{abc} + N_B \hat{N}^2_{abc} n^{``}_{abc}\right]$$
$$+ g_A g_B\left[N_A N_B N_C n^{```}_{abc} - N_C N_B \hat{N}_{abc} n^{```}_{abc} - N_C N_A \hat{N}_{abc} n^{```}_{abc} + N_C \hat{N}^2_{abc} n^{```}_{abc}\right]$$

$$B = \left[g_B g_C N_A n^{`}_{abc} + g_A g_C N_B n^{``}_{abc} + g_A g_B N_C n^{```}_{abc}\right]\hat{N}^2_{abc} - \left[g_B g_C (N_B N_A n^{`}_{abc} + N_C N_A n^{`}_{abc}) + g_A g_C (N_C N_B n^{``}_{abc} + N_C N_B n^{``}_{abc}) + g_A g_B (N_C N_B n^{```}_{abc} + N_C N_A n^{```}_{abc})\right]\hat{N}_{abc} + N_A N_B N_C\left[g_B g_C n^{`}_{abc} + g_A g_C n^{``}_{abc} + g_A g_B n^{``}_{abc}\right]$$

$$\hat{N}_{abc} A = B$$

$$[n_A g_B g_C + n_B g_A g_C + n_C g_A g_B]\hat{N}^3_{abc} - [g_B g_C (n_A N_B + n_A N_C) + g_A g_C (n_B N_C + n_B N_A) + g_A g_B (n_C N_A + n_C N_B)] + [n_A g_B g_C N_B N_C + n_B g_A g_C N_A N_C + n_C g_A g_B N_B N_A]\hat{N}_{abc} = \left[g_B g_C N_A n^{`}_{abc} + g_A g_C N_B n^{``}_{abc} + g_A g_B N_C n^{```}_{abc}\right]\hat{N}^2_{abc} - \left[g_B g_C (N_B N_A n^{`}_{abc} + N_C N_A n^{`}_{abc}) + g_A g_C (N_C N_B n^{``}_{abc} + N_A N_B n^{``}_{abc}) + g_A g_B (N_C N_B n^{```}_{abc} + N_C N_A n^{```}_{abc})\right]\hat{N}_{abc} + N_A N_B N_C\left[g_B g_C n^{`}_{abc} + g_A g_C n^{``}_{abc} + g_A g_B n^{``}_{abc}\right] \quad (5.2.10)$$

Hence, from (5.2.10) we can find $\hat{N}_{abc}$ by solving the following equation

$$[n_A g_B g_C + n_B g_A g_C + n_C g_A g_B]\hat{N}^{\mathbf{3}}_{\mathbf{abc}} - \left[g_B g_C (n_A N_B + n_A N_C + N_A n^{`}_{abc}) + g_A g_C\left(n_B N_C + n_B N_A + N_B n^{``}_{abc}\right) + g_A g_B (n_C N_A + n_C N_B + N_C n^{```}_{abc})\right]\hat{N}^{\mathbf{2}}_{\mathbf{abc}} + \left[g_B g_C (n_A N_B N_C + N_B N_A n^{`}_{abc} + N_C N_A n^{`}_{abc}) + g_A g_C (n_B N_A N_C + N_A N_B n^{``}_{abc} + N_C N_B n^{``}_{abc}) + g_A g_B (n_C N_B N_A + N_C N_B n^{```}_{abc} + N_C N_A n^{```}_{abc})\right]\hat{N}_{\mathbf{abc}} - N_A N_B N_C = 0$$

(5.2.11)

CHAPTER 6

APPLICATION OF MULTIPLE FRAME METHODS WITH ARTIFICAL DATA

A simple example has been made with an artificial three frames presented in Table A.1, A.2 and A.3 in Appendix A. These frames consist of nine times enlarged elements presented in Table 2. 01-02-03-04-05-06-07-08-09 was suffixed to these enlarged elements. In this thesis, the population element values were presented to compare with the expanded population total estimates. However generally, population element values (measurement or count) are not known to the researcher. Different estimators were calculated for estimation of population size and total for two overlapping frames A and B shown in Chapter 4 and for three overlapping frames A, B and C shown in Chapter 5. Sizes of frames were N_A = 300 and N_B = 250 and N_C = 210. In these examples, sample A, sample B and sample C presented in Table D.1, D.2 and D.3 in Appendix D were selected with simple random selection without replacement. Their sizes were n_A = 60, n_B = 50 and n_C = 30.

6.1 An Example For Dual Frame

SITUATION 1: In this situation, N_A, N_B and N_{ab} was known. In generally the knowledge of the domain sizes is a very restrictive assumption that is seldom verified. In this thesis, in order to calculate estimation of population total for situation 1, frame A has been matched with Frame B according to their ID Label to classify population elements to three domains. As results of this matching, population elements classification to three domains a, ab and b can be seen in Table B.1, B.2 and B.3 in Appendix B. The sizes, totals, variances of frame A, frame B, domains a, b and ab were as found below:

N_A = 300,	N_B = 250,	
Y_A = 790,	Y_B = 650,	
N_a = 150,	N_{ab} = 150,	N_b = 100
Y_a = 380,	Y_{ab} = 410,	Y_b = 240
$\bar{Y}_a$ = 2.533,	$\bar{Y}_{ab}$ = 2.733,	$\bar{Y}_b$ = 2.4
S_a^2 = 2.5324	S_{ab}^2 = 4.6263	S_b^2 = 1.4545
σ_a^2 = 2.5155,	σ_{ab}^2 = 4.5954	σ_b^2 = 1.4399

$N_a = N_A - N_{ab}$ = 300 – 150 = 150

$N_b = N_B - N_{ab}$ = 250 – 150 = 100

The size and total of population was obtained as follows:

$N = N_A + N_B - N_{ab} = N_a + N_b + N_{ab}$

N = 300 + 250 – 150 = 150 + 100 + 150

N = 400

$Y = Y_A + Y_B - Y_{ab} = Y_a + Y_b + Y_{ab}$

Y = 790 + 650 – 410 = 380 + 240 + 410

$\boldsymbol{Y = 1030}$

Each sample was classified into two domains by using algorithm in Figure 7. The results of classification sample A elements to domains a and ab were shown in Table E.1 in Appendix E and results of classification Sample B elements to domains a and ab were shown in Table E.2 in Appendix E. $\grave{n}_{ab}$ was the number of elements sampled from frame A and in the domain ab. $\grave{\bar{y}}_{ab}$ was the mean of these elements. Similarly $n_{ab}^{``}$ was the number of elements sampled from frame $B$ and in the domain $ab$. $\bar{y}_{ab}^{``}$ was the mean of these elements. The sizes, totals and means of independent simple random samples s_A , s_B and sample domains s_a, s_b,s_{ab} were found as below:

n_A = 60, n_B = 50

y_A = 180, y_B = 115,

n_a = 24 , n_b = 20

y_a = 66 , y_b = 45

$\bar{y}_a = y_a/n_a$=66/24 = 2.75 $\bar{y}_b = y_b/n_b$= 45/20 = 2.25

$n_{ab}^{`}$ = 36 , $n_{ab}^{``}$ = 30

$y_{ab}^{`}$ = 114 $y_{ab}^{``}$ = 70

$n_{ab} = n_{ab}^{`} + n_{ab}^{``}$ = 36 + 30 = 66

$\bar{y}_{ab}^{`} = y_{ab}^{`}/n_{ab}^{`}$ = 114/36 = 3.167,

$\bar{y}_{ab}^{``} = y_{ab}^{``}/n_{ab}^{``}$ = 70/30 = 2.333.

$y_{ab} = y_{ab}^{`} + y_{ab}^{``}$ = 114 + 70 = 184 ,

$\bar{y}_{ab} = y_{ab}/n_{ab}$ = 184/66 = 2.788,

Hartley`s unbiased estimator for the population total was

$$\hat{Y}_H = N_a\bar{y}_a + N_{ab}p\bar{y}_{ab}^{'} + N_{ab}q\bar{y}_{ab}^{''} + N_b\bar{y}_b$$

Lund(1968) showed Hartley`s bi-quadratic p value ,with a simple expression as below:

$$p_H = \frac{\alpha n_A}{\alpha n_A + \beta n_B} = \frac{(0.5)60}{(0.5)60+(0.6)50} = 0.50 \qquad q = 1 - p = 1 - 0.50 = 0.50$$

where $\alpha = N_{ab}/N_A$, $\beta = N_{ab}/N_B$,

α = 150/300 = 0.5 , β = 150/250 = 0.6,

$\hat{Y}_H$ = 150(2.75) + 150(0.50)(3.167) + 150(0.50)(2.333) + 100(2.25)

$\boldsymbol{\hat{Y}_H = 1050}$

Response Bias (RB) = $Y - \hat{Y}_H$ = 1030 – 1050

RB= $\boldsymbol{-20}$

The variance of this estimator was calculated as below:

$$Var\ (\hat{Y}_H) = \frac{N_A^2}{n_A}\left[(1-\alpha)\sigma_a^2 + \alpha p^2\sigma_{ab}^2\right] + \frac{N_B^2}{n_B}\left[(1-\beta)\sigma_b^2 + \beta q^2\sigma_{ab}^2\right]$$

$$Var\ (\hat{Y}_H) = \frac{300^2}{60}\ [(1-0.5)(2.5155) + (0.5)(0.50)^2(4.5954)] + \frac{250^2}{50}$$

$$[(1-0.6)(1.4399) + (0.6)(0.50)^2(4.5954)]$$

$$= 1500(1.8322) + 1250(1.2653) = 4329.93$$

$\boldsymbol{Var\ (\hat{Y}_H) = 4329.93}$

When N_A, N_B, N_{ab} were known, Lund`s estimator for population total was as calculated below:

$$\hat{Y}_L = N_a\bar{y}_a + N_{ab}\bar{y}_{ab} + N_b\bar{y}_b$$

$\hat{Y}_L$ = 150(2.75) + 150(2.788) + 100(2.25)

$\boldsymbol{\hat{Y}_L = 1055.7}$

Response Bias (RB) = $Y - \hat{Y}_L$ = 1030 – 1055,7

RB = $\boldsymbol{-25.7}$

The variance of this estimator was

$$Var\ (\hat{Y}_L) = \frac{N_A^2}{n_A}(1-\alpha)\sigma_a^2 + \frac{N_A N_B \alpha\beta}{\alpha n_A + \beta n_B}\sigma_{ab}^2 + \frac{N_B^2}{n_B}(1-\beta)\sigma_b^2$$

$$Var\ (\hat{Y}_L) = \frac{300^2}{60}\ (1 - 0.5)(2.5155) + \frac{300(250)(0.5)(0.6)}{(0.5)60 + (0.6)50}\ (4.5954)$$

$$+ \frac{250^2}{50}\ (1 - 0.6)(1.4399)$$

$$= 1500(1.258) + 375(4.5954) + 1250(0.576)$$

$$= 1887 + 1723.28 + 720$$

$$\boldsymbol{Var\ (\hat{Y}_L) = 4330.28.}$$

As a result of this application, it was seen that $Var\ (\hat{Y}_H) \leq Var\ (\hat{Y}_L)$ and response bias of Hartley`s estimator is also smaller than Lund`s estimator. Thus, using Hartley`s estimator can be better than using Lund`s estimator for Situation 1.

SITUATION 2: In this case, it was assumed that N_{ab} was unknown and duplicated units (i.e.,$s_A \cap s_B$) might not be identified. So, estimates of N_a, N_b and N_{ab} should be calculated for estimation of size of population. After estimating N_{ab}, the estimates of N_a and N_b might be found by subtraction as

$$\hat{N}_a = N_A - \hat{N}_{ab} \qquad \text{and} \qquad \hat{N}_b = N_B - \hat{N}_{ab}.$$

When N_a, N_b, N_{ab} were unknown, Hartley`s estimator for population total was as calculated below:

$$\hat{Y}_H = \frac{N_A}{n_A}\left\{y_a + py'_{ab}\right\} + \frac{N_B}{n_B}\left\{y_b + qy''_{ab}\right\} = \frac{300}{60}\{66 + (0.5)114\}$$

$$+ \frac{250}{50}\{45 + (0.5)70\}$$

$$\boldsymbol{\hat{Y}_H = 1015}$$

Response Bias (RB) = $Y - \hat{Y}_H$ = 1030 – 1015

RB= $\mathbf{15}$

The variance of this estimator was as below:

$$Var(\hat{Y}_H) = \frac{N_A^2}{n_A}\left\{(1-\alpha)\sigma_a^2 + \alpha p^2\sigma_{ab}^2 + \alpha(1-\alpha)(\bar{Y}_a - p\bar{Y}_{ab})^2\right\}$$

$$+ \frac{N_B^2}{n_B}\left\{(1-\beta)\sigma_b^2 + \beta q^2\sigma_{ab}^2 + \beta(1-\beta)(\bar{Y}_b - q\bar{Y}_{ab})^2\right\}$$

$$Var(\hat{Y}_H) = \frac{300^2}{60}\{(1-0.5)(2.5155) + (0.5)(0.5)^2(4.5954) + 0.5(1-0.5)$$

$$(2.533 - (0.5)2.733)^2\} + \frac{250^2}{50}\{(1-0.6)(1.4399)$$

$$+(0.6)(0.50)^2(4.5954) + 0.6(1-0.6)(2.4 - (0.50)2.733)^2\}$$

$$= 1500\{1.2578 + 0.5744 + 0.3402\} + 1250\{0.5760 + 0.6893 + 0.2563\}$$

$$= 1500(2.1724) + 1250(1.5216)$$

$$\boldsymbol{Var(\hat{Y}_H) = 5160.60}$$

Lund`s estimator for population total for situation 2 was calculated as below:

$$\hat{Y}_L = \frac{N_A}{n_A}n_a\bar{y}_a + \left\{\frac{N_A}{n_A}\,n_{ab}^{`}\,p + \frac{N_B}{n_B}\,n_{ab}^{``}\,(1-p)\right\}\bar{y}_{ab} + \frac{N_B}{n_B}n_b\bar{y}_b$$

where an estimator of optimum p from the sample data was

$$\hat{p}_L = \frac{\frac{N_A n_a}{n_A^2}\bar{y}_a + \frac{N_B n_b}{n_B^2}(\bar{y}_{ab}-\bar{y}_b)}{\left(\frac{N_A n_a}{n_A^2} + \frac{N_B n_b}{n_B^2}\right)\bar{y}_{ab}} = \frac{\frac{300(24)}{60^2}(2.75) + \frac{250(20)}{50^2}(2.788-2.25)}{\left(\frac{300(24)}{60^2} + \frac{250(20)}{50^2}\right)2.788}$$

$$= \frac{6.576}{11.152}$$

$\hat{p}_L$ = 0.589 where $\hat{q}_L$ = 1 – 0.589 = 0.411

$\hat{Y}_L$ = $\frac{300}{60}$ 24(2.75) + $\left\{\frac{300}{60}\right.$ 36 (0.589) + $\frac{250}{50}$ 30 (1 – 0.589)$\left.\right\}$ (2.788)

+ $\frac{250}{50}$ 20(2.25)

= 330 + {106.02 + 61.65}2.787 + 225

$\boldsymbol{\hat{Y}_L} = \mathbf{1022.30}$

Response Bias (RB) = $Y - \hat{Y}_L$ = 1030 – 1022.30

RB = $\mathbf{7.7}$

Variance of Lund`s estimator was

$$Var(\hat{Y}_L) = \frac{N_A^2}{n_A}(1-\alpha)\sigma_a^2 + \frac{N_A N_B \alpha\beta}{\alpha n_A + \beta n_B}\sigma_{ab}^2 + \frac{N_B^2}{n_B}(1-\beta)\sigma_b^2$$

$$+ \frac{N_A^2(1-\alpha)\alpha}{n_A}(\bar{Y}_a - p\bar{Y}_{ab})^2 + \frac{N_B^2(1-\beta)\beta}{n_B}(\bar{Y}_b - q\bar{Y}_{ab})^2$$

$$Var(\hat{Y}_L) = \frac{300^2}{60}(1-0.5)(2.5155) + \frac{300\ 250(0.5)(0.6)}{(0.5)60+(0.6)50}(4.5954)$$

$$+ \frac{250^2}{50}(1-0.6)(1.4399) + \frac{300^2(1-0.5)0.5}{60}(2.533-(0.589)(2.733))^2$$

$$+ \frac{250^2(1-0.6)0.6}{50}(2.4-(0.411)(2.733))^2$$

= 1886.625 + 1723.275 + 719.95 + 319.655 + 480.017

$\boldsymbol{Var(\hat{Y}_L)} = \mathbf{5138.54}$

According to Fuller and Burmeister (1972), when N_a, N_b, N_{ab} were unknown, estimation of population total was

$$\hat{Y}_{FB} = (N_A - \hat{N}_{ab,FB})\,\bar{y}_a + \hat{N}_{ab,FB}\bar{y}_{ab} + (N_B - \hat{N}_{ab,FB})\,\bar{y}_b$$

where the estimation of N_{ab} was the left (smallest) root of equation given by

$$[n_A g_B + n_B g_A]\,\hat{N}^2_{ab,FB} - [n_A N_B g_B + n_B N_A g_A + n^{`}_{ab} N_A g_B + n^{``}_{ab} N_B g_A]\hat{N}_{ab,FB}$$

$$+ [n^{`}_{ab} g_B + n^{``}_{ab} g_A]\, N_A N_B = 0$$

and $\bar{y}_{ab} = p\bar{y}^{`}_{ab} + (1-p)\bar{y}^{``}_{ab}$ where $p_{FB} = \frac{n^{`}_{ab}(1-f_B)}{n^{`}_{ab}(1-f_B)+n^{``}_{ab}(1-f_A)}$.

