

**Comparing approaches for combining data collected from
multiple complex surveys, adjusting for clustering and stratification**

for the degree PhD (Public Health)

Author: Loveness Nyaradzo Dziki

Student number: 04249550

Contact details:

Address: School of Health Systems and Public Health

31 Bophelo Rd

Pretoria 0001

Telephone: 012 356 3288

Cell: 084 479 1632

E-mail: loveness.dziki@up.ac.za

Supervisor: Prof Marcel de Toledo Vieira

Co-supervisor: Prof Brendan Girdler-Brown

Date: 15 December 2019

Ethics Clearance

The Research Ethics Committee, Faculty Health Sciences, University of Pretoria complies with ICH-GCP guidelines and has US Federal wide Assurance.

- FWA 00002567, Approved dd 22 May 2002 and Expires 20 Oct 2016.
- IRB 0000 2235 IORG0001762 Approved dd 22/04/2014 and Expires 22/04/2017.



UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Faculty of Health Sciences Research Ethics Committee

28/07/2016

Approval Certificate New Application

Ethics Reference No.: 270/2016

Title: Comparing approaches for combining data collected from repeated and multiple complex surveys, adjusting for clustering and stratification

Dear Mrs Loveness Dzikiti

The **New Application** as supported by documents specified in your cover letter dated 28/06/2016 for your research received on the 28/06/2016, was approved by the Faculty of Health Sciences Research Ethics Committee on its quorate meeting of 27/07/2016.

Please note the following about your ethics approval:

- Ethics Approval is valid for 2 years
- Please remember to use your protocol number (270/2016) on any documents or correspondence with the Research Ethics Committee regarding your research.
- Please note that the Research Ethics Committee may ask further questions, seek additional information, require further modification, or monitor the conduct of your research.

Ethics approval is subject to the following:

- The ethics approval is conditional on the receipt of **6 monthly written Progress Reports**, and
- The ethics approval is conditional on the research being conducted as stipulated by the details of all documents submitted to the Committee. In the event that a further need arises to change who the investigators are, the methods or any other aspect, such changes must be submitted as an Amendment for approval by the Committee.

Additional Conditions:

- For the researcher to note: The REC recommend qualifying the "approaches" in the title of the study, as not all approaches will be studied e.g. on ethical approach.

We wish you the best with your research.

Yours sincerely

*** Kindly collect your original signed approval certificate from our offices, Faculty of Health Sciences, Research Ethics Committee, Tswelopele Building, Room 4.59 / 4.60.*

Dr R Sommers; MBChB; MMed (Int); MPharm, PhD
Deputy Chairperson of the Faculty of Health Sciences Research Ethics Committee, University of Pretoria

The Faculty of Health Sciences Research Ethics Committee complies with the SA National Act 61 of 2003 as it pertains to health research and the United States Code of Federal Regulations Title 45 and 46. This committee abides by the ethical norms and principles for research, established by the Declaration of Helsinki, the South African Medical Research Council Guidelines as well as the Guidelines for Ethical Research: Principles Structures and Processes 2004 (Department of Health).

☎ 012 356 3084 ✉ fhsethics@up.ac.za 🌐 <http://www.up.ac.za/healthethics>
✉ Private Bag X323, Arcadia, 0007 - Tswelopele Building, Level 4, Room 60, Gezina, Pretoria

The Research Ethics Committee, Faculty Health Sciences, University of Pretoria complies with ICH-GCP guidelines and has US Federal wide Assurance.

- FWA 00002567, Approved dd 22 May 2002 and Expires 03/20/2022.
- IRB 0000 2235 IORG0001762 Approved dd 22/04/2014 and Expires 03/14/2020.



UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Faculty of Health Sciences Research Ethics Committee

1/03/2018

**Approval Certificate
Amendment
(to be read in conjunction with the main approval certificate)**

Ethics Reference No: 270/2016

Title: Comparing approaches for combining data collected from multiple complex surveys, adjusting for clustering and stratification

Dear Mrs Loveness Dzikiti

The **Amendment** as described in your documents specified in your cover letter dated 31/01/2018 received on 31/01/2018 was approved by the Faculty of Health Sciences Research Ethics Committee on its quorate meeting of 28/02/2018.

Please note the following about your ethics amendment:

- Please remember to use your protocol number (**270/2016**) on any documents or correspondence with the Research Ethics Committee regarding your research.
- Please note that the Research Ethics Committee may ask further questions, seek additional information, require further modification, or monitor the conduct of your research.

Ethics amendment is subject to the following:

- The ethics approval is conditional on the receipt of **6 monthly written Progress Reports**, and
- The ethics approval is conditional on the research being conducted as stipulated by the details of all documents submitted to the Committee. In the event that a further need arises to change who the investigators are, the methods or any other aspect, such changes must be submitted as an Amendment for approval by the Committee.

We wish you the best with your research.