$g_A = \frac{300-60}{300-1} = 0.802$ and $g_B = \frac{250-50}{250-1} = 0.803.$

$$[60(0.803 + 50(0.802)]\,\hat{N}^2_{ab,FB} - [60\,250(0.803) + 50\,300(0.802) +$$
$$36\,300(0.803) + 30\,250(0.802)]\hat{N}_{ab,FB} + [36(0.803) + 30(0.802)]\,250\,300 = 0$$

$$88.28\hat{N}^2_{ab,FB} - 38762.4\hat{N}_{ab,FB} + 3972600 = 0$$

$$\hat{N}_{ab,FB}{}^{1} = \frac{-b-\sqrt{b^2-4ac}}{2a} = \frac{38762.4-\sqrt{38762.4^2-4(88{,}28)3972600}}{2(88.28)}$$

$$= 162.98 \cong 163$$

$$\hat{N}_{ab,FB}{}^{2} = \frac{-b+\sqrt{b^2-4ac}}{2a} = \frac{38762.4+\sqrt{38762.4^2-4(88{,}28)3972600}}{2(88{,}28)}$$

$$\hat{N}_{ab,FB}{}^{2} = 276.10 \cong 276$$

$$\boldsymbol{\hat{N}_{ab,FB} = 163}$$

Estimation of N_a was

$$\hat{N}_{a,FB} = N_A - \hat{N}_{ab,FB} = 300 - 163 = 137$$

Estimation of N_b was

$$\hat{N}_{b,FB} = N_B - \hat{N}_{ab,FB} = 250 - 163 = 87$$

Estimation of N was

$$\widehat{N}_{FB} = N_A + N_B - \widehat{N}_{ab,FB} = N_{a,FB} + N_{b,FB} + \widehat{N}_{ab,FB}$$

$$= 300 + 250 - 163 = 137 + 87 + 163 = 387$$

$$f_A = \frac{n_A}{N_A} = \frac{60}{300} = 0.2, \quad f_B = \frac{n_B}{N_B} = \frac{50}{250} = 0.2$$

$$p_{FB} = \frac{36(1-0.2)}{36(1-0.2)+30(1-0.2)} = \frac{28.8}{52.8} = 0.545$$

$$\bar{y}_{ab} = 0.545(3.167) + (1 - 0.545)(2.333)$$

$$\bar{y}_{ab} = 2.788$$

$$\hat{Y}_{FB} = (300 - 163)(2.75) + 163(2.788) + (250 - 163)2.25$$

$$\hat{Y}_{FB} = 376.75 + 454.44 + 195.75$$

$$\boldsymbol{\hat{Y}_{FB} = 1026.95}$$

Response Bias (RB) = $Y - \hat{Y}_{FB}$ = 1030 – 1026.95

RB = $\mathbf{3.05}$

The variance of this estimator was

$$Var(\hat{Y}_{FB}) = N_a(f_A^{-1} - 1)S_a^2 + [(1 - f_B)f_A + (1 - f_A)f_B]^{-1}$$

$$(1 - f_A)(1 - f_B)N_{ab}S_{ab}^2 + N_b(f_B^{-1} - 1)S_b^2$$

$$+(\bar{Y}_{ab} - \bar{Y}_a - \bar{Y}_b)^2 \frac{N_{ab}N_aN_bg_Ag_B}{n_AN_bg_B + n_BN_ag_A}$$

$$Var(\hat{Y}_{FB}) = 150(5 - 1)(2.5324) + [(1 - 0.2)0.2 + (1 - 0.2)0.2]^{-1}$$

$$(1 - 0.2)(1 - 0.2)150(4.6263) + 100(5 - 1)(1.4545)$$

$$+(2.733 - 2.533 - 2.4)^2 \frac{150\ 150\ 100\ (0.802)(0.803)}{60\ 100\ (0.803) + 50\ 150\ (0.802)}$$

$Var\ (\hat{Y}_{FB})$ = 1521.24 + (3.125)(444.125) + 581.8 + (4.84)(133.759)

$\boldsymbol{Var\ (\hat{Y}_{FB})}$ = $\mathbf{4138.32}$

When N_a, N_b, N_{ab} were unknown, SF estimator of population total was calculated as follows:

$\hat{Y}_{SF}$ = $f_A^{-1}y_a$ + $(f_A + f_B)^{-1}y_{ab}$ + $f_B^{-1}y_b$

where

y_{ab} = $y_{ab}^{`}$ + $y_{ab}^{``}$ = 114 + 70 = 184

$\hat{Y}_{SF}$ = 5(66) + (0.2 + 0.2)$^{-1}$184 + 5(45)

$\hat{Y}_{SF}$ = 330 + 460 + 225

$\boldsymbol{\hat{Y}_{SF}}$ = $\mathbf{1015}$

Response Bias (RB) = $Y - \hat{Y}_{SF}$ = 1030 – 1015

RB= $\mathbf{15}$

In Situation 2, $Var\ (\hat{Y}_{FB})$ < $Var\ (\hat{Y}_L)$ < $Var\ (\hat{Y}_H)$

$\hat{Y}_{SF}$ was a special case of $\hat{Y}_H$ or of $\hat{Y}_L$ with $p = f_A(f_A + f_B)^{-1}$. Thus

Var ($\hat{Y}_{SF}$) > Var ($\hat{Y}_{FB}$) Skinner (1991).

Skinner`s raking ratio (RR) estimator for population total was as calculated below:

$\hat{Y}_{RR}^{(\infty)}$ = $\left(N_A - \hat{N}_{ab,RR}\right)\bar{y}_a$ + $\hat{N}_{ab,RR}\bar{y}_{ab}$ + $(N_B - \hat{N}_{ab,RR})\bar{y}_b$,

where $\tilde{N}_{ab,RR}$ was the smallest root of the quadratic equation

$n_{ab}\hat{N}_{ab,RR}{}^2 - [n_{ab}(N_A + N_B) + (f_A^{-1} + f_B^{-1})n_a n_b]\hat{N}_{ab,RR} + n_{ab}N_A N_B = 0$

and

$$66\widehat{N}_{ab,RR}{}^{2} - [66(300 + 250) + (5 + 5)24\ 20]\widehat{N}_{ab,RR} + 66(300)250 = 0$$

$$66\widehat{N}_{ab,RR}{}^{2} - [41100]\widehat{N}_{ab,RR} + 4950000 = 0$$

$$\widetilde{N}_{ab,RR}{}^{1} = \frac{-b-\sqrt{b^2-4ac}}{2a} = \frac{41100-\sqrt{41100^2-4(66)4950000}}{2(66)}$$

$$\widetilde{N}_{ab,RR}{}^{1} = 163.21 \cong 163$$

$$\widehat{N}_{ab,RR}{}^{2} = \frac{-b+\sqrt{b^2-4ac}}{2a} = \frac{41100+\sqrt{41100^2-4(66)4950000}}{2(66)}$$

$$\widehat{N}_{ab,RR}{}^{2} = 459.51 \cong 460$$

$$\boldsymbol{\widehat{N}_{ab,RR} = 163}$$

$$\widehat{Y}_{RR}^{(\infty)} = (300 - 163)2.75 + 163(2.788) + (250 - 163)2.25$$

$$\mathbf{\widehat{Y}_{RR}^{(\infty)} = 1026.95}.$$

Response Bias (RB) = $Y - \widehat{Y}_{RR}$ = 1030 – 1026.95

RB = $\mathbf{3.05}$

The variance of this estimator was

$$Var(\widehat{Y}_{RR}^{(\infty)}) \sim f_A^{-1}N_aS_a^2 + f_B^{-1}N_bS_b^2 + (f_A + f_B)^{-1}N_{ab}S_{ab}^2$$

$$+(\bar{Y}_a + \bar{Y}_b - \bar{Y}_{ab})^2\ \frac{N_{ab}N_aN_b}{n_AN_b+n_BN_a}(1+\lambda^2),$$

$$Var\ (\widehat{Y}_{RR}^{(\infty)}) \sim 5(150)2.5324 + 5(100)1.4545 + (0.2 + 0.2)^{-1}150(4.6263)$$

$$+(2.533 + 2.4 - 2.733)^2\ \frac{150\ 150\ 100}{60\ 100+50\ 150}\ (1 + 0.0014)$$

$$Var\ (\widehat{Y}_{RR}^{(\infty)}) \sim 1899.3 + 727.25 + 1734.87 + 807.80$$

$$\boldsymbol{Var\,(\hat{Y}_{RR}^{(\infty)})\sim 5169.22}$$

where

$$\lambda^2 = \frac{150\ 100\ 163^2(0.2^2 300-0.2^2 250)^2}{300\ 250(0.2)(0.2)(0.2+0.2)^2(163^2-300\ 250)^2} = \frac{1594140000}{480(2345561761)}$$

$$= 0.0014$$

$Var\,(\hat{Y}_{FB}) < Var\,(\hat{Y}_{RR}^{(\infty)})$.

Fulvia Mecatti (2007) proposed SF multiplicity estimator as

$$\hat{Y}_{SFMulti} = \sum_{q=1}^{Q} \sum_{i\in S_q} w_i^{(q)}\, y_i m_i^{-1}$$

where $w_i^{(q)} = f_q^{-1}$, $\forall i \in s_q$.

In this method, unit multiplicity was collected by asking sampled units "how many frames they belong to" as in Figure 9, results were shown in Table G.1 and G.2 in Appendix G.

$$\hat{Y}_{SFMulti} = f_A^{-1}\left(\sum_{i\in S_A} y_i m_i^{-1}\right) + f_B^{-1}\left(\sum_{i\in S_B} y_i m_i^{-1}\right)$$

$\hat{Y}_{SFMulti}$ = 5(123) + 5(80) = 1015

$\boldsymbol{\hat{Y}_{SFMulti} = 1015}$

Response Bias (RB) = $Y - \hat{Y}_{FMulti}$ = 1030 – 1015

RB = $\mathbf{15}$

For SRS of every frame the estimator variance was

$$Var(\hat{Y}_{SFMulti}) = \sum_{q=1}^{Q} \frac{N_q - n_q}{n_q(N_q-1)}\left[N_q \sum_{i\in A_q} y_i^2 m_i^{-2} - \left(\sum_{i\in A_q} y_i\, m_i^{-1}\right)^2\right]:$$

$$Var(\hat{Y}_{SFMulti}) = \frac{N_A - n_A}{n_A(N_A-1)}\left[N_A \sum_{i\in A} y_i^2 m_i^{-2} - \left(\sum_{i\in A} y_i\, m_i^{-1}\right)^2\right] +$$

$$\frac{N_B - n_B}{n_B(N_B-1)}\left[N_B \sum_{i\in B} y_i^2 m_i^{-2} - \left(\sum_{i\in B} y_i\, m_i^{-1}\right)^2\right]$$

where

$\sum_{i\in A} y_i\, m_i^{-1}$ = 585 and $\sum_{i\in A} y_i^{-2}\, m_i^{-2}$ = 1792.5

$\sum_{i\in B} y_i\, m_i^{-1}$ = 445 and $\sum_{i\in B} y_i^{-2}\, m_i^{-2}$ = 1172,5

Variance of this estimator was calculated as follows:

$$Var(\hat{Y}_{SFMulti}) = \frac{300-60}{60(299)}\,[300(1792.5) - (585)^2]$$

$$+ \frac{250-50}{50(249)}\,[250(1172.5) - (445)^2]$$

$Var\ (\hat{Y}_{SFMulti})$ = 2615.72 + 1527.71

$Var\ (\hat{Y}_{SFMulti})$ = 4143.43

$Var\ (\hat{Y}_{SFMulti})$ > Var $(\hat{Y}_{FB})$.

Estimates of population total and variance for situation 1 and situation 2 were summarized as Table 15 and 16.

As seen in Table 15, using Hartley`s estimator can be good solution for estimation of population total when $N_A, N_B, N_{ab}$ are known. Because the variance and response bias of Hartley`s estimator smaller than the variance of Lund`s estimator.

$Var(\hat{Y}_H) \leq Var(\hat{Y}_L)$.

Table 15: The Estimates of Population Total and Its Variance for Dual Frame in Situation 1

HARTLEY`S ESTIMATOR		
$\hat{Y}_{HT}$ = 1050	$Var(\hat{Y}_{HT})$ = 4329.93	Response Bias (RB) = -20
LUND`S ESTIMATOR		
$\hat{Y}_{L}$ = 1055.70	$Var(\hat{Y}_{L})$ = 4330.28	Response Bias (RB) = -25.7

As seen from the Table 16, Fuller-Burmeister`s estimator has minimum variance and response bias between others.

Table 16: The Estimates of Population Total and Its Variance for Dual Frame in Situation 2

HARTLEY`S ESTIMATOR		
$\hat{Y}_{HT}$ = 1015	$Var(\hat{Y}_{HT})$ = 5160.60	Response Bias (RB) = 15
LUND`S ESTIMATOR		
$\hat{Y}_{L}$ = 1022.30	$Var(\hat{Y}_{L})$ = 5138.5	Response Bias (RB) = 7.7
FULLER-BURMEISTER`S ESTIMATOR		
$\hat{Y}_{FB,s}$ = 1026.95	$Var(\hat{Y}_{FB,s})$ = 4138.32	Response Bias (RB) = 3.05
BANKIER`S SINGLE- FRAME ESTIMATOR		
$\hat{Y}_{SF}$ = 1015	*	Response Bias (RB) = 15
SKINNER`S RAKING RATIO ESTIMATOR		
$\hat{Y}_{RR}{}^{\infty}$ = 1026.95	$Var(\hat{Y}_{RR}{}^{\infty})$ = 5169.22	Response Bias (RB) = 3.05
MECATTI`S SINGLE FRAME MULTIPLICITY ESTIMATOR		
$\hat{Y}_{SFMulti}$ = 1015	$Var(\hat{Y}_{SFMulti})$ = 4143.43	Response Bias (RB) = 15

*$\hat{Y}_{SF}$ is a special case of $\hat{Y}_{H}$ or of $\hat{Y}_{L}$ with $p = f_A(f_A + f_B)^{-1}$ so Var($\hat{Y}_{SF}$) ≥ Var($\hat{Y}_{FB}$).(Skinner 1991)

Thus, when N_a, N_b and N_{ab} are unknown, using Fuller-Burmeister`s estimator may be good approach for dual frame. However, Fuller-Burmeister`s estimators become more complex as the number of frames is increased. Fulvia Mecatti`s estimator can be good alternative when the number of frames is increased and this estimator does not require domain classification. Therefore they are insensitive to misclassification.

6.2 An Example For Three Frames

In this study, three artificial frames A, B and C were used. (See Table A.1, A.2 and A.3 in Appendices A). As it is said in section 6.1, in order to calculate estimation of population total for situation 1, frame A, frame B and frame C have been matched with each other according to their ID Label to classified population elements to seven domains. Results of matching, population elements classification to seven domains a, b, c, ab, ac, bc and abc can be seen in Table C.1, C.2, C.3, C.4, C.5, C.6 and C.7 in Appendix C. The sizes, totals of frame A, frame B, and frame C and domains a, b, c, ab, ac, bc and abc were given as below:

N_A = 300, N_B = 250, N_C = 210
Y_A = 790, Y_B = 650 Y_C = 580

N_a = 110, N_b = 60, N_c = 30,
Y_a = 300, Y_b = 170, Y_c = 130,

N_{ab} = 50, N_{ac} = 40, N_{bc} = 40, N_{abc} = 100
Y_{ab} = 110, Y_{ac} = 80, Y_{bc} = 70, Y_{abc} = 300.

The size and total of population was calculated as follows:

$$N = N_a + N_b + N_c + N_{ab} + N_{ac} + N_{bc} + N_{abc}$$

$$N = 110 + 60 + 30 + 50 + 40 + 40 + 100 = 430$$

$$Y = Y_a + Y_b + Y_c + Y_{ab} + Y_{ac} + Y_{bc} + Y_{abc}$$

$$Y = 300 + 170 + 130 + 110 + 80 + 70 + 300$$

$$\boldsymbol{Y = 1160}$$

In this section Cochran`s estimators were used for estimation of domain sizes and Fuller and Burmeister`s estimators for dual frames were enlarged to three frames and domain sizes were estimated by using sample data. For this purpose, each sample was classified in four domains by using algorithms in Figure 9, 10 and 11. Results of classification sample A elements to domains a, ab, ac and abc were shown in Table F.1, F.2, F.3 and F.4 in Appendix F. Results of classification sample B elements to domains a, ab, bc and abc were shown in Table F.5, F.6, F.7 and F.8 in Appendix F and results of classification sample C elements to domains a, ab, ac and abc were shown in Table F.9, F.10, F.11 and F.12 in Appendix F. n'_{ab} was the number of elements sampled from frame A and in the domain ab. $\bar{y}'_{ab}$ was the mean of these elements. n'_{ac} was the number of elements sampled from frame A and in the domain ac. $\bar{y}'_{ac}$ was the mean of these elements. n'_{abc} was the number of elements sampled from frame A and in the domain abc. $\bar{y}'_{abc}$ was the mean of these elements. n''_{ab} was the number of elements sampled from frame B and in the domain ab. $\bar{y}''_{ab}$ was the mean of these elements. n'_{bc} was the number of elements sampled from frame B and in the domain bc. $\bar{y}'_{bc}$ was the mean of these elements. n''_{abc} was the number of elements sampled from frame B and in the domain abc. $\bar{y}''_{abc}$ was the mean of these elements. Similarly, n''_{ac} was the number of elements sampled from frame C and in the domain ac. $\bar{y}'_{ac}$ was the mean of these elements. n''_{bc} was the number of elements sampled from frame C and in the domain bc. $\bar{y}'_{bc}$ was the mean of these elements and n'''_{abc} was the number of elements sampled from frame C and in the domain abc. $\bar{y}'_{abc}$ was the mean of these elements. The sizes, totals and means of independent simple random samples s_A , s_B, s_C and sample domains s_a, s_b, s_c, s_{ab}, s_{bc}, s_{ac} and s_{abc} were found as below:

n_A = 60, n_B = 50, n_C = 30,
y_A = 180, y_B = 115, y_C = 87,

n_a = 12 , n'_{ab} = 24, n'_{ac} = 12, n'_{abc} = 12,
y_a = 18, y'_{ab} = 54, y'_{ac} = 48, y'_{abc} = 60,
$\bar{y}_a$ = 1.5, $\bar{y}'_{ab}$ = 2.25, $\bar{y}'_{ac}$ = 4, $\bar{y}'_{abc}$ = 5,

$n_b = 10$, $n_{ab}^{''} = 10$, $n_{bc}^{'} = 10$, $n_{abc}^{''} = 20$

$y_b = 35$, $y_{ab}^{''} = 15$, $y_{bc}^{'} = 10$, $y_{abc}^{''} = 55$,

$\bar{y}_b = 3.5$, $\bar{y}_{ab}^{''} = 1.5$, $\bar{y}_{bc}^{'} = 1$, $\bar{y}_{abc}^{''} = 2.75$,

$n_c = 6$, $n_{ac}^{''} = 6$, $n_{bc}^{''} = 9$, $n_{abc}^{'''} = 9$

$y_c = 24$ $y_{ac}^{''} = 15$, $y_{bc}^{''} = 21$, $y_{abc}^{'''} = 27$

$\bar{y}_c = 4$, $\bar{y}_{ac}^{''} = 2.5$, $\bar{y}_{bc}^{'} = 2.33$ $\bar{y}_{abc}^{'''} = 3$

Sample means of units in the areas of overlaps were calculated as follows:

$$\bar{y}_{ab} = \frac{y_{ab}^{'} + y_{ab}^{''}}{n_{ab}^{'} + n_{ab}^{''}} = \frac{54+15}{24+10} = 2.03,$$

$$\bar{y}_{bc} = \frac{y_{bc}^{'} + y_{bc}^{''}}{n_{bc}^{'} + n_{bc}^{''}} = \frac{10+21}{10+9} = 1.63,$$

$$\bar{y}_{ac} = \frac{y_{ac}^{'} + y_{ac}^{''}}{n_{ac}^{'} + n_{ac}^{''}} = \frac{48+15}{12+6} = 3.5,$$

$$\bar{y}_{abc} = \frac{y_{abc}^{'} + y_{abc}^{''} + y_{abc}^{'''}}{n_{abc}^{'} + n_{abc}^{''} + n_{abc}^{'''}} = \frac{60+55+27}{12+20+9} = 3.47$$

$$g_A = \frac{N_A - n_A}{N_A - 1} = \frac{300-60}{300-1} = 0.802,$$

$$g_B = \frac{N_B - n_B}{N_B - 1} = \frac{250-50}{250-1} = 0.803$$

$$g_C = \frac{N_C - n_C}{N_C - 1} = \frac{210-30}{210-1} = 0.861$$

The estimation of population size and total by using Cochran`s estimator were calculated as below:

Estimation of $N_{ab,}$ and its variance was as calculated below:

$$v\,(\hat{N}_{ab}^{`}) = \frac{300^2}{60}\,\frac{24}{60}\left(1-\frac{24}{60}\right) = 360 \quad , \quad v\,(\hat{N}_{ab}^{``}) = \frac{250^2}{50}\,\frac{10}{50}\left(1-\frac{10}{50}\right) = 200$$

$$p_{ab} = \frac{v\,(\hat{N}_{ab}^{``})}{v\,(\hat{N}_{ab}^{`}) + v\,(\hat{N}_{ab}^{``})} = \frac{200}{360+200} = 0.36$$

$$\hat{N}_{ab,COC} = p\hat{N}_{ab}^{`} + q\hat{N}_{ab}^{``} \qquad \text{where} \quad p + q = 1$$

$$\hat{N}_{ab}^{`} = \frac{N_A}{n_A}\,n_{ab}^{`} = \frac{300}{60}\,24 = 120 \quad , \quad \hat{N}_{ab}^{``} = \frac{N_B}{n_B}\,n_{ab}^{``} = \frac{250}{50}\,10 = 50$$

$$\hat{N}_{ab,COC} = 0.36(120) + 0.64(50) = 75.2 \cong 75$$

$$\boldsymbol{\hat{N}_{ab,COC} = 75}$$

$$v\,(\hat{N}_{ab,COC}) = {p_{ab}}^2\,v\,(\hat{N}_{ab}^{`}) + {q_{ab}}^2\,v\,(\hat{N}_{ab}^{``})$$

$$v\,(\hat{N}_{ab,COC}) = (0.36)^2\,360 + (0.64)^2.200$$

$$v\,(\hat{N}_{ab,COC}) = 128.57$$

Estimation of $N_{ac,}$ and its variance was as calculated below:

$$v\,(\hat{N}_{ac}^{`}) = \frac{300^2}{60}\,\frac{12}{60}\left(1-\frac{12}{60}\right) = 240 \quad , \quad v\,(\hat{N}_{ac}^{``}) = \frac{210^2}{30}\,\frac{6}{30}\left(1-\frac{6}{30}\right) = 235.2$$

$$p_{ac} = \frac{v\,(\hat{N}_{ac}^{``})}{v\,(\hat{N}_{ac}^{`}) + v\,(\hat{N}_{ac}^{``})} = \frac{235.2}{240+235.2} = 0.49$$

$$\hat{N}_{ac}^{`} = \frac{N_A}{n_A}\,n_{ac}^{`} = \frac{300}{60}\,12 = 60 \; , \quad \hat{N}_{ac}^{``} = \frac{N_C}{n_C}\,n_{ac}^{``} = \frac{210}{30}\,6 = 42$$

$$\hat{N}_{ac,COC} = (0.49)60 + (0.51)42 = 50.82 \cong 51$$

$$\boldsymbol{\hat{N}_{ac,COC} = 51}$$

$$v\,(\hat{N}_{ac,COC}) = p_{ac}{}^2\; v\,(\hat{N}^{`}_{ac}) + q_{ac}{}^2\; v\,(\hat{N}^{``}_{ac})$$

$$= (0.49)^2 240 + (0.51)^2 235.2$$

$$= 118.79$$

Estimation of N_{bc} and its variance was as calculated below:

$$v\,(\hat{N}^{`}_{bc}) = \frac{250^2}{50}\,\frac{10}{50}\left(1-\frac{10}{50}\right) = 200 \quad , \quad v\,(\hat{N}^{``}_{bc}) = \frac{210^2}{30}\,\frac{9}{30}\left(1-\frac{9}{30}\right) = 308.7$$

$$p_{bc} = \frac{v\,(\hat{N}^{``}_{bc})}{v\,(\hat{N}^{`}_{bc}) + v\,(\hat{N}^{``}_{bc})} = \frac{308.7}{200+308.7} = 0.61$$

$$\hat{N}^{`}_{bc} = \frac{N_B}{n_B}\, n^{`}_{bc} = \frac{250}{50}\, 10 = 50 \quad , \quad \hat{N}^{``}_{bc} = \frac{N_C}{n_C}\, n^{``}_{bc} = \frac{210}{30}\, 9 = 63$$

$$\hat{N}_{bc,COC} = (0.61)50 + (0.39)63 = 55\ 07 \cong 55$$

$$\boldsymbol{\hat{N}_{bc,COC} = 55}$$

$$v\,(\hat{N}_{bc,COC}) = p_{bc}{}^2\; v\,(\hat{N}^{`}_{bc}) + q_{bc}{}^2\; v\,(\hat{N}^{``}_{bc})$$

$$= (0.61)^2 200 + (0.39)^2 308.7$$

$$= 121.37$$

Estimation of N_{abc} and its variance was as calculated below:

$$v\,(\hat{N}^{`}_{abc}) = \frac{300^2}{60}\,\frac{12}{60}\left(1-\frac{12}{60}\right) = 240$$

$$v\,(\hat{N}^{``}_{abc}) = \frac{250^2}{50}\,\frac{20}{50}\left(1-\frac{20}{50}\right) = 300$$

$$v\,(\hat{N}^{```}_{abc}) = \frac{210^2}{30}\,\frac{9}{30}\left(1-\frac{9}{30}\right) = 308.7$$

$$p_A = \frac{\frac{1}{V(\hat{N}'_{abc})}}{\frac{1}{V(\hat{N}'_{abc})} + \frac{1}{V(\hat{N}''_{abc})} + \frac{1}{V(\hat{N}'''_{abc})}} = \frac{0.0042}{0.0042+0.0033+0.0032} = 0.40$$

$$p_B = \frac{\frac{1}{V(\hat{N}''_{abc})}}{\frac{1}{V\left(\hat{N}'_{abc}\right)} + \frac{1}{V(\hat{N}''_{abc})} + \frac{1}{V(\hat{N}'''_{abc})}} = \frac{0.0033}{0.0042+0.0033+0.0032} = 0.30$$

$$p_C = \frac{\frac{1}{V(\hat{N}'''_{abc})}}{\frac{1}{V(\hat{N}'_{abc})} + \frac{1}{V(\hat{N}''_{abc})} + \frac{1}{V(\hat{N}'''_{abc})}} = \frac{0.0032}{0.0042+0.0033+0.0032} = 0.30$$

$$\hat{N}_{abc,COC} = p_A \frac{N_A}{n_A} n'_{abc} + p_B \frac{N_B}{n_B} n''_{abc} + p_C \frac{N_C}{n_C} n'''_{abc}$$

$$\hat{N}_{abc,COC} = (0.40) \frac{300}{60} 12 + (0.30) \frac{250}{50} 20 + (0.30) \frac{210}{30} 9 = 72.9$$

$$\boldsymbol{\hat{N}_{abc,COC} = 72.9 \cong 73}$$

$$Var\left(\hat{N}_{abc,COC}\right) = p_A^2\, Var\left(\hat{N}'_{abc}\right) + p_B^2\, Var\left(N''_{abc}\right) + p_C^2\, Var(N'''_{abc}$$

$$= (0.40)^2 240 + (0.30)^2 300 + (0.30)^2 308.7 = 93.18$$

Estimation of N_a was obtained by subtraction of the estimates of $\hat{N}_{ab}$, $\hat{N}_{ac}$ and $\hat{N}_{abc}$ from the known frame size N_A,

$$\hat{N}_a = N_A - (\hat{N}_{ab,COC} + \hat{N}_{ac,COC} + \hat{N}_{abc,COC}) = 300 - (75 + 51 + 73)$$

$$= 300 - 199 = 101$$

$$\boldsymbol{\hat{N}_{a,COC} = 101}$$

Estimation of N_b was obtained by subtraction of the estimates of $\hat{N}_{ab}$, $\hat{N}_{bc}$ and $\hat{N}_{abc}$ from the known frame size N_B,

$$\hat{N}_{b,COC} = N_B - (\hat{N}_{ab,COC} + \hat{N}_{bc,COC} + \hat{N}_{abc,COC}) = 250 - (75 + 55 + 73)$$

$$= 250 - 203 = 47$$

$\boldsymbol{\widehat{N}_{b,COC}} = \mathbf{47}$

Estimation of N_c was obtained by subtraction of the estimates of $\widehat{N}_{ac}$, $\widehat{N}_{bc}$ and $\widehat{N}_{abc}$ from the known frame size N_C,

$$\widehat{N}_{c,COC} = N_C - (\widehat{N}_{ac,COC} + \widehat{N}_{bc,COC} + \widehat{N}_{abc,COC}) = 210 - (51 + 55 + 73)$$

$$= 210 - 179 = 31$$

$\boldsymbol{\widehat{N}_{c,COC}} = \mathbf{31}$

Estimation of size and total of population by using Cochran estimators were as calculated below:

$$\widehat{N}_{COC} = \widehat{N}_{a,COC} + \widehat{N}_{b,COC} + \widehat{N}_{c,COC} + \widehat{N}_{ab,COC} + \widehat{N}_{ac,COC} + \widehat{N}_{bc,COC} + \widehat{N}_{abc,COC}$$

$$= 101 + 47 + 31 + 72 + 55 + 51 + 73$$

$\boldsymbol{\widehat{N}_{COC}} = \mathbf{433}$

$$\hat{Y}_{COC} = \widehat{N}_{a,COC}\bar{y}_a + \widehat{N}_{b,COC}\bar{y}_b + \widehat{N}_{c,COC}\bar{y}_c + \widehat{N}_{ab,COC}\bar{y}_{ab} + \widehat{N}_{ac,COC}\bar{y}_{ac} + \widehat{N}_{bc,COC} + \widehat{N}_{abc,COC}\bar{y}_{abc}$$

$$\hat{Y}_{COC} = 101(1.5) + 47(3.5) + 31(3) + 75(2.03) + 51(3.5) + 55(1.63) + 73(3.47)$$

$\boldsymbol{\hat{Y}_{COC}} = \mathbf{1082.71}$

Response Bias (RB) = $Y - \hat{Y}_{COC}$ = 1160 – 1082.71

RB = $\mathbf{77.29}$

In this section, Fuller-Burmeister`s estimators for dual frame given section 4.2.3 has been enlarged to three frames. These estimators were applied to obtain estimation of size and total of population. Thus, the results were obtained as follows:

Estimation of N_{ab} was the smallest root of quadratic equation

$$[n_A g_B + n_B g_A]\,\widehat{N}^2_{ab,FB} - [n_A N_B g_B + n_B N_A g_A + n^{`}_{ab} N_A g_B + n^{``}_{ab} N_B g_A]\widehat{N}_{ab,FB}$$

$$+ [n^{`}_{ab} g_B + n^{``}_{ab} g_A]\, N_A N_B = 0$$

$$[60(0.803) + 50(0.802)]\widehat{N}^2_{ab,FB}$$
$$- [60(250)0.803 + 50(300)0.802 + 24(300)0.803$$
$$+ 10(250)0.802]\widehat{N}_{ab,FB} + [24(0.803) + 10(0.802)]300\,250 = 0$$

$$88.28\widehat{N}^2_{ab,FB} - 31861.6\widehat{N}_{ab,FB} + 2046900$$

$$\widehat{N}_{ab,FB}{}^1 = \frac{-b-\sqrt{b^2-4ac}}{2a} = \frac{31861.1-\sqrt{31861.1^2-4(88.28)2046900}}{2(88.28)}$$

$$= \frac{14763.05}{176.56} = 83.61 \cong 84$$

$$\widehat{N}_{ab,FB}{}^2 = \frac{-b+\sqrt{b^2-4ac}}{2a} = \frac{31861.1+\sqrt{31861.1^2-4(88.28)2046900}}{2(88.28)}$$

$$= \frac{48960.14}{176.56} = 277.30$$

$\boldsymbol{\widehat{N}_{ab,FB} = 84}$

The estimation of N_{ac} was the smallest root of quadratic equation,

$$[n_A g_C + n_C g_A]\,\widehat{N}^2_{ac,FB} - [n_A N_C g_C + n_C N_A g_A + n^{`}_{ac} N_A g_C + n^{``}_{ac} N_C g_A]\widehat{N}_{ac,FB}$$

$$+ [n^{`}_{ac} g_C + n^{``}_{ac} g_A]\, N_A N_C = 0$$

$$[60(0.861) + 30(0.802)]\widehat{N}^2_{ac,FB}$$
$$- [60(210)0.861 + 30(300)0.802 + 12(300)0.861$$
$$+ 6(210)0.802]\widehat{N}_{ac,FB} + [12(0.861) + 6(0.802)]300\,210 = 0$$

$$75.72\widehat{N}^2_{ac,FB} - 22176.72\widehat{N}_{ac,FB} + 954072 = 0$$

$$\hat{N}_{ac,FB}{}^{1} = \frac{-b-\sqrt{b^2-4ac}}{2a} = \frac{22176.72-\sqrt{22176.72^2-4(75.72)954072}}{2(75.72)}$$

$$= \frac{7934.61}{151.44} = 52{,}39$$

$$\hat{N}_{ac,FB}{}^{2} = \frac{-b-\sqrt{b^2-4ac}}{2a} = \frac{22176.72+\sqrt{22176.72^2-4(75.72)954072}}{2(75.72)}$$

$$= \frac{36418.82}{151.44} = 240.48$$

Thus

$$\boldsymbol{\hat{N}_{ac,FB} = 52}$$

The estimation of N_{bc} was the smallest root of quadratic equation,

$$[n_B g_C + n_C g_B]\,\hat{N}^2_{bc,FB} - [n_B N_C g_C + n_C N_B g_B + n^{`}_{bc} N_B g_C + n^{``}_{bc} N_C g_B]\hat{N}_{bc,FB}$$

$$+ [n^{`}_{bc} g_C + n^{``}_{bc} g_B]\, N_B N_C = 0$$

$$[50(0.861) + 30(0.803)]\hat{N}^2_{bc,FB}$$
$$- [50(210)0.861 + 30(250)0.803 + 10(250)0.861$$
$$+ 9(210)0.803]\hat{N}_{bc,FB} + [10(0.861) + 9(0.803)]250\ 210 = 0$$

$$67.14\hat{N}^2_{bc,FB} - 18733.17\hat{N}_{bc,FB} + 831442.5 = 0$$

$$\hat{N}_{bc,FB}{}^{1} = \frac{-b-\sqrt{b^2-4ac}}{2a} = \frac{18733.17-\sqrt{18733.17^2-4(67.14)831442.5}}{2(67.14)}$$

$$= \frac{7435.40}{134.28} = 55.37$$

$$\boldsymbol{\hat{N}_{bc,FB} = 55.37}$$

Estimation of N_{abc}

$$[n_A g_B g_C + n_B g_A g_C + n_C g_A g_B]\boldsymbol{N^3_{abc,FB}} - [g_B g_C(n_A N_B + n_A N_C + N_A n^{`}_{abc}) + g_A g_C(n_B N_C + n_B N_A + N_B n^{``}_{abc}) + g_A g_B(n_C N_A + n_C N_B + N_C n^{```}_{abc})]\boldsymbol{N^2_{abc,FB}} + [g_B g_C(n_A N_B N_C + N_B N_A n^{`}_{abc} + N_C N_A n^{`}_{abc}) + g_A g_C(n_B N_A N_C + N_A N_B n^{``}_{abc} + N_C N_B n^{``}_{abc}) + g_A g_B(n_C N_B N_A + N_C N_B n^{```}_{abc} + N_C N_A n^{```}_{abc})]\boldsymbol{N_{abc,FB}} - N_A N_B N_C[g_B g_C n^{`}_{abc} + g_A g_C n^{``}_{abc} + g_A g_B n^{```}_{abc}] = 0$$

$$[60(0.803)0.861 + 50(0.802)0.861 + 30(0.802)0.803]\boldsymbol{N^3_{abc,FB}} - [0.803(0.861)(60\ 250 + 60\ 210 + 300\ 12) + 0.802(0.861)(50\ 210 + 50\ 300 + 250\ 20) + 0.802(0.803)(30\ 300 + 30\ 250 + 210\ 9)]\boldsymbol{N^2_{abc,FB}} + [0.803(0.861)(60\ 250\ 210 + 250\ 300\ 12 + 210\ 300\ 12) + 0.802(0.861)(50\ 300\ 210 + 300\ 250\ 20 + 210\ 250\ 20) + 0.802(0.803)(30\ 250\ 300 + 210\ 250\ 9 + 210\ 300\ 9)]\boldsymbol{N_{abc,FB}} - [0.803(0.861)12 + 0.802(0.861)20 + 0.802(0.803)9]300\ 250\ 210 = 0$$

$$95.33N^3_{abc,FB} - 54475.34N^2_{abc,FB} + 9377219.84N_{abc,FB} - 439473667.5 = 0$$

$$\boldsymbol{\hat{N}_{abc,FB} = 75.86 = 76}$$

The estimation of N_a, N_b and N_c can be found by

$$\hat{N}_{a,FB} = N_A - (\hat{N}_{ab,FB} + \hat{N}_{ac,FB} + \hat{N}_{abc,FB}) = 300 - (84 + 52 + 76)$$

$$\boldsymbol{\hat{N}_{a,FB} = 88}$$

$$\hat{N}_{b,FB} = N_B - (\hat{N}_{ab,FB} + \hat{N}_{bc,FB} + \hat{N}_{abc,FB}) = 250 - (84 + 55 + 76)$$

$$\boldsymbol{\hat{N}_{b,FB} = 35}$$

$$\hat{N}_{c,FB} = N_C - (\hat{N}_{bc,FB} + \hat{N}_{ac,FB} + \hat{N}_{abc,FB}) = 210 - (55 + 52 + 76)$$

$$\boldsymbol{\hat{N}_{c,FB} = 27}$$

Estimation of N was found easily using these estimators as below:

$$\widehat{N}_{FB} = \widehat{N}_{a,FB} + \widehat{N}_{b,FB} + \widehat{N}_{c,FB} + \widehat{N}_{ab,FB} + \widehat{N}_{ac,FB} + \widehat{N}_{bc,FB} + \widehat{N}_{abc,FB}$$