Yours sincerely

*** Kindly collect your original signed approval certificate from our offices, Faculty of Health Sciences, Research Ethics Committee, Tswelopele Building, Level 4-60*

Dr R Sommers; MBChB; MMed (Int); MPharmD; PhD

Deputy Chairperson of the Faculty of Health Sciences Research Ethics Committee, University of Pretoria

The Faculty of Health Sciences Research Ethics Committee complies with the SA National Act 61 of 2003 as it pertains to health research and the United States Code of Federal Regulations Title 45 and 46. This committee abides by the ethical norms and principles for research, established by the Declaration of Helsinki, the South African Medical Research Council Guidelines as well as the Guidelines for Ethical Research: Principles Structures and Processes, Second Edition 2015 (Department of Health).

☎ 012 356 3084 ✉ deepeka.behari@up.ac.za / fhsethics@up.ac.za 🌐 <http://www.up.ac.za/healthethics>
✉ Private Bag X323, Arcadia, 0007 - Tswelopele Building, Level 4, Room 60 / 61, 31 Bophelo Road, Gezina, Pretoria

Acknowledgements

I would like to express my gratitude to the following people, without whom my journey would have been impossible.

To my supervisor Prof Marcel de Toledo Vieira, I am grateful for all your patience and guidance. Your sacrifices, leaving a busy schedule at your university and leaving your family to travel to Pretoria have humbled me. I have learnt more than Statistics from you.

To my co-supervisor Prof Brendan Girdler-Brown, I thank you from the bottom of my heart, for all the guidance and mentorship. I greatly appreciate your relieving me of my teaching duties so that I could focus on my studies. You taught me Statistics, English and etiquette. I will always cherish our morning debriefing sessions. I also appreciate your generous donations to the Biostatistics Development Fund.

Prof Kuku Voyi, thank you for the leadership, guidance and emotional support.

Maria L Van Wyk, for all the help in the computer lab, I thank you. You went above the call of duty to help me find the best computers in the university.

Estelle Grobler, thank you for all the assistance in the library during my literature review.

Dr Liz Wolvaardt and Dr Astrid Turner, thank you for the English editing and being there to offer support.

My husband Tarisai Brighton Dzikiti, for absorbing all the frustrations of my journey, offering support, love and guidance, you are much appreciated.

My mother, Mrs Nyembezi Roseline Chamisa, thank you for the emotional support and guidance.

Above all, I thank my children Kayla and Carlin Dzikiti, who sat with me in the office while I worked during the weekends.

Funding Acknowledgements

The University of Pretoria Visiting Professor Programme funded Prof Vieira's trips between Brazil and South Africa for onsite supervision.

The School of Health Systems and Public Health RESCOM partially paid for a more powerful laptop which I required for simulations.

List of Tables

Table 3. 1: Table Summary of the literature review where surveys were combined.	53
Table 4. 1: Combining D samples of size $n=1000$, from a population of size $N=1000\ 000$, distributed as $N(100,25)$, $\theta=99.996$.	64
Table 4. 2: Combining D samples of size $n=5000$, from a population of size $N=1000\ 000$, distributed as $N(100,25)$, $\theta=99.996$.	66
Table 4. 3: Combining D samples of size $n=10000$, from a population of size $N=1000\ 000$, distributed as $N(100,25)$, $\theta=99.996$.	67
Table 4. 4: MSE comparing weighting strategies for a normal distribution superpopulation, $D=2$ and $n=1000$ across variances and population sizes.	68
Table 4. 5 Scenario 1: Combining D samples of size $n=1000$, from a population of size $N=1000\ 000$, distributed as the skewed normal distribution $SN(100,25)$, $\theta=100.082$.	71
Table 4. 6: Scenario 2: Combining D samples of size $n=5000$, from a population of size $N=1000\ 000$, distributed as the skewed normal distribution $SN(100,25)$, $\theta=100.082$.	73
Table 4. 7: Scenario 3: Combining D samples of size $n=10\ 000$, from a population of size $N=1000\ 000$, distributed as the skewed normal distribution, $SN(100,25)$, $\theta=100.082$.	74
Table 4. 8: MSE for the skewed Normal distribution, mean=100, $N=1\ 000\ 000$ across sample sizes n and number of samples combined D.	76

Table 4. 9: Scenario 1: Combining D samples of size $n=1000$, from a population of size $N=1000\ 000$, distributed as the t-distribution with three degrees of freedom, $t(3)(100,25)$, $\theta=99.998$	78
Table 4. 10: Scenario 2: Combining D samples of size $n=5\ 000$, from a population of size $N=1000\ 000$, distributed as the t-distribution with three degrees of freedom, $t(3)(100,25)$, $\theta=99.998$	80
Table 4. 11: Scenario 3: Combining D samples of size $n=10\ 000$, from a population of size $N=1000\ 000$, distributed as the t-distribution with three degrees of freedom, $t(3)(100,25)$, $\theta=99.998$	81
Table 4. 12: MSE for the t distribution with three degrees of freedom, mean=100, $N=1\ 000\ 000$ across sample sizes n and number of samples combined D.	82
Table 4. 13: Scenario 1: Combining D samples of size $n=1000$, from a population of size $N=1000\ 000$, distributed as the skewed t-distribution with three degrees of freedom, $at(3)(100,25)$, $\theta=99.955$	85
Table 4. 14: Scenario 2: Combining D samples of size $n=5\ 000$, from a population of size $N=1000\ 000$, distributed as the skewed t-distribution with three degrees of freedom, $at(3)(100,25)$, $\theta=99.955$	86
Table 4. 15: Scenario 3: Combining D samples of size $n=10\ 000$, from a population of size $N=1000\ 000$, distributed as the skewed t-distribution with three degrees of freedom, $at(3)(100,25)$, $\theta=99.955$	88
Table 4. 16: MSE for the skewed t distribution with three degrees of freedom, mean=100, $N=1\ 000\ 000$ across sample sizes n and number of samples combined D.....	89
Table 4. 17 : Misspecification effects for all the distributions and sample sizes considered under stratified sampling.....	92

Table 4. 18 : Scenario 1: Combining D samples of size $n=1000$, from a population of size $N=1000\ 000$, distributed as the normal distribution $N(100,25)$, $\theta=99.999$	94
Table 4. 19: Scenario 2: Combining D samples of size $n=5000$, from a population of size $N=1000\ 000$, distributed as the normal distribution $N(100,25)$, $\theta=99.999$	96
Table 4. 20: Scenario 3: Combining D samples of size $n=10000$, from a population of size $N=1000\ 000$, distributed as the normal distribution $N(100,25)$, $\theta=99.999$	98
Table 4. 21: Scenario 1: Comparing weighting strategies for a finite population generated using the skewed normal distribution, $SN(100,25)$, $N=1000\ 000$ and $n=1000$ $\theta=100.056$	101
Table 4. 22: Scenario 2: Comparing weighting strategies for a finite population generated using the skewed normal distribution, $SN(100,25)$, $N=1000\ 000$ and $n=5000$ $\theta=100.056$	103
Table 4. 23: Scenario 3: Comparing weighting strategies for a finite population generated using the skewed normal distribution, $SN(100,25)$, $N=1000\ 000$ and $n=10000$ $\theta=100.056$	105
Table 4. 24: Scenario 1: Comparing weighting strategies for a finite population generated using the t distribution with three degrees of freedom, $t(3)(100,25)$, $N=1000\ 000$ and $n=1000$ $\theta=99.998$	108
Table 4. 25: Scenario 2: Comparing weighting strategies for a finite population generated using the t distribution with three degrees of freedom, $t(3)(100,25)$, $N=1000\ 000$ and $n=5000$ $\theta=99.998$	110
Table 4. 26: Scenario 3: Comparing weighting strategies for a finite population generated using the t distribution with three degrees of freedom, $t(3)(100,25)$, $N=1000\ 000$ and $n=10000$ $\theta=99.998$	112

Table 4. 27: Scenario 1: Comparing weighting strategies for a finite population generated using a skewed t distribution with three degrees of freedom, $t(3)(100,25)$, $N=1000\ 000$ and $n=1000$ $\theta=99.970$	115
Table 4. 28: Scenario 2: Comparing weighting strategies for a finite population generated using a skewed t distribution with three degrees of freedom, $t(3)(100,25)$, $N=1000\ 000$ and $n=5000$ $\theta=99.970$	117
Table 4. 29: Scenario 3: Comparing weighting strategies for a finite population generated using a skewed t distribution with three degrees of freedom, $t(3)(100,25)$, $N=1000\ 000$ and $n=10000$ $\theta=99.970$	119
Table 4. 30: Misspecification effects for all the distributions and sample sizes considered under cluster sampling.....	123
Table 4. 31: Scenario 1: Combining D samples of size $n=1000$, from a population of size $N=1000\ 000$, distributed as the normal distribution $N(100,25)$, $\theta=100.031$	125
Table 4. 32: Scenario 2: Combining D samples of size $n=5000$, from a population of size $N=1000\ 000$, distributed as the normal distribution $N(100,25)$, $\theta=100.031$	127
Table 4. 33: Scenario 3: Combining D samples of size $n=10000$, from a population of size $N=1000\ 000$, distributed as the normal distribution $N(100,25)$, $\theta=100.031$	129
Table 4. 34: Scenario 1: Combining D samples of size $n=1000$, from a population of size $N=1000\ 000$, distributed as the skewed normal distribution $SN(100,25)$, $\theta=99.914$	132
Table 4. 35: Scenario 2: Combining D samples of size $n=5000$, from a population of size $N=1000\ 000$, distributed as the skewed normal distribution $SN(100,25)$, $\theta=99.914$	134
Table 4. 36: Scenario 3: Combining D samples of size $n=10000$, from a population of size $N=1000\ 000$, distributed as the skewed normal distribution $SN(100,25)$, $\theta=99.914$	136