$\widehat{N}_{FB}$ = 88 + 35 + 27 + 84 + 52 + 55 + 76 = 417

$\boldsymbol{\widehat{N}_{FB} = 417}$

Estimation of Y was found using these estimators as below:

$$\widehat{Y}_{FB} = \widehat{N}_{a,FB}\bar{y}_a + \widehat{N}_{b,FB}\bar{y}_b + \widehat{N}_{c,FB}\bar{y}_c + \widehat{N}_{ab,FB}\bar{y}_{ab} + \widehat{N}_{ac,FB}\bar{y}_{ac} + \widehat{N}_{bc,FB}\bar{y}_{bc} + \widehat{N}_{abc,FB}\bar{y}_{abc}$$

$\widehat{Y}_{FB}$ = 88(1.5) + 35(3.5) + 27(3) + 84(2.03) + 52(3.5) + 55(1.63) + 76(3.47)

$\boldsymbol{\widehat{Y}_{FB} = 1041.39}$

Response Bias (RB) = $Y - \widehat{Y}_{FB}$ = 1160 – 1041.39

RB= $\mathbf{118.61}$

The multiplicity estimator was also used for estimation of population total. The units multiplicity were collected by asking sampled elements “how many frames they belong to". The multiplicity units of sample A, B and C were shown in Table H.1, H.2, H.3 and H.4 in Appendix H respectively.

By using the single frame multiplicity estimator, total value of population were obtained as follows:

$$\widehat{Y}_{SFMulti} = f_A^{-1}\left(\sum_{i \in S_A} y_i m_i^{-1}\right) + f_B^{-1}\left(\sum_{i \in S_B} y_i m_i^{-1}\right) + f_C^{-1}\left(\sum_{i \in S_C} y_i m_i^{-1}\right)$$

$\widehat{Y}_{SFMulti}$ = 5(88.98) + 5(65.75) + 7(50.97)

$\boldsymbol{\widehat{Y}_{SFMulti} = 1130.44}$

Response Bias (RB) = 1160 – 1130.44

RB = $\mathbf{29.56}$

As seen in Table 17, Mecatti`s the single frame multiplicity estimator gives smallest response bias for estimation of population total.

Table 17: The Estimates of Population Total for Three Frames

COCHRAN`S ESTIMATOR	
$\hat{Y}_{COC}$ = 1082.71	Response Bias (RB) = 77.29
FULLER AND BURMEISTER`S ESTIMATOR	
$\hat{Y}_{FB}$ = 1041.39	Response Bias (RB) = 118.61
MECATTI`S SINGLE FRAME MULTIPLICITY ESTIMATOR	
$\hat{Y}_{SFMulti}$ = 1130.44	Response Bias (RB) = 29.56

When the number of frames is increased, using Mecatti`s single frame multiplicity estimator is more reasonable than using Cochran`s and Fuller and Burmeister`s estimators.

CHAPTER 7

CONCLUSION

Multiple Frame Surveys can be adopted only if the different frames contribute with essential information or when a complete frame is very expensive to sample. In a typical application, such as farm survey, there is a list frame (incomplete, names, addresses, less costly) and an area frame (complete, insensitive to changes, expensive to sample). Hartley (1962 and 1974) was concentrate on this circumstance. By taking a large sample from list frame and a smaller sample from area frame, can be obtained a more efficient estimator for a fixed cost than the area frame was used alone. If the list frame was used alone, the results would be seriously biased due to its incompleteness.

When the articles in literature were examined, it was seen that different estimators were proposed by the author under the several assumptions. These assumptions were that every unit in the population of interest should belong to at least one of the frames and it should be possible to record for each sampled unit whether or not it belongs to one or more of the other frames. All of the methods used for estimation with overlapping frames were examined in this thesis. Various author, such as Hartley (1962 and 1974), Lund (1968), Fuller and Burmeister (1972), proposed optimal estimators of the population total when domain sizes were not known. (Skinner 1991 and Lohr and Rao 2000) showed that optimal estimators had optimal properties but practical problems. These estimators are sensitive to misclassification of observation into domains and estimates can be biased due to misclassification. In addition, optimum weights depend on unknown population parameter so they should be estimated from sample data and these estimators become more complex as the number of frames is increased.

Single frame (SF) approach has been proposed in order to reduce this kind of complexity by using fixed weights. Fixed weights usually differ from optimum estimator and they are also sensitive to misclassification. Hence, optimum and SF estimators can be applied only if the correct classification of sample units into the domains is achieved.

A single frame multiplicity approach was proposed by Mecatti (2005 and 2007). The multiplicity estimator, unlike optimum and SF estimators, does not require domain classification. So they are insensitive to misclassification. Multiplicity estimator is recommended when the risk of misclassification of sampled unit into the domains is possible.

In this thesis, different estimators of the total were applied to the artificial two and three frames. Fuller and Burmeister`s estimator was the most efficient, however due to additional calculations and complexity SF multiplicity estimator may be preferred.

Multiple frame surveys can improve efficiency and reduce undercoverage bias, but should be used carefully. The number of used frames should not be high, otherwise the sample size per domain will be small, the domain sizes will probably be only approximately known and the population total estimator can be seriously biased. Moreover, with many frames, some of which out of date, record matching is very difficult and errors in record matching are another sources of bias.

REFERENCES

Ayhan, H.Ö. (2000) Estimators of Vital Events in Dual Record Systems, *Journal of Applied Statistics,* 27(2), 157-169

Ayhan, H.Ö. and Ekni, S. (2003), Coverage Error in Population Censuses: The Case of Turkey, *Survey Methodology,* 29(2), 155-165

Bankier, M.D. (1986) Estimators Based on Several Stratified Samples with Applications to Multiple Frame Surveys, *Journal of the American Statistical Association,* 81, 1074-1079.

Cochran, Robert S. (1967) The estimation of Domain Sizes When Sampling Frames Are Interlocking, *Mimeographed paper at the American Statistical Association meetings,* Social Science Section, Washington, D.C.

Fuller, W.A., and Burmeister, L.F. (1972) Estimators for Sample Selected From Two Overlapping Frames, *Proceedings of the Social Statistics Section,* American Statistical Association, pp. 245-249.

Hartley, H.O. (1962) Multiple Frame Surveys, *Proceedings of the Social Statistics Section,* American Statistical Association, pp. 203-206.

Hartley, H.O. (1974) Multiple Frame Methodology and Selected Applications, *Sankhya,* Series C, 36, 99-118.

Kalton, G., and Anderson, D.W. (1986) Sampling Rare Population, *Journal of the Royal Statistical Society,* Series A, 149, 65-82.

Kish, L. (1965) *Survey Sampling*, New York: John Wiley and Sons.

Lund, R.E. (1968) Estimators in Multiple Frame Surveys, *Proceedings of the Social Statistics Section,* American Statistical Association, pp. 282-288

Lohr, S.L., and Rao, J.N.K. (2000) Inference in Dual Frame Surveys, *Journal of the American Statistical Association,* 95, 271-280.

Lohr, S.L., and Rao, J.N.K. (2006) Estimation in Multiple Frame Surveys, *Journal of the American Statistical Association,* 101, 1019-1030.

Maia, M., and Vincente, P. (2009) Indirect Sampling in the Contenxt of Dual Frame Surveys, *Faculdade de Economia e Gesto,* Universidade Catolica Portuquesa (Porto)

Marks, E.S. Seltzer, W. & Krotki, K.J. (1974) *Population Growth Estimation: A Handbook of Vital Statistics Measurement,* New York, The Population Council.

Mecatti, F. (2005) Single Frame Estimation in Multiple Frame Survey, *Proceedings of the Statistics Canada Symposium.*

Mecatti, F. (2007) A Single Frame Multiplicity Estimator for Multiple Frame Surveys, *Survey Methodology,* Vol.33, No.2, pp 151-157.

Rao, J.N.K., and Skinner, C.J.(1996) Estimation in Dual Frame Surveys With Complex Designs, *Proceedings of the Survey Methods Section,* Statistical Society of Canada, pp 63-68.

Skinner, C.J. (1991) On the efficiency of Raking Ratio Estimation for Multiple Frame Surveys, *Journal of the American Statistical Association,* 86, 779-784.

Skinner , C.J., and Rao, J.N.K. (1996) Estimation in Dual Frame Surveys with Complex Design, *Journal of the American Statistical Association,* 91, 349-356.

APPENDIX A

TABLES OF POPULATION ELEMENTS FOR APPLICATION

Table A.1: List of Population Elements for Frame A

FRAME A								
Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i
1	Aylin	1	36	Hakan-01	2	71	Aysegul-02	0
2	Suheyla	1	37	Ebru-01	2	72	Deniz-02	5
3	Arif	2	38	Ozcan-01	3	73	Nilgun-02	2
4	Aynur	3	39	Birsen-01	4	74	Zehra-02	3
5	Ayse	5	40	Kamil-01	2	75	Sule-02	4
6	Hakan	2	41	Aysegul-01	0	76	Ahmet-02	2
7	Ebru	2	42	Deniz-01	5	77	Bora-02	1
8	Ozcan	3	43	Nilgun-01	2	78	Hacer-02	6
9	Birsen	4	44	Zehra-01	3	79	Nuretin-02	2
10	Kamil	2	45	Sule-01	4	80	Asli-02	1
11	Aysegul	0	46	Ahmet-01	2	81	Vural-02	1
12	Deniz	5	47	Bora-01	1	82	Haydar-02	6
13	Nilgun	2	48	Hacer-01	6	83	Tuna-02	7
14	Zehra	3	49	Nurettin-01	2	84	Ozlem-02	4
15	Sule	4	50	Asli-01	1	85	Berrin-02	1
16	Ahmet	2	51	Vural-01	1	86	Umit-02	5
17	Bora	1	52	Haydar-01	6	87	Ilsu-02	0
18	Hacer	6	53	Tuna-01	7	88	Aysenaz-02	0
19	Nurettin	2	54	Ozlem-01	4	89	Osman-02	3
20	Asli	1	55	Berrin-01	1	90	Mehmet-02	1
21	Vural	1	56	Umit-01	5	91	Aylin-03	1
22	Haydar	6	57	Ilsu-01	0	92	Suheyla-03	1
23	Tuna	7	58	Aysenaz01	0	93	Arif-03	2
24	Ozlem	4	59	Osman-01	3	94	Aynur-03	3
25	Berrin	1	60	Mehmet-01	1	95	Ayse-03	5
26	Umit	5	61	Aylin-02	1	96	Hakan-03	2
27	Ilsu	0	62	Suheyla02	1	97	Ebru-03	2
28	Aysenaz	0	63	Arif-02	2	98	Ozcan-03	3
29	Osman	3	64	Aynur-02	3	99	Birsen-03	4
30	Mehmet	1	65	Ayse-02	5	100	Kamil-03	2
31	Aylin-01	1	66	Hakan-02	2	101	Aysegul-03	0
32	Suheyla-01	1	67	Ebru-02	2	102	Deniz-03	5
33	Arif-01	2	68	Ozcan-02	3	103	Nilgun-03	2
34	Aynur-01	3	69	Birsen-02	4	104	Zehra-03	3
35	Ayse-01	5	70	Kamil-02	2	105	Sule-03	4

Table A.1 (continued)

FRAME A								
Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i
106	Ahmet-03	2	160	Kamil-05	2	214	Aynur-07	3
107	Bora-03	1	161	Aysegul-05	0	215	Ayse-07	5
108	Hacer-03	6	162	Deniz-05	5	216	Hakan-07	2
109	Nurettin-03	2	163	Nilgun-05	2	217	Ebru-07	2
110	Asli-03	1	164	Zehra-05	3	218	Ozcan-07	3
111	Vural-03	1	165	Sule-05	4	219	Birsen-07	4
112	Haydar-03	6	166	Ahmet-05	2	220	Kamil-07	2
113	Tuna-03	7	167	Bora-05	1	221	Aysegul-07	0
114	Ozlem-03	4	168	Hacer-05	6	222	Deniz-07	5
115	Berrin-03	1	169	Nurettin-05	2	223	Nilgun-07	2
116	Umit-03	5	170	Asli-05	1	224	Zehra-07	3
117	Ilsu-03	0	171	Vural-05	1	225	Sule-07	4
118	Aysen03az	0	172	Haydar-05	6	226	Ahmet-07	2
119	Osman-03	3	173	Tuna-05	7	227	Bora-07	1
120	Mehmet-03	1	174	Ozlem-05	4	228	Hacer-07	6
121	Aylin-04	1	175	Berrin-05	1	229	Nurettin-07	2
122	Suheyla-04	1	176	Umit-05	5	230	Asli-07	1
123	Arif-04	2	177	Ilsu-05	0	231	Vural-07	1
124	Aynur-04	3	178	Aysenaz05	0	232	Haydar-07	6
125	Ayse-04	5	179	Osman-05	3	233	Tuna-07	7
126	Hakan-04	2	180	Mehmet-05	1	234	Ozlem-07	4
127	Ebru-04	2	181	Aylin-06	1	235	Berrin-07	1
128	Ozcan-04	3	182	Suheyla-06	1	236	Umit-07	5
129	Birsen-04	4	183	Arif-06	2	237	Ilsu-07	0
130	Kamil-04	2	184	Aynur-06	3	238	Aysenaz-07	0
131	Aysegul-04	0	185	Ayse-06	5	238	Osman-07	3
132	Deniz-04	5	186	Hakan-06	2	240	Mehmet-07	1
133	Nilgun-04	2	187	Ebru-06	2	241	Aylin-08	1
134	Zehra-04	3	188	Ozcan-06	3	242	Suheyla08	1
135	Sule-04	4	189	Birsen-06	4	243	Arif-08	2
136	Ahmet-04	2	190	Kamil-06	2	244	Aynur-08	3
137	Bora-04	1	191	Aysegul-06	0	245	Ayse-08	5
138	Hacer-04	6	192	Deniz-06	5	246	Hakan-08	2
139	Nurettin-04	2	193	Nilgun-06	2	247	Ebru-08	2
140	Asli-04	1	194	Zehra-06	3	248	Ozcan-08	3
141	Vural-04	1	195	Sule-06	4	249	Birsen-08	4
142	Haydar-04	6	196	Ahmet-06	2	250	Kamil-08	2
143	Tuna-04	7	197	Bora-06	1	251	Aysegul-08	0
144	Ozlem-04	4	198	Hacer-06	6	252	Deniz-08	5
145	Berrin-04	1	199	Nurettin-06	2	253	Nilgun-08	2
146	Umit-04	5	200	Asli-06	1	254	Zehra-08	3
147	Ilsu-04	0	201	Vural-06	1	255	Sule-08	4
148	Aysenaz-04	0	202	Haydar-06	6	256	Ahmet-08	2
149	Osman-04	3	203	Tuna-06	7	257	Bora-08	1
150	Mehmet-04	1	204	Ozlem-06	4	258	Hacer-08	6
151	Aylin-05	1	205	Berrin-06	1	259	Nurettin-08	2
152	Suheyla-05	1	206	Umit-06	5	260	Asli-08	1
153	Arif-05	2	207	Ilsu-06	0	261	Vural-08	1
154	Aynur-05	3	208	Aysenaz06	0	262	Haydar-08	6
155	Ayse-05	5	209	Osman-06	3	263	Tuna-08	7
156	Hakan-05	2	210	Mehmet-06	1	264	Ozlem-08	4
157	Ebru-05	2	211	Aylin-07	1	265	Berrin-08	1
158	Ozcan-05	3	212	Suheyla-07	1	266	Umit-08	5
159	Birsen-05	4	213	Arif-07	2	267	Ilsu-08	0

Table A.1 (continued)

FRAME A								
Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i
268	Aysenaz08	0	279	Birsen-09	4	290	Asli-09	1
269	Osman-08	3	280	Kamil-09	2	291	Vural-09	1
270	Mehmet-08	1	281	Aysegul09	0	292	Haydar-09	6
271	Aylin-09	1	282	Deniz-09	5	293	Tuna-09	7
272	Suheyla-09	1	283	Nilgun-09	2	294	Ozlem-09	4
273	Arif-09	2	284	Zehra-09	3	295	Berrin-09	1
274	Aynur-09	3	285	Sule-09	4	296	Umit-09	5
275	Ayse-09	5	286	Ahmet-09	2	297	Ilsu-09	0
276	Hakan-09	2	287	Bora-09	1	298	Aysenaz-09	0
277	Ebru-09	2	288	Hacer-09	6	299	Osman-09	3
278	Ozcan-09	3	289	Nurettin09	2	300	Mehmet09	1

Table A.2: List of Population Elements for Frame B

FRAME B								
Pop. ID i	**ID Label**	**Pop. Element Value y_i**	**Pop. ID i**	**ID Label**	**Pop. Element Value y_i**	**Pop. ID i**	**ID Label**	**Pop. Element Value y_i**
1	Suheyla	1	55	Mahmut-02	3	109	Aylin-04	1
2	Vural	1	56	Murat-02	0	110	Ebru-04	2
3	Aysegul	0	57	Hacer-02	6	111	Didem-04	1
4	Berrin	1	58	Ayse-02	5	112	Tuna-04	7
5	Mahmut	3	59	Aylin-02	1	113	Hakan-04	2
6	Murat	0	60	Ebru-02	2	114	Zehra-04	3
7	Hacer	6	61	Didem-02	1	115	Ozcan-04	3
8	Ayse	5	62	Tuna-02	7	116	Necla-04	2
9	Aylin	1	63	Hakan-02	2	117	Ismail-04	4
10	Ebru	2	64	Zehra-02	3	118	Eda-04	3
11	Didem	1	65	Ozcan-02	3	119	Ferit-04	2
12	Tuna	7	66	Necla-02	2	120	Bora-04	1
13	Hakan	2	67	Ismail-02	4	121	Nurettin-04	2
14	Zehra	3	68	Eda-02	3	122	Gizem-04	3
15	Ozcan	3	69	Ferit-02	2	123	Nurcan-04	2
16	Necla	2	70	Bora-02	1	124	Doga-04	4
17	Ismail	4	71	Nurettin-02	2	125	Haydar-04	6
18	Eda	3	72	Gizem-02	3	126	Suheyla-05	1
19	Ferit	2	73	Nurcan-02	2	127	Vural-05	1
20	Bora	1	74	Doga-02	4	128	Aysegul-05	0
21	Nurettin	2	75	Haydar-02	6	129	Berrin-05	1
22	Gizem	3	76	Suheyla-03	1	130	Mahmut-05	3
23	Nurcan	2	77	Vural-03	1	131	Murat-05	0
24	Doga	4	78	Aysegul-03	0	132	Hacer-05	6
25	Haydar	6	79	Berrin-03	1	133	Ayse-05	5
26	Suheyla-01	1	80	Mahmut-03	3	134	Aylin-05	1
27	Vural-01	1	81	Murat-03	0	135	Ebru-05	2
28	Aysegul-01	0	82	Hacer-03	6	136	Didem-05	1
29	Berrin-01	1	83	Ayse-03	5	137	Tuna-05	7
30	Mahmut-01	3	84	Aylin-03	1	138	Hakan-05	2
31	Murat-01	0	85	Ebru-03	2	139	Zehra-05	3
32	Hacer-01	6	86	Didem-03	1	140	Ozcan-05	3
33	Ayse-01	5	87	Tuna-03	7	141	Necla-05	2
34	Aylin-01	1	88	Hakan-03	2	142	Ismail-05	4
35	Ebru-01	2	89	Zehra-03	3	143	Eda-05	3
36	Didem-01	1	90	Ozcan-03	3	144	Ferit-05	2
37	Tuna-01	7	91	Necla-03	2	145	Bora-05	1
38	Hakan-01	2	92	Ismail-03	4	146	Nurettin-05	2
39	Zehra-01	3	93	Eda-03	3	147	Gizem-05	3
40	Ozcan-01	3	94	Ferit-03	2	148	Nurcan-05	2
41	Necla-01	2	95	Bora-03	1	149	Doga-05	4
42	Ismail-01	4	96	Nurettin-03	2	150	Haydar-05	6
43	Eda-01	3	97	Gizem-03	3	151	Suheyla-06	1
44	Ferit-01	2	98	Nurcan-03	2	152	Vural-06	1
45	Bora-01	1	99	Doga-03	4	153	Aysegul-06	0
46	Nurettin-01	2	100	Haydar-03	6	154	Berrin-06	1
47	Gizem-01	3	101	Suheyla-04	1	155	Mahmut-06	3
48	Nurcan-01	2	102	Vural-04	1	156	Murat-06	0
49	Doga-01	4	103	Aysegul-04	0	157	Hacer-06	6
50	Haydar-01	6	104	Berrin-04	1	158	Ayse-06	5
51	Suheyla-02	1	105	Mahmut-04	3	159	Aylin-06	1
52	Vural-02	1	106	Murat-04	0	160	Ebru-06	2
53	Aysegul-02	0	107	Hacer-04	6	161	Didem-06	1
54	Berrin-02	1	108	Ayse-04	5	162	Tuna-06	7

Table A.2 (continued)

FRAME B								
Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i
163	Hakan-06	2	193	Eda-07	3	223	Nurcan-08	2
164	Zehra-06	3	194	Ferit-07	2	224	Doga-08	4
165	Ozcan-06	3	195	Bora-07	1	225	Haydar-08	6
166	Necla-06	2	196	Nurettin07	2	226	Suheyla-09	1
167	Ismail-06	4	197	Gizem-07	3	227	Vural-09	1
168	Eda-06	3	198	Nurcan-07	2	228	Aysegul-09	0
169	Ferit-06	2	199	Doga-07	4	229	Berrin-09	1
170	Bora-06	1	200	Haydar-07	6	230	Mahmut-09	3
171	Nurettin-06	2	201	Suheyla-08	1	231	Murat-09	0
172	Gizem-06	3	202	Vural-08	1	232	Hacer-09	6
173	Nurcan-06	2	203	Aysegul-08	0	233	Ayse-09	5
174	Doga-06	4	204	Berrin-08	1	234	Aylin-09	1
175	Haydar-06	6	205	Mahmut-08	3	235	Ebru-09	2
176	Suheyla-07	1	206	Murat-08	0	236	Didem-09	1
177	Vural-07	1	207	Hacer-08	6	237	Tuna-09	7
178	Aysegul-07	0	208	Ayse-08	5	238	Hakan-09	2
179	Berrin-07	1	209	Aylin-08	1	238	Zehra-09	3
180	Mahmut-07	3	210	Ebru-08	2	240	Ozcan-09	3
181	Murat-07	0	211	Didem-08	1	241	Necla-09	2
182	Hacer-07	6	212	Tuna-08	7	242	Ismail-09	4
183	Ayse-07	5	213	Hakan-08	2	243	Eda-09	3
184	Aylin-07	1	214	Zehra-08	3	244	Ferit-09	2
185	Ebru-07	2	215	Ozcan-08	3	245	Bora-09	1
186	Didem-07	1	216	Necla-08	2	246	Nurettin-09	2
187	Tuna-07	7	217	Ismail-08	4	247	Gizem-09	3
188	Hakan-07	2	218	Eda-08	3	248	Nurcan-09	2
189	Zehra-07	3	219	Ferit-08	2	249	Doga-09	4
190	Ozcan-07	3	220	Bora-08	1	250	Haydar-09	6
191	Necla-07	2	221	Nurettin-08	2			
192	Ismail-07	4	222	Gizem-08	3			

Table A.3: List of Population Elements for Frame C

FRAME C								
Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i
1	Suheyla	1	55	Nurcan-02	2	109	Hakan-05	2
2	Aylin	1	56	Hacer-02	6	110	Ozgur-05	5
3	Berrin	1	57	Umit-02	5	111	Necla-05	2
4	Hakan	2	58	Tuna-02	7	112	Haydar-05	6
5	Ozgur	5	59	Osman-02	3	113	Zehra-05	3
6	Necla	2	60	Gizem-02	3	114	Murat-05	0
7	Haydar	6	61	Vural-02	1	115	Kaya-05	6
8	Zehra	3	62	Ilsu-02	0	116	Nurettin-05	2
9	Murat	0	63	Banu-02	2	117	Aysenaz05	0
10	Kaya	6	64	Suheyla-03	1	118	Nurcan-05	2
11	Nurettin	2	65	Aylin-03	1	119	Hacer-05	6
12	Aysenaz	0	66	Berrin-03	1	120	Umit-05	5
13	Nurcan	2	67	Hakan-03	2	121	Tuna-05	7
14	Hacer	6	68	Ozgur-03	5	122	Osman-05	3
15	Umit	5	69	Necla-03	2	123	Gizem-05	3
16	Tuna	7	70	Haydar-03	6	124	Vural-05	1
17	Osman	3	71	Zehra-03	3	125	Ilsu-05	0
18	Gizem	3	72	Murat-03	0	126	Banu-05	2
19	Vural	1	73	Kaya-03	6	127	Suheyla-06	1
20	Ilsu	0	74	Nurettin-03	2	128	Aylin-06	1
21	Banu	2	75	Aysenaz-03	0	129	Berrin-06	1
22	Suheyla01	1	76	Nurcan-03	2	130	Hakan-06	2
23	Aylin-01	1	77	Hacer-03	6	131	Ozgur-06	5
24	Berrin-01	1	78	Umit-03	5	132	Necla-06	2
25	Hakan-01	2	79	Tuna-03	7	133	Haydar-06	6
26	Ozgur-01	5	80	Osman-03	3	134	Zehra-06	3
27	Necla-01	2	81	Gizem-03	3	135	Murat-06	0
28	Haydar-01	6	82	Vural-03	1	136	Kaya-06	6
29	Zehra-01	3	83	Ilsu-03	0	137	Nurettin-06	2
30	Murat-01	0	84	Banu-03	2	138	Aysenaz06	0
31	Kaya-01	6	85	Suheyla-04	1	139	Nurcan-06	2
32	Nurettin01	2	86	Aylin-04	1	140	Hacer-06	6
33	Aysenaz01	0	87	Berrin-04	1	141	Umit-06	5
34	Nurcan-01	2	88	Hakan-04	2	142	Tuna-06	7
35	Hacer-01	6	89	Ozgur-04	5	143	Osman-06	3
36	Umit-01	5	90	Necla-04	2	144	Gizem-06	3
37	Tuna-01	7	91	Haydar-04	6	145	Vural-06	1
38	Osman-01	3	92	Zehra-04	3	146	Ilsu-06	0
39	Gizem-01	3	93	Murat-04	0	147	Banu-06	2
40	Vural-01	1	94	Kaya-04	6	148	Suheyla07	1
41	Ilsu-01	0	95	Nurettin-04	2	149	Aylin-07	1
42	Banu-01	2	96	Aysenaz-04	0	150	Berrin-07	1
43	Suheyla02	1	97	Nurcan-04	2	151	Hakan-07	2
44	Aylin-02	1	98	Hacer-04	6	152	Ozgur-07	5
45	Berrin-02	1	99	Umit-04	5	153	Necla-07	2
46	Hakan-02	2	100	Tuna-04	7	154	Haydar-07	6
47	Ozgur-02	5	101	Osman-04	3	155	Zehra-07	3
48	Necla-02	2	102	Gizem-04	3	156	Murat-07	0
49	Haydar-02	6	103	Vural-04	1	157	Kaya-07	6
50	Zehra-02	3	104	Ilsu-04	0	158	Nurettin-07	2
51	Murat-02	0	105	Banu-04	2	159	Aysenaz07	0
52	Kaya-02	6	106	Suheyla-05	1	160	Nurcan-07	2
53	Nurettin-02	2	107	Aylin-05	1	161	Hacer-07	6
54	Aysenaz-02	0	108	Berrin-05	1	162	Umit-07	5

Table A.3 (continued)

FRAME C								
Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i
163	Tuna-07	7	179	Nurettin-08	2	195	Necla-09	2
164	Osman-07	3	180	Aysenaz08	0	196	Haydar-09	6
165	Gizem-07	3	181	Nurcan-08	2	197	Zehra-09	3
166	Vural-07	1	182	Hacer-08	6	198	Murat-09	0
167	Ilsu-07	0	183	Umit-08	5	199	Kaya-09	6
168	Banu-07	2	184	Tuna-08	7	200	Nurettin09	2
169	Suheyla08	1	185	Osman-08	3	201	Aysenaz-09	0
170	Aylin-08	1	186	Gizem-08	3	202	Nurcan-09	2
171	Berrin-08	1	187	Vural-08	1	203	Hacer-09	6
172	Hakan-08	2	188	Ilsu-08	0	204	Umit-09	5
173	Ozgur-08	5	189	Banu-08	2	205	Tuna-09	7
174	Necla-08	2	190	Suheyla09	1	206	Osman-09	3
175	Haydar-08	6	191	Aylin-09	1	207	Gizem-09	3
176	Zehra-08	3	192	Berrin-09	1	208	Vural-09	1
177	Murat-08	0	193	Hakan-09	2	209	Ilsu-09	0
178	Kaya-08	6	194	Ozgur-09	5	210	Banu-09	2

APPENDIX B

TABLES OF POPULATION ELEMENTS CLASSIFICATION TO THREE DOMAINS FOR DUAL FRAME

Table B.1: List of Population Elements in Domain a for Dual Frame

Population Elements In Domain a								
Pop. ID i	**ID Label**	**Pop. Element Value y_i**	**Pop. ID I**	**ID Label**	**Pop. Element Value y_i**	**Pop. ID i**	**ID Label**	**Pop. Element Value y_i**
3	Arif	2	70	Kamil-02	2	135	Sule-04	4
4	Aynur	3	72	Deniz-02	5	136	Ahmet-04	2
9	Birsen	4	73	Nilgun-02	2	140	Asli-04	1
10	Kamil	2	75	Sule-02	4	144	Ozlem-04	4
12	Deniz	5	76	Ahmt-02	2	146	Umit-04	5
13	Nilgun	2	80	Asli-02	1	147	Ilsu-04	0
15	Sule	4	84	Ozlem-02	4	148	Aysenaz-04	0
16	Ahmet	2	86	Umit-02	5	149	Osman-04	3
20	Asli	1	87	Ilsu-02	0	150	Mehmet-04	1
24	Ozlem	4	88	Aysenaz-02	0	151	Arif-05	2
26	Umit	5	89	Osman-02	3	154	Aynur-05	3
27	Ilsu	0	90	Mehmet-02	1	159	Birsen-05	4
28	Aysenaz	0	93	Arif-03	2	160	Kamil-05	2
29	Osman	3	94	Aynur	3	162	Deniz-05	5
30	Mehmet	1	99	Birsen-03	4	163	Nilgun-05	2
33	Arif-01	2	100	Kamil-03	2	165	Sule-05	4
34	Aynur-01	3	102	Deniz-03	5	166	Ahmet-05	2
39	Birsen-01	4	103	Nilgun-03	2	170	Asli-05	1
40	Kamil-01	2	105	Sule-03	4	174	Ozlem-05	4
42	Deniz-01	5	106	Ahmet-03	2	176	Umit-05	5
43	Nilun-01	2	110	Asli-03	1	177	Ilsu-05	0
45	Sule-01	4	114	Ozlem-03	4	178	Aysenaz-05	0
46	Ahmet-01	2	116	Umit-03	5	179	Osman-05	3
50	Asli-01	1	117	Ilsu-03	0	180	Mehmet-05	1
54	Ozlem-01	4	118	Aysenaz-03	0	183	Arif-06	2
56	Umit-01	5	119	Osman-03	3	184	Aynur-06	3
57	Ilsu-01	0	120	Mehmet-03	1	189	Birsen-06	4
58	Aysenaz-01	0	123	Arif-04	2	190	Kamil-06	2
59	Osman-01	3	124	Aynur-04	3	192	Deniz-06	5
60	ehmet-01	1	129	Birsen-04	4	193	Nilgun-06	2
63	Arif-02	2	130	Kamil-04	2	195	Sule-06	4
64	Aynur-02	3	132	Deniz-04	5	196	Ahmet-06	2
69	Birsen-02	4	133	Nilgun-04	2	200	Asli-06	1

Table B.1 (continued)

Population Elements In Domain a								
Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID I	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i
204	Ozlem-06	4	237	Ilsu-07	0	269	Osman-08	3
206	Umit-06	5	238	Aysenaz-07	0	270	Mehmet-08	1
207	Ilsu-06	0	239	Osman-07	3	273	Arif-09	2
208	Aysenaz-06	0	240	Mehmet-07	1	274	Aynur-09	3
209	Osman-06	3	243	Arif-08	2	279	Birsen-09	4
210	Mehmet-06	1	244	Aynur-08	3	280	Kamil-09	2
213	Arif-07	2	249	Birsen-08	4	282	Deniz-09	5
214	Aynur-07	3	250	Kamil-08	2	283	Nilgun-09	2
219	Birsen-07	4	252	Deniz-08	5	285	Sule-09	4
220	Kamil-07	2	253	Nilgun-08	2	286	Ahmet-09	2
222	Deniz-07	5	255	Sule-08	4	290	Asli-09	1
223	Nilgun-07	2	256	Ahmet-08	2	294	Ozlem-09	4
225	Sule-07	4	260	Asli-08	1	296	Umit-09	5
226	Ahmet-07	2	264	Ozlem-08	4	297	Ilsu-09	0
230	Asli-07	1	266	Umit-08	5	298	Aysenaz-09	0
234	Ozlem-07	4	267	Ilsu-08	0	299	Osman-09	3
236	Umit-07	5	268	Aysenaz-08	0	300	Mehmet-09	1