Table 4. 37: Scenario 1: Combining D samples of size $n=1000$, from a population of size $N=1000\ 000$, distributed as the t-distribution with three degrees of freedom $t(3)(100,25)$, $\theta=100.010$	139
Table 4. 38: Scenario 2: Combining D samples of size $n=5000$, from a population of size $N=1000\ 000$, distributed as the t- distribution with three degrees of freedom $t(3)(100,25)$, $\theta=100.010$	141
Table 4. 39: Scenario 3: Combining D samples of size $n=10000$, from a population of size $N=1000\ 000$, distributed as the t- distribution with three degrees of freedom $t(3)(100,25)$, $\theta=100.010$	143
Table 4. 40: Scenario 1: Combining D samples of size $n=1000$, from a population of size $N=1000\ 000$, distributed as the skewed t-distribution with three degrees of freedom $at(3)(100,25)$, $\theta=99.940$	146
Table 4. 41: Scenario 2: Combining D samples of size $n=5000$, from a population of size $N=1000\ 000$, distributed as the skewed t-distribution with three degrees of freedom $at(3) (100,25)$, $\theta=99.940$	148
Table 4. 42: Scenario 3: Combining D samples of size $n=10000$, from a population of size $N=1000\ 000$, distributed as the skewed t-distribution with three degrees of freedom $at(3) (100,25)$, $\theta=99.940$	150
Table 6. 1: Comparing Relative differences across weighting strategies and number of samples combined, for each variance estimator for the normal distribution under simple random sampling, $n=1000$	162
Table 6. 2: Comparing Relative differences across weighting strategies and number of samples combined, for each variance estimator for the normal distribution under simple random sampling, $n=5000$	164

Table 6. 3: Comparing Relative differences across weighting strategies and number of samples combined, for each variance estimator for the normal distribution under simple random sampling, $n=10000$	166
Table 6. 4: Comparing standard errors across sampling strategies, weighting strategies and number of samples combined, for a finite population distributed the normal distribution $N(100,25)$, $n=10000$	169
Table 6. 5: Comparing standard errors across sampling strategies, weighting strategies and number of samples combined, for the finite population distributed as the skewed normal distribution, $SN(100,25)$, $n=10000$	171
Table 6. 6: Comparing standard errors across sampling strategies, weighting strategies and number of samples combined, for the finite population distributed as the t distribution, 3d.f, $t(100,25)$, $n=10000$	173
Table 6. 7: Comparing standard errors across sampling strategies, weighting strategies and number of samples combined, for the finite population distributed as the skewed t distribution, 3d.f, $at(3)(100,25)$, $n=10000$	175
Table 7. 1: Combined estimates of mean age of the household head and their confidence intervals under simple random sampling.....	180
Table 7. 2: Comparison of standard errors from one sample to standard errors resulting from combined samples.....	183
Table 7. 3: Combined estimates of mean age of household head and confidence intervals under stratified random sampling	185
Table 7. 4: Comparison of standard errors from one sample to standard errors resulting from combined samples under stratified random sampling.....	188

List of Figures

Figure 1: Schematic Diagram of a Meta-analysis.....	15
Figure 2: Combining surveys: Scenario 1.....	16
Figure 3: Combining Surveys: Scenario 2.....	17
Figure 4: Combining Surveys: Scenario 3.....	18
Figure 5: Separate Approach: Adapted from G. Roberts and D. Binder 2009.....	28
Figure 6: Pooled Analysis: Adapted from G. Roberts and D. Binder 2009.....	29
Figure 7: Convergence of combined estimates as number of samples increases....	58
Figure 8: Comparing weighting strategies for combining surveys under simple random sampling.....	60
Figure 9: Evaluating variance estimators	159
Figure 10: Serial computing	160
Figure 11: Parallel computing.....	160

Abstract

Even though there is substantial literature on studies which pool survey data, it is still not clear which are the most efficient methodologies for pooling data from different surveys. For example, it is important to know whether the surveys involved should be given equal importance in the calculation of the combined statistics or not. If they are not given equal importance, then it should be clear how they should be weighted and why. In this research project, alternative methods used to combine survey data were evaluated and new methods proposed.

A literature review of methods that are currently being used in combining repeated and multiple surveys was presented. New methods were proposed or adapted from meta-analysis methodology to try and improve the calculation of weights and precision measures when multiple surveys are combined. Different variance estimators for the proposed point estimators were evaluated through simulation. Only the separate approach was considered in this study. Simple random samples and complex samples were drawn from simulated finite population data and used to evaluate current and proposed methods of combining surveys. Simple super-population models were used to simulate finite population data. The South African Community Survey of 2016 and the General Household Survey of 2016 were used to simulate finite populations which were then used for evaluating the different methods of combining simple random sampling and stratified surveys respectively.

Our results suggest that the choice of weighting method when combining surveys should depend on the super-population model assumed to have generated the finite population. The sample size used appeared to influence the choice of the method used to combine surveys, but the variance of the super-population did not influence the choice. Under simple random sampling, the strength of the skewness and kurtosis also appeared to affect the performance of the weighting strategies. Weighting by the inverse of the sample size, the inverse of variance

and the inverse of the coefficient of variation appeared to work for most super-population models. Combining samples appeared to yield better estimates with lower mean square errors compared to single sample estimates. The number of samples combined appeared not to influence the choice of weighting strategy although the mean square errors decreased with increased number of samples combined.