Table B.2: List of Population Elements in Domain ab for Dual Frame

Population Elements In Domain ab							
Pop.ID			Pop. Element Value	Pop.ID			Pop. Element Value
i(A)	i(B)	ID Label	y_i	i(A)	i(B)	ID Label	y_i
1	9	Aylin	1	107	95	Bora-03	1
2	1	Suheyla	1	108	82	Hacer-03	6
5	8	Ayse	5	109	96	Nurettin-03	2
6	13	Hakan	2	111	77	Vural-03	1
7	10	Ebru	2	112	100	Haydar-03	6
8	15	Ozcan	3	113	87	Tuna-03	7
11	3	Aysegul	0	115	79	Berrin-03	1
14	14	Zehra	3	121	109	Aylin-04	1
17	20	Bora	1	122	101	Suheyla-04	1
18	7	Hacer	6	125	108	Ayse04	5
19	21	Nurettin	2	126	112	Hakan-04	2
21	2	Vural	1	127	110	Ebru-04	2
22	25	Haydar	6	128	115	Ozcan-04	3
23	12	Tuna	7	131	103	Aysegul-04	0
25	4	Berrin	1	134	114	Zehra-04	3
31	34	Aylin-01	1	137	120	Bora-04	1
32	26	Suheyla-01	1	138	107	Hacer-04	6
35	33	Ayse-01	5	139	121	Nurettin-04	2
36	38	Hakan-01	2	141	102	Vural-04	1
37	35	Ebru-01	2	142	125	Haydar-04	6
38	40	Ozcan-01	3	143	112	Tuna-04	7
41	28	Aysegul-01	0	145	104	Berrin-04	1
44	39	Zehra-01	3	151	134	Aylin-05	1
47	45	Bora-01	1	152	126	Suheyla-05	1
48	32	Hacer-01	6	155	133	Ayse-05	5
49	46	Nurettin-01	2	156	138	Hakan-05	2
51	27	Vural-01	1	157	135	Ebru-05	2
52	50	Haydar-01	6	158	140	Ozcan-05	3
53	37	Tuna-01	7	161	128	Aysegul-05	0
55	29	Berrin-01	1	164	139	Zehra5	3
61	59	Aylin-02	1	167	145	Bora-05	1
62	51	Suheyla-02	1	168	132	Hacer-05	6
65	58	Ayse-02	5	169	145	Nurettin-05	2
66	63	Hakan-02	2	171	127	Vural-05	1
67	60	Ebru-02	2	172	150	Haydar-05	6
68	65	Ozcan-02	3	173	137	Tuna-05	7
71	53	Aysegul-02	0	175	129	Berrin-05	1
74	64	Zehra-02	3	181	159	Aylin-06	1
77	70	Bora-02	1	182	151	Suheyla-06	1
78	57	Hacer-02	6	185	158	Ayse-06	5
79	71	Nurettin-02	2	186	163	Hakan-06	2
81	52	Vural-02	1	187	160	Ebru-06	2
82	75	Haydar-02	6	188	165	Ozcan-06	3
83	62	Tuna-02	7	191	153	Aysegul-06	0
85	54	Berrin-02	1	194	164	Zehra-06	3
91	84	Aylin-03	1	197	170	Bora-06	1
92	76	Suheyla-03	1	198	157	Hacer-06	6
95	83	Ayse-03	5	199	171	Nurettin-06	2
96	88	Hakan-03	2	201	152	Vural-06	1
97	85	Ebru-03	2	202	175	Haydar-06	6
98	90	Ozcan-03	3	203	162	Tuna-06	7
101	78	Aysegul-03	0	205	154	Berrin-06	1
104	79	Zehra-03	3	211	184	Aylin-07	1

Table B.2 (continued)

Population Elements In Domain ab							
Pop.ID			Pop. Element Value	Pop.ID			Pop. Element Value
i(A)	i(B)	ID Label	y_i	i(A)	i(B)	ID Label	y_i
212	176	Suheyla-07	1	257	220	Bora-08	1
215	183	Ayse-07	5	258	207	Hacer-08	6
216	188	Hakan-07	2	259	221	Nurettin-08	2
217	185	Ebru-07	2	261	202	Vural-08	1
218	190	Ozcan-07	3	262	25	Haydar-08	6
221	178	Aysegul-07	0	263	212	Tuna-08	7
224	189	Zehra-07	3	265	204	Berrin-08	1
227	195	Bora-07	1	271	234	Aylin-09	1
228	182	Hacer-07	6	272	226	Suheyla-09	1
229	196	Nurettin-07	2	275	233	Ayse-09	5
231	177	Vural-07	1	276	238	Hakan-09	2
232	200	Haydar-07	6	277	235	Ebru-09	2
233	187	Tuna-07	7	278	240	Ozan-09	3
235	179	Berrin-07	1	281	228	Aysegul-09	0
241	209	Aylin-08	1	284	239	Zehra-09	3
242	201	Suheyla-08	1	287	245	Bora-09	1
245	208	Ayse-08	5	288	232	Hacer-09	6
246	213	Hakan-08	2	289	246	Nurettin-09	2
247	210	Ebru-08	2	291	227	Vural-09	1
248	215	Ozcan-08	3	292	250	Haydar-09	6
251	203	Aysegul-08	0	293	237	Tuna-09	7
254	214	Zehra-08	3	295	229	Berrin-09	1

Table B.3: List of Population Elements in Domain b for Dual Frame

Population Elements In Domain b					
Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i
5	Mahmut	3	130	Mahmut-05	3
6	Murat	0	131	Murat-05	0
11	Didem	1	136	Didem-05	1
16	Necla	2	141	Necla-05	2
17	Ismail	4	142	Ismail-05	4
18	Eda	3	123	Eda-05	3
19	Ferit	2	144	Ferit-05	2
22	Gizem	3	147	Gizem -05	3
23	Nurcan	2	148	Nurcan-05	2
24	Doga	4	149	Doga-05	4
30	Mahmut-01	3	155	Mahmut-06	3
31	Murat-01	0	156	Murat-06	0
36	Didem-01	1	161	Didem-06	1
41	Necla-01	2	166	Necla-06	2
42	Ismail-01	4	167	Ismail-06	4
43	Eda-01	3	168	Eda-06	3
44	Ferit-01	2	169	Ferit-06	2
47	Gizem -01	3	172	Gizem -06	3
48	Nurcan-01	2	173	Nurcan-06	2
49	Doga-01	4	174	Doga-06	4
55	Mahmut-02	3	180	Mahmut-07	3
56	Murat-02	0	181	Murat-07	0
61	Didem-02	1	186	Didem-07	1
66	Necla-02	2	191	Necla-07	2
67	Ismail-02	4	192	Ismail-07	4
68	Eda-02	3	193	Eda-07	3
69	Ferit-02	2	194	Ferit-07	2
72	Gizem -02	3	197	Gizem -07	3
73	Nurcan-02	2	198	Nurcan-07	2
74	Doga-02	4	199	Doga-07	4
80	Mahmut-03	3	205	Mahmut-08	3
81	Murat-03	0	206	Murat-08	0
86	Didem-03	1	211	Didem-08	1
91	Necla-03	2	216	Necla-08	2
92	Ismail-03	4	217	Ismail-08	4
93	Eda-03	3	218	Eda-08	3
94	Ferit-03	2	219	Ferit-08	2
97	Gizem -03	3	222	Gizem 08	3
98	Nurcan-03	2	223	Nurcan-08	2
99	Doga-03	4	224	Doga-08	4
105	Mahmt-04	3	230	Mahmut-09	3
106	Murat-04	0	231	Murat-09	0
111	Didem-04	1	236	Didem-09	1
116	Necla-04	2	241	Necla-09	2
117	Ismail-04	4	242	Ismail-09	4
118	Eda-04	3	243	Eda-09	3
119	Ferit-04	2	244	Ferit-09	2
122	Gizem -04	3	247	Gizem -09	3
123	Nurcan-04	2	248	Nurcan-09	2
124	Doga-04	4	249	Doga-09	4

APPENDIX C

TABLES OF POPULATION ELEMENTS CLASSIFICATION TO SEVEN DOMAINS FOR THREE FRAMES

Table C.1: List of Population Elements in Domain a for Three Frames

Population Elements In Domain a								
Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i
3	Arif	2	93	Arif-03	2	183	Arif-06	2
4	Aynur	3	94	Aynur-03	3	184	Aynur-06	3
9	Birsen	4	99	Birsen-03	4	189	Birsen-06	4
10	Kamil	2	100	Kamil-03	2	190	Kamil-06	2
12	Deniz	5	102	Deniz-03	5	192	Deniz-06	5
13	Nilgun	2	103	Nilgun-03	2	193	Nilgu-06	2
15	Sule	4	105	Sule-03	4	195	Sule-06	4
16	Ahmet	2	106	Ahmet-03	2	196	Ahmet-06	2
20	Asli	1	110	Asli-03	1	200	Asli-06	1
24	Ozlem	4	114	Ozlem-03	4	204	Ozlem-06	4
30	Mehmet	1	120	Mehmet-3	1	210	Mehmet-06	1
33	Arif-01	2	123	Arif-04	2	213	Arif-07	2
34	Aynur-01	3	124	Aynur-04	3	214	Aynur-07	3
39	Birsen-01	4	129	Birsen-04	4	219	Birsen-07	4
40	Kamil-01	2	130	Kamil-04	2	220	Kamil-07	2
42	Deniz-01	5	132	Deniz-04	5	222	Deniz-07	5
43	Nilgun-01	2	133	Nilgun-04	2	223	Nilun-07	2
45	Sule-01	4	135	Sule-04	4	225	Sule-07	4
46	Ahmet-01	2	136	Ahmet-04	2	226	Ahmet-07	2
50	Asli-01	1	140	Asli-04	1	230	Asli-07	1
54	Ozlem-01	4	144	Ozlem-04	4	234	Ozlem-07	4
60	Mehmet-01	1	150	Mehmet-04	1	240	Mehmet-07	1
63	Arif-02	2	153	Arif-05	2	243	Arif-08	2
64	Aynur-02	3	154	Aynur-05	3	244	Aynur-08	3
69	Birsen-02	4	159	Birsen-05	4	249	Birsen-08	4
70	Kamil-02	2	160	Kamil-05	2	250	Kamil-08	2
72	Deniz-02	5	162	Deniz-05	5	252	Deniz-08	5
73	Nilgun-02	2	163	Nilgun-5	2	253	Nilgun-08	2
75	Sule-02	4	165	Sule-05	4	255	Sule-08	4
76	Ahmet-02	2	166	Ahmet-05	2	256	Ahmet-08	2
80	Asli-02	1	170	Asli-05	1	260	Asli-08	1
84	Ozlem-02	4	174	Ozlem-05	4	264	Ozlem-08	4
90	Mehmet-02	1	180	Mehmet-05	1	270	Mehmet-08	1

Table C.1 (continued)

Population Elements In Domain a								
Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i
273	Arif-09	2	282	Deniz-09	5	290	Asli-09	1
274	Aynur-09	3	283	Nilgun-09	2	294	Ozlem-09	4
279	irsen-09	4	285	Sule-09	4	300	Mehmet09	1
280	Kamil-09	2	286	Ahmet-09	2			

Table C.2: List of Population Elements in Domain b for Three Frames

Population Elements In Domain b					
Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i
5	Mahmut	3	130	Mahmut-05	3
11	Didem	1	136	Didem-05	1
17	Ismail	4	142	Ismail-05	4
18	Eda	3	123	Eda-05	3
19	Ferit	2	144	Ferit-05	2
24	Doga	4	149	Doga-05	4
30	Mahmut-01	3	155	Mahmut-06	3
36	Didem-01	1	161	Didem-06	1
42	Ismail-01	4	167	Ismail-06	4
43	Eda-01	3	168	Eda-06	3
44	Ferit-01	2	169	Ferit-06	2
49	Doga-01	4	174	Doga-06	4
55	Mahmut-02	3	180	Mahmut-07	3
61	Didem-02	1	186	Didem-07	1
67	Ismail-02	4	192	Ismail-07	4
68	Eda-02	3	193	Eda-07	3
69	Ferit-02	2	194	Ferit-07	2
74	Doga-02	4	199	Doga-07	4
80	Mahmut-03	3	205	Mahmut-08	3
86	Didem-03	1	211	Didem-08	1
92	Ismail-03	4	217	Ismail-08	4
93	Eda-03	3	218	Eda-08	3
94	Ferit-03	2	219	Ferit-08	2
99	Doga-03	4	224	Doga-08	4
105	Mahmut-04	3	230	Mahmut-09	3
111	Didem-04	1	236	Didem-09	1
117	Ismail-04	4	242	Ismail-09	4
118	Eda-04	3	243	Eda-09	3
119	Ferit-04	2	244	Ferit-09	2
124	Doga-04	4	249	Doga-09	4

Table C.3: List of Population Elements in Domain c for Three Frames

Population Elements In Domain c					
Pop. ID i	ID Label	Pop. Element Value y_i	Pop. ID i	ID Label	Pop. Element Value y_i
5	Ozgur	5	110	Ozgur-05	5
10	Kaya	6	115	Kaya-05	6
21	Banu	2	126	Banu-05	2
26	Ozgur-01	5	131	Ozgur-06	5
31	Kaya-01	6	136	Kaya-06	6
42	Banu-01	2	147	Banu-06	2
47	Ozgur-02	5	152	Ozgur-07	5
52	Kaya-02	6	157	Kaya-07	6
63	Banu-02	2	168	Banu-07	2
68	Ozgur-03	5	173	Ozgur-08	5
73	Kaya-03	6	178	Kaya-08	6
84	Banu-03	2	189	Banu-08	2
89	Ozgur-04	5	194	Ozgur-09	5
94	Kaya-04	6	199	Kaya-09	6
105	Banu-04	2	210	Banu-09	2

Table C.4: List of Population Elements in Domain ac for Three Frames

Population Elements In Domain ac							
Pop.ID			Pop. Element Value	Pop.ID			Pop. Element Value
i(A)	i(C)	ID Label	y_i	i(A)	i(C)	ID Label	y_i
26	15	Umit	5	176	120	Umit-05	5
27	20	Ilsu	0	177	125	Ilsu-05	0
28	12	Aysenaz	0	178	117	Aysenaz-05	0
29	17	Osman	3	179	122	Osman-05	3
56	36	Umit-01	5	206	141	Umit-06	5
57	41	Ilsu-01	0	207	146	Ilsu-06	0
58	33	Aysenaz-01	0	208	138	Aysenaz-06	0
59	38	Osman-01	3	209	143	Osman-06	3
86	57	Umit-02	5	236	162	Umit-07	5
87	62	Ilsu-02	0	237	167	Ilsu-07	0
88	54	Aysenaz-02	0	238	159	Aysenaz-07	0
89	59	Osman-02	3	239	164	Osman-07	3
116	78	Umit-03	5	266	183	Umit-08	5
117	83	Ilsu-03	0	267	188	Ilsu-08	0
118	75	Aysenaz-03	0	268	180	Aysenaz-08	0
119	80	Osman-03	3	269	185	Osman-08	3
146	99	Umit-04	5	296	204	Umit-09	5
147	104	Ilsu-04	0	297	209	Ilsu-09	0
148	96	Aysenaz-04	0	298	201	Aysenaz-09	0
149	101	Osman-04	3	299	206	Osman-09	3

Table C.5: List of Population Elements in Domain ab for Three Frames

Population Elements In Domain ab							
Pop.ID i(A)	Pop.ID i(B)	ID Label	Pop. Element Value y_i	Pop.ID i(A)	Pop.ID i(B)	ID Label	Pop. Element Value y_i
5	8	Ayse	5	155	133	Ayse-05	5
7	10	Ebru	2	157	135	Ebru-05	2
8	15	Ozcan	3	158	140	Ozcan-05	3
11	3	Aysegul	0	161	128	Aysegul-05	0
17	20	Bora	1	167	145	Bora-05	1
35	33	Ayse-01	5	185	158	Ayse-06	5
37	35	Ebru-01	2	187	160	Ebru-06	2
38	40	Ozcan-01	3	188	165	Ozcan-06	3
41	28	Aysegul-01	0	191	153	Aysegu-06	0
47	45	Bora-01	1	197	170	Bora-06	1
65	58	Ayse-02	5	215	183	Ayse-07	5
67	60	Ebru-02	2	217	185	Ebru-07	2
68	65	Ozcan-02	3	218	190	Ozcan-07	3
71	53	Aysegul-02	0	221	178	Aysegul-07	0
77	70	Bora-02	1	227	195	Bora-07	1
95	83	Ayse-03	5	245	208	Ayse-08	5
97	85	Ebru-03	2	247	210	Ebru-08	2
98	90	Ozcan-03	3	248	215	Ozcan-08	3
101	78	Aysegul-03	0	251	203	Aysegul-08	0
107	95	Bora-03	1	257	220	Bora-08	1
125	108	Ayse-04	5	275	233	Ayse-09	5
127	110	Ebru-04	2	277	235	Ebru-09	2
128	115	Ozcan-04	3	278	240	Ozcan-09	3
131	103	Aysegul-04	0	281	228	Aysegul-09	0
137	120	Bora-04	1	287	245	Bora-09	1

Table C.6: List of Population Elements in Domain bc for Three Frames

Population Elements in Domain bc							
Pop.ID i(B)	Pop.ID i(C)	ID Label	Pop. Element Value y_i	Pop.ID i(B)	Pop.ID i(C)	ID Label	Pop. Element Value y_i
6	9	Murat	0	131	114	Murat-05	0
22	18	Gizem	3	147	123	Gizem-05	3
23	13	Nurcan	2	148	118	Nurcan-05	2
16	6	Necla	2	141	111	Necla-05	2
31	30	Murat-01	0	156	135	Murat-06	0
47	39	Gizem-01	3	172	144	Gizem-06	3
48	34	Nurcan-01	2	173	139	Nurcan-06	2
41	27	Necla-01	2	166	132	Necla-06	2
56	51	Murat-02	0	181	156	Murat-07	0
72	60	Gizem-02	3	197	165	Gizem-07	3
73	55	Nurcan-02	2	198	160	Nurcan-07	2
66	48	Necla-02	2	191	153	Necla-07	2
81	72	Murat-03	0	206	177	Murat-08	0
97	81	Gizem-03	3	222	186	Gizem-08	3
98	76	Nurcan-03	2	223	181	Nurcan-08	2
91	69	Necla-03	2	216	174	Necla-08	2
106	93	Murat-04	0	231	198	Murat-09	0
122	102	Gizem-04	3	247	207	Gizem-09	3
123	97	Nurcan-04	2	248	202	Nurcan-09	2
116	90	Necla-04	2	241	195	Necla-09	2