Under simple random sampling, the meta-analysis variance estimator appeared to work the best with the inverse of variance weighting method as expected.

Table of Contents

Ethics clearance.....	ii
Ethics Clearance	ii
Acknowledgements.....	iv
Funding Acknowledgements.....	iv
List of Tables	v
List of Figures	xi
Abstract.....	xii
Table of Contents.....	xiv
Chapter 1.....	1
Introduction	1
1.1 Health surveys.....	1
1.2 Sampling in health surveys	4
1.3 Meta-analysis and combining of surveys.....	13
1.4 Weighting implications in combining surveys and meta-analysis	19
1.5 Motivation and aims of the thesis	24
1.6 Outline of the thesis.....	25
Chapter 2.....	27
Methods for combining surveys and applications – A review	27
2.1 Separate and pooled approaches	27
2.2 Analysis of survey data in the context of combining surveys	32
2.3 Application in health and social surveys.....	35
Chapter 3.....	38
Methods for combining surveys: Point estimators under the separate approach.....	38
3.1 A review	38
3.2 Contributions under simple random sampling	44
3.3 Criteria for evaluating estimators	47
3.3 Contributions under complex sampling: stratification and clustering	50
Chapter 4.....	57
Simulation study 1	57
4.1 Simulation set up	58
4.2 Simulation results under simple random sampling	62
4.3 Simulation results under stratified sampling.....	91

4.4	Simulation results under two-stage simple cluster sampling	122
Chapter 5.....		152
Methods for combining surveys: Variance estimators under the separate approach.....		152
5.1	Traditional methods.....	152
5.2	Replication methods	155
Chapter 6.....		157
Simulation study 2 (Variance estimators in chapter 5).....		157
6.1	Simulation set up	158
6.2	Simulation results for variance estimators under simple random sampling.....	161
6.3	Supplementary Analysis.....	168
Chapter 7 (Application).....		177
7.1	Community Survey 2016.....	177
7.2	General Household Survey 2016.....	178
7.3	Simple random sampling application.....	179
7.4	Stratified random sampling application.....	184
7.5	Conclusion.....	190
Chapter 8.....		191
Conclusion.....		191
References		194
Appendix A1		201
Comparing weighting strategies: Simple Random Sampling: Normal distribution		201
Appendix A2		226
Comparing weighting strategies: Stratified Random Sampling: Normal distribution		226
Appendix A3		233
Comparing weighting strategies: Cluster Random Sampling: Normal distribution.....		233
Appendix B1		240
Comparing weighting strategies: Stratified Random Sampling: Skewed normal distribution		240
Appendix B2		247
Comparing weighting strategies: Cluster Random Sampling: Skewed normal distribution.....		247
Appendix C1		253
Comparing weighting strategies: Stratified Random Sampling: t-distribution, 3 d.f.....		253
Appendix C2		259
Comparing weighting strategies: Cluster Random Sampling: t-distribution, 3 d.f.....		259
Appendix D1.....		265
Comparing weighting strategies: Stratified Random Sampling: Skewed t-distribution, 3 d.f.....		265

Appendix D2.....	271
Comparing weighting strategies: Cluster Random Sampling: Skewed t-distribution, 3 d.f.....	271
Appendix E	277
R code used in simulations for simple random sampling, point estimators.....	277
E1 Normal Distribution	277
E2 T Distribution, 3 degrees of freedom.....	283
E3 Skewed Normal Distribution	290
E4 Skewed T Distribution, 3 degrees of freedom	296
Appendix F	302
R code used in simulations for simple random sampling, variance estimators	302
F1 Normal Distribution.....	302
Appendix G.....	308
R code used in simulations for stratified random sampling, point estimators.....	308
G1 Normal Distribution	308
G2 T Distribution, 3 degrees of freedom	317
G3 Skewed Normal Distribution	326
G4 Skewed T Distribution, 3 degrees of freedom.....	335
Appendix H.....	344
R code used in simulations for cluster random sampling, point estimators	344
H1 Normal Distribution	344
H2 T Distribution, 3 degrees of freedom	377
H3 Skewed Normal Distribution	410
H4 Skewed T Distribution, 3 degrees of freedom.....	443

Chapter 1

Introduction

Different surveys usually adopt alternative sampling designs and consequently require the consideration of different sets of sampling weights. When surveys are combined for calculation of estimates, questions arise on how to handle the different survey weights and sampling designs for producing combined point and variance estimators. Combined estimates may be essential in the formulation of national, regional or global health and social policies. This thesis discusses some of the ways in which surveys can be weighted when calculating combined estimates. Chapter 1 presents a brief introduction to types of surveys and the use of surveys in the health and social sectors.