Table C.7: List of Population Elements in Domain abc for Three Frames

Population Elements in Domain abc									
	Pop.ID			Pop. Element Value		Pop.ID			Pop. Element Value
i(A)	i(B)	i(C)	ID Label	y_i	i(A)	i(B)	i(C)	ID Label	y_i
1	9	2	Aylin	1	151	134	107	Aylin-05	1
2	1	1	Suheyla	1	152	126	106	Suheyla-05	1
6	13	4	Hakan	2	156	138	109	Hakan-05	2
14	14	8	Zehra	3	164	139	113	Zehra-05	3
18	7	14	Hacer	6	168	132	119	Hacer-05	6
19	21	11	Nurettin	2	169	146	116	Nurettin-05	2
21	2	19	Vural	1	171	127	124	Vural-05	1
22	25	7	Haydar	6	172	150	112	Haydar-05	6
23	12	16	Tuna	7	173	137	121	Tuna-05	7
25	4	3	Berrin	1	175	129	108	Berrin-05	1
31	34	23	Al-01	1	181	159	128	Aylin-06	1
32	26	22	Suheyla-01	1	182	151	127	Suheyla-06	1
36	38	25	Hakan-01	2	186	163	130	Hakan-06	2
44	39	29	Zehra-01	3	194	164	134	Zehra-06	3
48	32	35	Hacer-01	6	198	157	140	Hacer-06	6
49	46	32	Nurettin-01	2	199	171	137	Nurettin-06	2
51	27	40	Vural-01	1	201	152	145	Vural-06	1
52	50	28	Haydar-01	6	202	175	140	Haydar-06	6
53	37	37	Tuna-1	7	203	162	142	Tuna-06	7
55	29	24	Berrin-01	1	205	154	129	Berrin-06	1
61	59	44	Aylin-02	1	211	184	149	Aylin-07	1
62	51	43	Suheyla-02	1	212	176	148	Suheyla-07	1
66	63	46	Hakan-02	2	216	188	151	Hakan-07	2
74	64	50	Zehra-02	3	224	189	155	Zehra-07	3
78	57	56	Hacer-02	6	228	182	161	Hacer-07	6
79	71	53	Nurettin-02	2	229	196	158	Nurettin-07	2
81	52	61	Vural-02	1	231	177	166	Vural-07	1
82	75	46	Haydar-02	6	232	200	154	Haydar-07	6
83	62	58	Tuna-02	7	233	187	163	Tuna-07	7
85	54	45	Berrin-02	1	235	179	150	Berrin-07	1
91	84	65	Aylin-03	1	241	209	170	Aylin-08	1
92	76	64	Suheyla-03	1	242	201	169	Suheyla-08	1
96	88	67	Hakan-03	2	246	213	172	Hakan-08	2
104	89	71	Zehra-03	3	254	214	176	Zehra-08	3
108	82	77	Hacer-03	6	258	207	182	Hacer-08	6
109	96	74	Nuretin-03	2	259	221	179	Nurettin-08	2
111	77	82	Vural-03	1	261	202	187	Vural-08	1
112	100	70	Haydar-03	6	262	225	175	Haydar-08	6
113	87	79	Tuna-03	7	263	212	184	Tuna-08	7
115	79	66	Berrin-3	1	265	204	171	Berrin-08	1
121	109	86	Aylin-04	1	271	234	191	Aylin-09	1
122	101	85	Suheyla-04	1	272	226	190	Suheyla-09	1
126	113	88	Hakan-04	2	276	238	193	Hakan-09	2
134	114	92	Zehra-04	3	284	239	197	Zehra-09	3
138	107	98	Her-04	6	288	232	203	Hacer-09	6
139	121	95	Nurettin-04	2	289	246	200	Nurettin-09	2
141	102	103	Vural-04	1	291	227	208	Vural-09	1
142	125	91	Haydar-04	6	292	250	196	Haydar-09	6
143	112	100	Tuna-04	7	293	237	205	Tuna-09	7
145	104	87	Berrin-04	1	295	229	192	Berrin-09	1

APPENDIX D

TABLES OF SAMPLE ELEMENTS FOR APPLICATION

Table D.1: List of Sample Elements for Independent Simple Random Sample s_A Selected from Frame A

Sampe A s_A							
Sample ID j	**Pop. ID i**	**ID Label**	**Sample Element Value y_i**	**Sample ID j**	**Pop. ID i**	**ID Label**	**Sample Element Value y_i**
1	71	Aysegul-02	0	31	173	Tuna-05	7
2	143	Tuna-04	7	32	203	Tuna-06	7
3	296	Umit-09	5	33	134	Zehra-04	3
4	170	Asli-05	1	34	53	Tuna-01	7
5	200	Asli-06	1	35	123	Arif-04	2
6	161	Aysegul-05	0	36	77	Bora-02	1
7	248	Ozcan-08	3	37	176	Umit-05	5
8	131	Aysegul-04	0	38	158	Ozcan-05	3
9	83	Tuna-02	7	39	213	Arif-07	2
10	119	Osman-03	3	40	149	Osman-04	3
11	146	Umit-04	5	41	153	Arif-05	2
12	56	Umit-01	5	42	80	Asli-02	1
13	290	Asli-09	1	43	116	Umit-03	5
14	215	Ayse-07	5	44	167	Bora-05	1
15	179	Osman-05	3	45	293	Tuna-09	7
16	197	Bora-06	1	46	140	Asli-04	1
17	287	Bora-09	1	47	8	Ozcan	3
18	33	Arif-01	2	48	254	Zehra-08	3
19	164	Zehra-05	3	49	125	Ayse-04	5
20	68	Ozcan-02	3	50	251	Aysegul-08	0
21	98	Ozcan-03	3	51	155	Ayse-05	5
22	104	Zehra-03	3	52	59	Osman-01	3
23	95	Ayse-03	5	53	5	Ayse	5
24	238	Osman-07	3	54	107	Bora-03	1
25	93	Arif-03	2	55	299	Osman-09	3
26	47	Bora-01	1	56	128	Ozcan-04	3
27	65	Ayse-02	5	57	50	Asli-01	1
28	101	Aysegul-03	0	58	3	Arif	2
29	206	Umit-06	5	59	74	Zehra-02	3
30	14	Zehra	3	60	11	Aysegul	0

Table D.2: List of Sample Elements for Independent Simple Random Sample s_B Selected from Frame B

Sampe B s_B							
Sample ID j	Pop. ID i	ID Label	Sample Element Value y_i	Sample ID j	Pop. ID i	ID Label	Sample Element Value y_i
1	137	Tuna-05	7	26	41	Necla-01	2
2	74	Doga-02	4	27	59	Aylin-02	1
3	90	Ozcan-03	3	28	6	Murat	0
4	241	Necla-09	2	29	246	Nurettin-09	2
5	166	Necla-06	2	30	37	Tuna-01	7
6	127	Vural-05	1	31	206	Murat-08	0
7	171	Nurettin-06	2	32	243	Eda-09	3
8	128	Aysegul-05	0	33	146	Nurein-05	2
9	93	Eda-03	3	34	53	Vural-02	1
10	2	Vural	1	35	116	Necla-04	2
11	115	Ozcan-04	3	36	99	Doga-03	4
12	53	Aysegul-02	0	37	56	Murat-02	0
13	62	Tuna-02	7	38	168	Eda-06	3
14	71	Nurettin-02	2	39	249	Doga-09	4
15	212	Tuna-08	7	40	149	Doga-05	4
16	215	Ozcan-08	3	41	165	Ozcan-06	3
17	9	Aylin	1	42	118	Eda-04	3
18	40	Ozcan-01	3	43	43	Eda-01	3
19	177	Vural-07	1	44	178	Aysegul-07	0
20	34	Aylin-05	1	45	199	Doga-07	4
21	102	Vural-04	1	46	103	Aysegul-04	0
22	106	Murat-04	0	47	131	Murat-05	0
23	96	Nurettin-03	2	48	209	Aylin-08	1
24	109	Aylin-04	1	49	91	Necla-03	2
25	112	Tuna-04	7	50	3	Aysegul	0

Table D.3: List of Sample Elements for Independent Simple Random Sample s_C Selected from Frame C

Sampe C s_C			
Sample ID j	**Pop. ID i**	**ID Label**	**Sample Element Value y_i**
1	153	Necla-07	2
2	175	Hayar-08	6
3	36	Umit-01	5
4	178	Kaya-08	6
5	97	Nurcan-04	2
6	207	Gizem-09	3
7	4	Hakan	2
8	210	Banu-09	2
9	146	Banu-06	2
10	120	Umit-05	5
11	13	Nurcan	2
12	42	Banu-01	2
13	69	Necla-03	2
14	70	Haydar-03	6
15	181	Nurcan-08	2
16	204	Umit-09	5
17	10	Kaya	6
18	180	Aysenaz-08	0
19	151	Hakan-07	2
20	39	Gizem-01	3
21	123	Gizem-05	3
22	127	Suheyla-06	1
23	73	Kaya-03	6
24	96	Aysenaz-04	0
25	1	Suheyla	1
26	27	Necla-01	2
27	46	Hakan-02	2
28	7	Haydar	6
29	43	Suheyla-02	1
30	12	Aysenaz	0

APPENDIX E

TABLES OF SAMPLE ELEMENTS CLASSIFICATION TO TWO DOMAINS FOR DUAL FRAME

Table E.1: List of Sample A Elements in Domains a and ab for Dual Frame

Sampe A s_A								
Sample Elements in Domain a s_a				**Sample Elements in Domain ab** s_{ab}				
Sample ID j	**Pop. ID i**	**ID Label**	**Sample Element Value y_i**	**Sample ID j**	**Pop. ID i**		**ID Label**	**Sample Element Value y_i**
3	296	Umit-09	5	1	71	53	Aysegul--02	0
4	170	Asli-05	1	2	143	112	Tuna-04	7
5	200	Asli-06	1	6	161	128	Aysegul-05	0
10	119	Osman-03	3	7	248	215	Ozcan-08	3
11	146	Umit-04	5	8	131	103	Aysegul-04	0
12	56	Umit-01	5	9	83	62	Tuna-02	7
13	290	Asli-09	1	14	215	183	Ayse-07	5
15	179	Osman-05	3	16	197	170	Bora-06	1
18	33	Arif-01	2	17	287	245	Bora-09	1
24	238	Osman-07	3	19	164	139	Zehra-05	3
25	93	Arif-03	2	20	68	65	Ozcan-02	3
29	206	Umit-06	5	21	98	90	Ozcan-03	3
35	123	Arif-04	2	22	104	89	Zehra-03	3
37	176	Umit-05	5	23	95	83	Ayse-03	5
39	213	Arif-07	2	26	47	45	Bora-01	1
40	149	Osman-04	3	27	65	58	Ayse-02	5
41	153	Arif-05	2	28	101	78	Aysegul-03	0
42	80	Asli-02	1	30	14	14	Zehra	3
43	116	Umit-03	5	31	173	137	Tuna-05	7
46	140	Asli-04	1	32	203	162	Tuna-06	7
52	59	Osman-01	3	33	134	114	Zehra-04	3
55	299	Osman-09	3	34	53	37	Tuna-01	7
57	50	Asli-01	1	36	77	70	Bora-02	1
58	3	Arif	2	38	158	140	Ozcan-05	3
				44	167	145	Bora-05	1
				45	293	237	Tuna-09	7
				47	8	15	Ozcan	3
				48	254	214	Zehra-08	3
				49	125	108	Ayse-04	5
				50	251	203	Aysegul-08	0

Table E.1 (continued)

Sampe A s_A								
Sample Elements in Domain a s_a				Sample Elements in Domain ab s_{ab}				
Sample ID j	Pop. ID i	ID Label	Sample Element Value y_i	Sample ID j	Pop. ID i		ID Label	Sample Element Value y_i
				51	155	133	Ayse-05	5
				53	5	8	Ayse	5
				54	107	95	Bora-03	1
				56	128	115	Ozcan-04	3
				59	74	64	Zehra-02	3
				60	11	3	Aysegul	0

Table E.2: List of Sample B Elements in Domains a and ab for Dual Frame

Sampe B s_B									
Sample Elements in Domain b s_b				**Sample Elements in Domain ab** s''_{ab}					
Sample ID j	**Pop. ID i**	**ID Label**	**Sample Element Value y_i**	**Sample ID j**	**Pop. ID i**		**ID Label**	**Sample Element Value y_i**	
2	74	Doga-02	4	1	137	173	Tuna-05	7	
4	241	Necla-09	2	3	90	98	Ozcan-03	3	
5	166	Necla-06	2	6	127	171	Vural-05	1	
9	93	Eda-03	3	7	171	199	Nurettin-06	2	
22	106	Murat-04	0	8	128	161	Aysegul-05	0	
26	41	Necla-01	2	10	2	21	Vural	1	
28	6	Murat	0	11	115	128	Ozcan-04	3	
31	206	Murat-08	0	12	53	71	Aysegul-2	0	
32	243	Eda-09	3	13	62	83	Tuna-02	7	
35	116	Necla-04	2	14	71	79	Nurettin-02	2	
36	99	Doga-03	4	15	212	263	Tuna-08	7	
37	56	Murat-02	0	16	215	248	Ozcan-08	3	
38	168	Eda-06	3	17	9	1	Aylin	1	
39	249	Doga-09	4	18	40	38	Ozcan-01	3	
40	149	Doga-05	4	19	177	231	Vural-07	1	
42	118	Eda-04	3	20	134	151	Aylin-05	1	
43	43	Eda-01	3	21	102	141	Vural-04	1	
45	199	Doga-07	4	23	96	109	Nurettin-03	2	
47	131	Murat-05	0	24	109	121	Aylin-04	1	
49	91	Necla-03	2	25	112	143	Tuna-04	7	
				27	59	61	Aylin-02	1	
				29	246	289	Nurettin-09	2	
				30	37	53	Tuna-01	7	
				33	146	169	Nurettin-05	2	
				34	52	81	Vural-02	1	
				41	165	188	Ozcan-06	3	
				44	178	221	Aysegul-07	0	
				46	103	131	Aysegul-04	0	
				48	209	241	Aylin-08	1	
				50	3	11	Aysegul	0	

APPENDIX F

TABLES OF SAMPLE ELEMENTS CLASSIFICATION TO FOUR DOMAINS FOR THREE FRAMES

Table F.1: List of Sample A Elements in Domains a for Three Frames

Sample A Elements in Domain a s_a							
Sample ID j	Pop. ID i	ID Label	Sample Element Value y_i	Sample ID j	Pop. ID i	ID Label	Sample Element Value y_i
4	170	Asli-05	1	39	213	Arif-07	2
5	200	Asli-06	1	41	153	Arif-05	2
13	290	Asli-09	1	42	80	Asli-02	1
18	33	Arif-01	2	46	140	Asli-04	1
25	93	Arif-03	2	57	50	Asli-01	1
35	123	Arif-04	2	58	3	Arif	2

Table F.2: List of Sample A Elements in Domains ab for Three Frames

Sample A Elements in Domain ab s_{ab}									
Sample ID	Pop.ID			Sample Element Value	Sample ID	Pop.ID			Sample Element Value
j	i(A)	i(B)	ID Label	y_i	j	i(A)	i(B)	ID Label	y_i
1	71	53	Aysegul-02	0	28	101	78	Aysegul-03	0
6	161	128	Aysegul-05	0	36	77	70	Bora-02	1
7	248	215	Ozcan-08	3	38	158	140	Ozcan-05	3
8	131	103	Aysegul-04	0	44	167	145	Bora-05	1
14	215	183	Ayse-07	5	47	8	15	Ozcan	3
16	197	170	Bora-06	1	49	125	108	Ayse-04	5
17	287	245	Bora-09	1	50	251	203	Aysegul-08	0
20	68	65	Ozcan-02	3	51	155	133	Ayse-05	5
21	98	90	Ozcan-03	3	53	5	8	Ayse	5
23	95	83	Ayse-03	5	54	107	95	Bora-03	1
26	47	45	Bora-01	1	56	128	115	Ozcan-04	3
27	65	58	Ayse-02	5	60	11	3	Aysegul	0

Table F.3: List of Sample A Elements in Domains ac for Three Frames

Sample A Elements in Domain ac $\grave{s}_{ac}$									
Sample ID j	**Pop.ID i(A)**	**i(C)**	**ID Label**	**Sample Element Value y_i**	**Sample ID j**	**Pop.ID i(A)**	**i(C)**	**ID Label**	**Sample Element Value y_i**
3	296	204	Umit-09	5	29	206	141	Umit-06	5
10	119	80	Osman-03	3	37	176	120	Umit-05	5
11	146	99	Umit-04	5	40	149	101	Osman-04	3
12	56	36	Umit-01	5	43	116	78	Umit-09	5
15	179	122	Osman-05	3	52	59	38	Osman-01	3
24	209	164	Osman-07	3	55	299	206	Osman-09	3

Table F.4: List of Sample A Elements in Domains abc for Three Frames

Sample A Elements in Domain abc $\grave{s}_{abc}$					
Sample ID j	**Pop.ID i(A)**	**i(B)**	**i(C)**	**ID Label**	**Sample Element Value y_i**
2	143	112	100	Tuna-04	7
9	83	62	58	Tuna-02	7
19	164	139	113	Zehra-05	3
22	104	89	71	Zehra-03	3
30	14	14	8	Zehra	3
31	173	137	121	Tuna-05	7
32	203	162	142	Tuna-06	7
33	134	114	92	Zehra-04	3
34	53	37	37	Tuna-01	7
45	293	237	205	Tuna-09	7
48	254	214	176	Zehra-08	3
59	74	64	50	Zehra-02	3

Table F.5: List of Sample B Elements in Domains b for Three Frames

Sample B Elements in Domain b s_b							
Sample ID j	Pop. ID i	ID Label	Sample Element Value y_i	Sample ID j	Pop. ID i	ID Label	Sample Element Value y_i
2	74	Doga-02	4	39	249	Doga-09	4
9	93	Eda-03	3	40	149	Doga-05	4
32	243	Eda-09	3	42	118	Eda-04	3
36	99	Doga-03	4	43	43	Eda-01	3
38	168	Eda-06	3	45	199	Doga-07	4

Table F.6: List of Sample B Elements in Domains ab for Three Frames

Sample B Elements in Domain ab $s_{ab}^{``}$									
Sample ID j	Pop.ID i(A)	i(B)	ID Label	Sample Element Value y_i	Sample ID j	Pop.ID i(A)	i(B)	ID Label	Sample Element Value y_i
3	98	90	Ozcan-03	3	18	38	40	Ozcan-01	3
8	161	128	Aysegul-05	0	41	188	165	Ozcan-06	3
11	128	115	Ozcan-04	3	44	221	178	Aysegul-07	0
12	71	53	Aysegul-02	0	46	131	103	Aysegul-04	0
16	248	203	Ozcan-08	3	50	11	3	Aysegul	0

Table F.7: List of Sample B Elements in Domains bc for Three Frames

Sample A Elements in Domain bc $s_{bc}^{`}$									
Sample ID j	Pop.ID i(B)	i(C)	ID Label	Sample Element Value y_i	Sample ID j	Pop.ID i(B)	i(C)	ID Label	Sample Element Value y_i
4	241	195	Necla-09	2	31	206	177	Murat-08	0
5	166	132	Necla-06	2	35	116	90	Necla-04	2
22	106	94	Murat-04	0	37	56	51	Murat-02	0
26	41	27	Necla-01	2	47	131	114	Murat-05	0
28	6	9	Murat	0	49	91	69	Necla-03	2

Table F.8: List of Sample B Elements in Domains abc for Three Frames

Sample A Elements in Domain abc $\ddot{s}_{abc}$					
Sample ID j	Pop.ID i(A)	i(B)	i(C)	ID Label	Sample Element Value y_i
1	173	137	121	Tuna-05	7
6	171	127	124	Vural-05	1
7	199	171	137	Nurettin-06	2
10	21	2	19	Vural	1
13	83	62	58	Tuna-02	7
14	79	71	53	Nurettin-02	2
15	263	212	184	Tuna-08	7
17	1	9	2	Aylin	1
19	231	177	166	Vural-07	1
20	151	134	107	Aylin-05	1
21	141	102	103	Vural-04	1
23	109	96	74	Nurettin-03	2
24	121	109	86	Aylin-04	1
25	143	112	100	Tuna-04	7
27	61	59	44	Aylin-02	1
29	289	246	200	Nurettin-09	2
30	53	37	37	Tuna-01	7
33	169	146	116	Nurettin-05	2
34	81	52	61	Vural-02	1
48	241	209	170	Aylin-08	1

Table F.9: List of Sample C Elements in Domains c for Three Frames

Sample C Elements in Domain c s_c							
Sample ID j	Pop. ID i	ID Label	Sample Element Value y_i	Sample ID j	Pop. ID i	ID Label	Sample Element Value y_i
4	178	Kaya-08	6	12	42	Banu-01	2
8	210	Banu-09	2	17	10	Kaya	6
9	147	Banu-06	2	23	73	Kaya-03	6

Table F.10: List of Sample C Elements in Domains ac for Three Frames

Sample B Elements in Domain ac $s_{ac}^{``}$									
Sample ID j	Pop.ID i(A)	i(C)	ID Label	Sample Element Value y_i	Sample ID j	Pop.ID i(A)	i(C)	ID Label	Sample Element Value y_i
3	56	36	Umit-01	5	18	268	180	Aysenaz-08	0
10	176	120	Umit-05	5	24	148	96	Aysenaz-04	0
16	296	204	Umit-09	5	30	28	12	Aysenaz	0

Table F.11: List of Sample B Elements in Domains bc for Three Frames

Sample A Elements in Domain bc $s_{bc}^{``}$									
Sample ID j	Pop.ID i(B)	i(C)	ID Label	Sample Element Value y_i	Sample ID j	Pop.ID i(B)	i(C)	ID Label	Sample Element Value y_i
1	191	153	Necla-07	2	15	223	181	Nurcan-08	2
5	123	97	Nurcan-04	2	20	47	39	Gizem-01	3
6	247	207	Gizem-09	3	21	147	123	Gizem-05	3
11	23	13	Nurcan	2	26	41	27	Necla-01	2
13	91	69	Necla-03	2					

Table F.12: List of Sample B Elements in Domains abc for Three Frames

Sample A Elements in Domain abc $s_{abc}^{```}$					
Sample ID j	Pop.ID i(A)	i(B)	i(C)	ID Label	Sample Element Value y_i
2	262	225	175	Haydar-08	6
7	6	13	4	Hakan	2
14	112	100	70	Haydar-03	6
19	216	188	151	Hakan-07	2
22	182	151	127	Suheyla-06	1
25	2	1	1	Suheyla	1
27	66	63	46	Hakan-02	2
28	22	25	7	Haydar	6
29	62	51	43	Suheyla-02	1

APPENDIX G

TABLES OF MULTIPLICITY UNITS OF SAMPLES A AND B FOR DUAL FRAME

Table G.1: List of Multiplicity Units of Sample A for Dual Frame

SAMPLE A s_A						
Sample ID j	**Pop. ID i**	**ID Label**	**Sample Element Value y_i**	**Multiplicity Unit m_i**	$y_i m_i^{-1}$	$y_i^2 m_i^{-2}$
1	71	Aysegul-02	0	2	0	0
2	143	Tuna-04	7	2	3.5	12.25
3	296	Umit-09	5	1	5	25
4	170	Asli-05	1	1	1	1
5	200	Asli-06	1	1	1	1
6	161	Aysegul-05	0	2	0	0
7	248	Ozcan-08	3	2	1.5	2.25
8	131	Aysegul-04	0	2	0	0
9	83	Tuna-02	7	2	3.5	12.25
10	119	Osman-03	3	1	3	9
11	146	Umit-04	5	1	5	25
12	56	Umit-01	5	1	5	25
13	290	Asli-09	1	1	1	1
14	215	Ayse-07	5	2	2.5	6.25
15	179	Osman-05	3	1	3	9
16	197	Bo-06	1	2	0.5	0.25
17	287	Bora-09	1	2	0.5	0.25
18	33	Arif-01	2	1	2	4
19	164	Zehra-05	3	2	1.5	2.25
20	68	Ozcan-02	3	2	1.5	2.25
21	98	Ozcan-03	3	2	1.5	2.25
22	104	Zehra-03	3	2	1.5	2.25
23	95	Ayse-03	5	2	2.5	6.25
24	238	Osman-07	3	1	3	9
25	93	Arif-03	2	1	2	4
26	47	Bora-01	1	2	0.5	0.25
27	65	Ayse-02	5	2	2.5	6.25
28	101	Aysegul-03	0	2	0	0
29	206	Umit-06	5	1	5	25
30	14	Zehra	3	2	1.5	2.25
31	173	Tuna-05	7	2	3.5	12.25
32	203	Tuna-06	7	2	3.5	12.25
33	134	Zehra-04	3	2	1.5	2.25

Table G.1 (continued)

SAMPLE A s_A						
Sample ID j	**Pop. ID i**	**ID Label**	**Sample Element Value y_i**	**Multiplicity Unit m_i**	$y_i m_i^{-1}$	$y_i^2 m_i^{-2}$
34	53	Tuna-01	7	2	3.5	12.25
35	123	Arif-04	2	1	2	4
36	77	Bora-02	1	2	0.5	0.25
37	176	Umit-05	5	1	5	25
38	158	Ozcan-05	3	2	1.5	2.25
39	213	Arif-07	2	1	2	4
40	149	Osman-04	3	1	3	9
41	153	Arif-05	2	1	2	4
42	80	Asli-02	1	1	1	1
43	116	Umit-03	5	1	5	25
44	167	Bora-05	1	2	0.5	0.25
45	93	Tuna-09	7	2	3.5	12.25
46	140	Asli-04	1	1	1	1
47	8	Ozcan	3	2	1.5	2.25
48	254	Zehra-08	3	2	1.5	2.25
49	125	Ayse-04	5	2	2.5	6.25
50	251	Aysegul-08	0	2	0	0
51	155	Ayse-05	5	2	2.5	6.25
52	59	Osman-01	3	1	3	9
53	5	Ayse	5	2	2.5	6.25
54	107	Bora-03	1	2	0.5	0.25
55	299	Osman-09	3	1	3	9
56	128	Ozcan-04	3	2	1.5	2.25
57	50	Asli-01	1	1	1	1
58	3	Arif	2	1	2	4
59	74	Zehra-02	3	2	1.5	2.25
60	11	Aysegul	0	2	0	0

Table G.2: List of Multiplicity Units of Sample B for Dual Frame

SAMPLE B s_B						
Sample ID j	Pop. ID i	ID Label	Sample Element Value y_i	Multiplicity Unit m_i	$y_i m_i^{-1}$	$y_i^2 m_i^{-2}$
1	137	Tuna-05	7	2	3.5	12.25
2	74	Doga-02	4	1	4	16
3	90	Ozcan-03	3	2	1.5	2.25
4	241	Necla-09	2	1	2	4
5	166	Necla-06	2	1	2	4
6	127	Vural-05	1	2	0.5	0.25
7	171	Nurettin-06	2	2	1	1
8	128	Aul-05	0	2	0	0
9	93	Eda-03	3	1	3	9
10	2	Vural	1	2	0.5	0.25
11	115	Ozcan-04	3	2	1.5	2.25
12	53	Aysegul-02	0	2	0	0
13	62	Tuna-02	7	2	3.5	12.25
14	71	Nurettin-02	2	2	1	1
15	212	Tuna-08	7	2	3.5	12.25
16	215	Ozcan-08	3	2	1.5	2.25
17	9	Aylin	1	2	0.5	0.25
18	40	Ozcan-01	3	2	1.5	2.25
19	177	Vural-07	1	2	0.5	0.25
20	134	Aylin-05	1	2	0.5	0.25
21	102	Vural-04	1	2	0.5	0.25
22	106	Murat-04	0	1	0	0
23	96	Nurettin-03	2	2	1	1
24	109	Aylin-04	1	2	0.5	0.25
25	112	Tuna-04	7	2	3.5	12.25
26	41	Ncla-01	2	1	2	4
27	59	Aylin-02	1	2	0.5	0.25
28	6	Murat	0	1	0	0
29	246	Nurettin-09	2	2	1	1
30	37	Tuna-01	7	2	3.5	12.25
31	206	Murat-08	0	1	0	0
32	243	Eda-09	3	1	3	9
33	146	Nurettin-05	2	2	1	1
34	53	Vural-02	1	2	0.5	0.25
35	116	Necla-04	2	1	2	4
36	99	Doga-03	4	1	4	16
37	56	Murat-02	0	1	0	0
38	168	Eda-06	3	1	3	9
39	249	Doga-09	4	1	4	16
40	149	Doga-05	4	1	4	16
41	165	Ozcan-06	3	2	1.5	2.25
42	118	Eda-04	3	1	3	9
43	43	Eda-01	3	1	3	9
44	178	Aysegl-07	0	2	0	0
45	19	Doga-07	4	1	4	16
46	103	Aysegul-04	0	2	0	0
47	131	Murat-05	0	1	0	0
48	209	Aylin-08	1	2	0.5	0.25
49	91	Necla-03	2	1	2	4
50	3	Aysegul	0	2	0	0

APPENDIX H

TABLES OF MULTIPLICITY UNITS OF SAMPLES A, B AND C FOR THREE FRAMES

Table H.1: List of Multiplicity Units of Sample A for Three Frames

SAMPLE A s_A						
Sample ID j	**Pop. ID i**	**ID Label**	**Sample Element Value y_i**	**Multiplicity Unit m_i**	$y_i m_i^{-1}$	$y_i^2 m_i^{-2}$
1	71	Aysegul-02	0	2	0	0
2	143	Tuna-04	7	3	2.33	5.44
3	296	Umit-09	5	2	2.5	6.25
4	170	Asli-05	1	1	1	1
5	200	Asli-06	1	1	1	1
6	161	Aysegul-05	0	2	0	0
7	248	Ozcan-08	3	2	1.5	2.25
8	131	Aysegul-04	0	2	0	0
9	83	Tuna-02	7	3	2.33	5.44
10	119	Osman-03	3	2	1.5	2.25
11	146	Umit-04	5	2	2.5	6.25
12	56	Umit-01	5	2	2.5	6.25
13	290	Asli-09	1	1	1	1
14	215	Ayse-07	5	2	2.5	6.25
15	179	Osman-05	3	2	1.5	2.25
16	197	Bora-06	1	2	0.5	0.25
17	287	Bora-09	1	2	0.5	0.25
18	33	Arif-01	2	1	2	4
19	164	Zehra-05	3	3	1	1
20	68	Ozcan-02	3	2	1.5	2.25
21	98	Ozcan-03	3	2	1.5	2.25
22	104	Zehra-03	3	3	1	1
23	95	Ayse-03	5	2	2.5	6,25
24	238	Osman-07	3	2	1.5	2.25
25	93	Arif-03	2	1	2	4
26	47	Bora-01	1	2	0.5	0.25
27	65	Ayse-02	5	2	2.5	6.25
28	101	Aysegul-03	0	2	0	0
29	06	mit-06	5	2	2.5	6.25
30	14	Zehra	3	3	1	1
31	173	Tuna-05	7	3	2.33	5.44
32	203	Tuna-06	7	3	2.33	5.44
33	134	Zehra-04	3	3	1	1

Table H.1 (Continued)

SAMPLE A s_A						
Sample ID j	**Pop. ID i**	**ID Label**	**Sample Element Value y_i**	**Multiplicity Unit m_i**	$y_i m_i^{-1}$	$y_i^2 m_i^{-2}$
34	53	Tuna01	7	3	2.33	5.44
35	123	Arif-04	2	1	2	4
36	77	Bora-02	1	2	0.5	0.25
37	176	Umit-05	5	2	2.5	6.25
38	158	Ozcan-05	3	2	1.5	2.25
39	213	Arif-07	2	1	2	4
40	149	Osman-04	3	2	1.5	2.25
41	153	Arif-05	2	1	2	4
42	80	Asli-02	1	1	1	1
43	116	Umit-03	5	2	2.5	6.25
44	167	Boa-05	1	2	0.5	0.25
45	293	Tuna-09	7	3	2.33	5.44
46	140	Asli-04	1	1	1	1
47	8	Ozcan	3	2	1.5	2.25
48	254	Zehra-08	3	3	1	46
49	125	Ayse-04	5	2	2.5	6.25
50	251	Aysegul-08	0	2	0	0
51	155	Ayse-05	5	2	2.5	6.25
52	59	Osman-01	3	2	1.5	2.25
53	5	Ayse	5	2	2.5	6.25
54	107	Bo03	1	2	0.5	0.25
55	299	Osman-09	3	2	1.5	2.25
56	128	Ozcan-04	3	2	1.5	2.25
57	50	Asli-01	1	1	1	1
58	3	Arif	2	1	2	4
59	74	Zehra-02	3	3	1	1
60	11	Aysegul	0	2	0	0

Table H.2: List of Multiplicity Units of Sample B for Three Frames

SAMPLE B s_B						
Sample ID j	Pop. ID i	ID Label	Sample Element Value y_i	Multiplicity Unit m_i	$y_i m_i^{-1}$	$y_i^2 m_i^{-2}$
1	137	Tuna-05	7	3	2.33	5.44
2	74	Doga-02	4	1	4	16
3	90	Ozcan-03	3	2	1.5	2.25
4	241	Necla-09	2	2	1	1
5	166	Necla-06	2	2	1	1
6	127	Vural-05	1	3	0.33	0.11
7	171	Nurettin-06	2	3	0.66	0.44
8	128	Aysegul-05	0	2	0	0
9	93	Eda-03	3	1	3	9
10	2	Vural	1	3	0.33	0.11
11	15	Ozcan-04	3	2	1.5	2.25
12	53	Aysegul-02	0	2	0	0
13	62	Tuna-02	7	3	2.33	5.44
14	71	Nurettin-02	2	3	0.66	0.44
15	212	Tuna-08	7	3	2.33	5.44
16	215	Ozcan-08	3	2	1.5	2.25
17	9	Aylin	1	3	0.33	0.11
18	40	Ozcan-01	3	2	1.5	2.25
19	177	Vural-07	1	3	0.33	0.11
20	134	Aylin-05	1	3	0.33	0.11
21	102	Vural-04	1	3	0.33	0.11
22	106	Murat-4	0	2	0	0
23	96	Nurettin-03	2	3	0.66	0.44
24	109	Aylin-04	1	3	0.33	0.11
25	112	Tuna-04	7	3	2.33	5.44
26	41	Necla-01	2	2	1	1
27	59	Aylin-02	1	3	0.33	0.11
28	6	Murat	0	2	0	0
29	246	Nurettin-09	2	3	0.66	0.44
30	37	Tuna-01	7	3	2.33	5.44
31	206	Murat-08	0	2	0	0
32	243	Eda-09	3	1	3	9
33	146	Nurettin-05	2	3	0.66	0.44
34	53	Vural02	1	3	0.33	0.11
35	116	Necla-04	2	2	1	1
36	99	Doga-03	4	1	4	16
37	56	Murat-02	0	2	0	0
38	168	Eda-06	3	1	3	9
39	249	Doga-09	4	1	4	16
40	149	Doga-05	4	1	4	16
41	165	Ozcan-06	3	2	1.5	2.25
42	118	Eda-04	3	1	3	9
43	43	Eda-01	3	1	3	9
44	178	Aysegul-07	0	2	0	0
45	199	Doga-07	4	1	4	16
46	103	Aysegul-04	0	2	0	0
47	131	Mura-05	0	2	0	0
48	209	Aylin-08	1	3	0.33	0.11
49	91	Necla-03	2	2	1	1
50	3	Aysegul	0	2	0	0

Table H.3: List of Multiplicity Units of Sample C for Three Frames

SAMPLE C s_C						
Sample ID j	**Pop. ID i**	**ID Label**	**Sample Element Value** y_i	**Multiplicity Unit** m_i	$y_i m_i^{-1}$	$y_i^2 m_i^{-2}$
1	153	Necla-07	2	2	1	1
2	175	Haydar-08	6	3	2	4
3	36	Umit-01	5	2	2.5	6.25
4	178	Kaya-08	6	1	6	36
5	97	Nurcan-04	2	2	1	1
6	207	Gizem-09	3	2	1.5	2.25
7	4	Hakan	2	3	0.66	0.44
8	210	Banu9	2	1	2	4
9	146	Banu-06	2	1	2	4
10	120	Umit-05	5	2	2.5	6.25
11	13	Nurcan	2	2	1	1
12	42	Banu-01	2	1	2	4
13	69	Necla-03	2	2	1	1
14	70	Haydar-03	6	3	2	4
15	181	Nurcan-08	2	2	1	1
16	204	Umit-09	5	2	2.5	6.25
17	10	Kaya	6	1	6	36
18	180	Ayseaz-08	0	2	0	0
19	151	Hakan-07	2	3	0.66	0.44
20	39	Gizem-01	3	2	1.5	2.25
21	123	Gizem-05	3	2	1.5	2.25
22	127	Suheyla-06	1	3	0.33	0.11
23	73	Kaya-03	6	1	6	36
24	96	Aysenaz-04	0	2	0	0
25	1	Suheyla	1	3	0.33	0.11
26	27	Necla1	2	2	1	1
27	46	Hakan-02	2	3	0.66	0.44
28	7	Haydar	6	3	2	4
29	43	Suheyla-02	1	3	0.33	0.11
30	12	Aysenaz	0	2	0	0