1.1 Health surveys

This section gives a brief introduction to types of surveys commonly used in health-related settings. Sample surveys can generally be categorized into cross-sectional and longitudinal surveys. Cross-sectional surveys are carried out at one point in time. They are usually conducted to estimate the prevalence of the outcome of interest for a given population, commonly for the purposes of public health planning. They cannot be used to measure incidence rates or new occurrence rates for the outcome of interest. Data can be collected on individual characteristics, including exposure to risk factors, alongside information about outcomes. In this way, cross-sectional studies provide a 'snapshot' of the outcome and the characteristics associated with it, at a specific point in time^[1]. Cross sectional studies cannot in general be used to study causal relationships because the outcome of interest and possible risk factors are observed at the same time; they may, however, be used to evaluate relationships between exposures and outcomes. A possible exception is when exposures are documented for lagged explanatory variables; in such a case, it is sometimes clear that the exposure occurred before the outcome.^[2]

In longitudinal surveys, a sample is selected and followed over time. This enables measurement of incidence and progression of disease over time. Changes can be observed both at group and individual level^[3]. Possible causal relationships can be evaluated using data from longitudinal surveys and suitable methodology. When conducting a longitudinal survey, measurements can be made on several occasions, on the same sample or panel over time for the duration of the survey. This is referred to as a fixed panel longitudinal survey. Alternatively, subgroups of the panel may be removed from the survey sample at a stipulated maximum time-in-sample, in which case, we have a rotating panel survey. The advantages of using a fixed panel survey design are:^[3-5]

1. Reduced sampling variation in the measurement of change.
2. They allow change to be measured at the level of the individual respondent.
3. Respondents can be trained to perform relatively complex tasks as, for example a consumption diary.
4. A wide range of data can be collected compared to a cross-sectional survey.
5. They may be less expensive compared to repeated cross-sectional surveys because the first time a unit is included there are higher costs.^[4]

However, fixed panel longitudinal surveys have some disadvantages too. For example, the initial co-operation rate is lower than that for single contact (cross-sectional) surveys. They are also prone to attrition over time as some respondents become lost to follow up for various reasons. Response and behavioural conditioning can also be problematic. For ethical reasons, counselling may be part of a survey and that may result in respondents behaving in a way that is not representative of the population at large. A study examining the effects of irresponsible alcohol consumption on HIV incidence among youths would be expected to provide counselling on alcohol abuse at the onset and during the course of the study for ethical reasons. As a result, some participants may start changing their consumption patterns. One way of dealing with these problems is the use of rotating panels. The extent to which rotating panels correct demographic imbalances due to attrition depends on the differentials in attrition in the demographic subgroups. Similarly, the extent to which fatigue is addressed by rotation

of panels depends on the rate at which fatigue occurs after joining the panel. Rotating panels also have the disadvantage of loss of sample continuity resulting in an increase in the sampling error in estimates of change. The loss of sample continuity also decreases the sample size available for analysis of change at the level of the individual respondents.^[3-5]

When longitudinal surveys cannot be carried out, repeated cross-sectional surveys may be carried out to give pseudo-longitudinal surveys, where the individuals included in the survey are either selected from a unique sampling frame or from multiple sampling frames. The dictionary of epidemiology defines a pseudo-longitudinal survey as a sequence of samples resembling a regularly measured cohort with the exception that it consists of a different sample at each occasion of sampling^[6]. In a repeated cross-sectional survey, an independent sample is selected on each occasion, and so there may be essentially no overlap in the samples between time periods. Valid estimates of changes at the population level can be calculated from independent samples. Let θ be a parameter of interest. If $\hat{\theta}_t$ is an estimator for θ at time t and $\hat{\theta}_{t-s}$ and estimator for θ at time $t - s$, then the change or movement between the two time periods can be estimated using $\hat{\theta}_t - \hat{\theta}_{t-s}$ ^[4]. Change over time is required to measure the effect of some forces known to have acted on the population. For example, estimating the effect of an HIV intervention program on the incidence rate of HIV in a population could be a measurement of interest.

The main advantage of repeated cross-sectional surveys over longitudinal surveys is that they tend to give more precise estimates of prevalence. This is because they generally have larger sample sizes and are more likely to cover the whole population. Many major longitudinal studies follow only a particular group of people. Moreover, in longitudinal studies, sampling is usually only done at the beginning of the study making long term longitudinal studies less sensitive to changes in the constituents of the population, especially for minority groups^[7,8]. The population may change due to people moving into the area or out of the area.^[7] The advantage of using longitudinal studies is that they give better estimates of change. A longitudinal study of the same sample size will give better estimates of change compared to repeated cross-sectional surveys

because of the reduction in variance of the change estimator brought by using the same individuals.^[9]

When analysing survey data, be it cross-sectional or longitudinal, some adjustments have to be made to account for the way the data were collected. For example, when cluster sampling or stratified random sampling is used, the assumption of equal probability of inclusion and independence among observations in the sample do not hold. Sampling designs that involve stratification, clustering, unequal selection probabilities and, consequently, survey weighting, are usually called complex sampling designs. Statistical techniques developed for simple random sampling will not be appropriate under complex sampling. The use of weights in the analysis will also be required. Over the years, some statistical software with built-in packages for analysing surveys have been developed (e.g. STATA svy commands, SAS Survey Procedures, SUDAAN, R Survey Package, SPSS Complex samples module among others.). These software packages are generally equipped to analyse single and cross-sectional complex sampling surveys, but still have mostly not implemented tools for analysing longitudinal surveys fully allowing for the complex sampling design.

In the following sections of this chapter, some health-related surveys, meta-analysis and the importance of combining surveys will be discussed. A distinction between classical meta-analysis and combining surveys is also given. Different scenarios where surveys may be combined are discussed together with the weighting implications of combining surveys.

1.2 Sampling in health surveys

To gain information about large populations, for example the health status of a population, a sample is usually collected and analysed. This saves on costs, time and human resources. For the sample to give accurate information, it should be selected using a probability sampling design. By using a probability sampling, a sample statistician avoids subjective judgements (bias) in sample selection. Probability

sampling and probability theory ensure proper statistical inference from the sample to the population and possibility to calculate quality indicators such as standard errors and other precision measures. Probability sample surveys may also suffer from coverage and response bias but there are widely accepted ways of dealing with the bias problem.^[10]

Survey research is important to health professionals and health policy makers. It provides information about the incidence rates and prevalence of diseases. Additionally, health surveys can be used to study the occurrence of healthy and unhealthy behaviours, exposures to potential risk factors, dietary intake, physiologic measures of the population, and costs and utilization of health services.^[11,12] The Health and Retirement (HRS) study in the United States of America was designed to follow a probability sample of age-eligible individuals and their life partners as they transitioned from active working to retirement, measuring physical and mental health among other things. This study has been used to address important questions about the challenges and opportunities of aging.^[13] The National Health and Nutrition Examination Survey (NHANES) studied the prevalence of major diseases in the United States of America. Change in prevalence over time and risk factors associated with the diseases were also studied.^[13]

The Demographic and Health Surveys (DHS) have been carried out in more than 90 countries including South Africa. Data have been collected, analysed, and disseminated on population, health, HIV, and nutrition through more than 300 surveys. The aim of the DHS is to provide a better understanding of the health status of the population in the different countries. Key topics include child health, reproductive health, adult health and nutrition. The information collected assists the various departments of health to plan and prioritise health programmes and service delivery. They also provide an opportunity for household members to understand their individual health statuses. The DHS are also useful as they provide potential baseline measures against which the impact of an intervention may be measured, including the use of propensity scores to identify controls who might not have been exposed to the intervention. They are a valuable

resource for monitoring and evaluation in countries like Ghana where they are repeated every two years without fail.^[14]

Statistics South Africa conducts the Community Survey (CS) to provide data between censuses. The Community Survey of 2016 aimed to provide indicators informing implementation, monitoring and evaluation of developmental programmes at local municipality levels. Demographic factors including fertility and mortality were measured. Access to services and facilities and the extent of poverty in households were also measured. We used this survey in the application of our proposed methods together with the South African General Household survey of 2016.

The General Household Survey (GHS) is conducted annually by Statistics South Africa. The survey measures the progress of development in the country, the performance of government programmes together with the quality of service delivery in important service sectors. These important service delivery sectors include the health and education sectors. Surveys inform important decision making and policy formulation. It is important that proper sampling procedures are followed in conducting surveys. Probability sampling allows the sample survey to be representative of the population and ensures that all members of the population have a known non-zero probability of inclusion in the sample survey and, therefore, allows the estimation of precision measures.^[4] Examples of probability sampling methods include simple random sampling, stratified random sampling, probability of selection proportional to size, cluster random sampling and multi-stage sampling.

Simple random sampling is conceptually the simplest probability sampling technique. The procedure ensures that every sample of the same size has the same probability of being selected, and consequently each element in the population has an equal probability of being selected. Most statistical tests and statistical inference methods were developed for simple random sampling with replacement data, assuming that the observations are independent and identically distributed.^[13,15,16]

However, it is usually more statistically efficient to perform stratified random sampling when the population has subgroups which may differ with regards to the measurements of interest, or perform systematic random sampling if a complete enumeration of the population is not known before the commencement of sample selection.^[15] Furthermore, simple random sampling may be cumbersome for surveys covering large geographic areas. In such situations, cluster sampling may be more attractive in terms of costs and fieldwork logistics. In many large-scale surveys, a large number of variables is collected from a large number of individuals which allows for many hypotheses to be addressed. Some of these hypotheses are formulated long after the surveys have been collected.^[12] The survey designs used are usually complex and multistage with differential sampling probabilities for the survey respondents. Clustering and stratification are frequently used. Clustering may be used when the available sampling frame is of clusters, and not individuals. For example the electoral register or census data may list households, as people are usually mobile.^[17] It may also be preferable to randomly select some clusters and then interview all the individuals or randomly selected individuals from those clusters to reduce the cost and time of travelling. Additionally, the population is concentrated in natural clusters, for example schools or hospitals, making it convenient to select these instead of using simple random sampling.^[13] However, individuals in a cluster usually have some common features leading to intra-cluster correlation.

In cluster sampling, the population is divided into clusters of a certain size, and the clusters may be selected by simple random sampling or other probability sampling methods. In practise, clusters may be of different sizes and in that situation if sample sizes within clusters are unequal, then individuals in larger clusters have a smaller probability of selection into the sample if one uses simple random sampling to select the clusters. In that case, sampling clusters with probability proportional to size may be adopted for adjusting the probability of selection such that the individuals have a similar probability of being selected into the sample.^[18,19]

When conducting a survey, there may be a rare sub-group that may play a big influence in the estimation of overall means or other parameters. The effects of the rare subgroup

may be easily missed, leading to a biased estimate for the overall mean (e.g. commercial sex workers or IV drug users). This is not very notable when one wants to estimate a median for an ordinal variable. There will usually be oversampling unless one uses a quota system with numbers in sample proportional to estimated numbers in the population. On the other hand, one may specifically wish to obtain a precise estimate of the sub-group parameter to make inference about the sub-group parameter. In this case one could only do this by stratification along with oversampling so that there are sufficient numbers in each sub-group (not just proportionally representative numbers). We may consider, for example, a desert district with a low population relative to other districts. The population is divided into strata and a relatively large number of survey sampling units are selected from each stratum.^[17] Large scale health surveys often make use of both clustering and stratification. Furthermore, there might be special interest in some minority groups who might not be well represented in a simple random sample. To ensure inference can be done for these minority groups, they are usually oversampled in surveys. Weights will be required in these situations.^{[13][17]} The HRS oversampled African Americans and Hispanic adults to increase their presence in the sample and hence allow inference on these particular racial groups. Weights were then used to correct for the influence of the disproportionate representation of the African Americans and Hispanic people when population estimates were calculated.^[13]

Weights are also used to mitigate the effects of differential non-response across the sample. In most surveys, some individuals selected into the sample do not consent to be interviewed and the non-response rate varies across the sample. In the case of some household surveys, only the people found at home when the interviewer arrives are interviewed. The elderly, people who do not work and children are more likely to be at home compared to the economically active age-group. This may cause bias in the estimation of population estimates. Weights have to be introduced to reduce bias in the estimates from the survey.^[13]

Other sampling techniques which are convenient under different scenarios are available but will not be discussed here. The use of other sampling methods other than simple random sampling can lead to either an increase or a decrease in the precision of

estimates. Stratification usually increases precision; while clustering generally decreases precision. Individuals in a cluster are generally more homogeneous than those in a simple random sample. This homogeneity is commonly known as intra-cluster or intra-class correlation. Homogeneity within a cluster may be caused by selective factors in the cluster, joint exposure to similar influences or effects of mutual interaction. It reduces the effective sample size and, hence, should increase the variance of the estimators. When estimating the variance of the estimators, disregarding intra-cluster correlation generally results in the underestimation of the variance and overestimation of the precision, unless the effective sample size is considered to perform the estimations.

One measure of the efficiency of a sampling design is the design effect. The design effect denoted def can be defined as the ratio of actual variance for a given sampling design to the variance of a simple random sample of the same size. That is, the ratio of the variance of an estimate obtained from a specific sampling design to the variance of the same estimate if a simple random sample had been selected. The term design effect was introduced by Kish in 1965.^[18,20] He calculated design effect as

$$def_{Kish}(\hat{\theta}) = \frac{Var_{true}(\hat{\theta})}{Var_{srs}(\hat{\theta})}$$

where $Var_{true}(\hat{\theta})$ is the *true* variance of $\hat{\theta}$ that considers the *true* sampling scheme used for the selection of the sample, and $Var_{srs}(\hat{\theta})$ is the hypothetical variance of $\hat{\theta}$ when it is assumed that the sample was selected by simple random sampling with replacement.^[20,21] It gives an idea of how much precision is gained or lost by the use of a particular complex sampling design^[18,22].

Under single stage cluster sampling and under simple random sampling of clusters, the design effect may be estimated using the following formula:

$$def = 1 + \rho(c - 1), \quad (1.1)$$

where ρ is the intra-cluster correlation and c is the average cluster size of the clusters. If cluster sizes were equal, equation 1.1 would give the exact value of the design effect. As a result, for such designs the intra-cluster correlation, may be calculated as

$$\rho = \frac{deff - 1}{c - 1} \quad (1.2)$$

When the intra-cluster correlation is positive, the design effect is larger than one implying that the use of a cluster sample yields a larger variance compared to a simple random sample of the same size. When it is equal to zero, the design effect is equal to one and the variable is distributed completely at random among the clusters and the use of a cluster sample results in the same variance as would a simple random sample of the same size^[18,22].

The original design effect measure presents some disadvantages for analysis purposes. It has its applicability related mainly to the evaluation of precision gains or losses, when comparing alternative sampling schemes before conducting the survey. The measure $deff_{Kish}(\hat{\theta})$ would not be appropriate for analysis under one single design, according to Skinner, Holt and Smith.^[20,23]

Skinner proposed the *misspecification effect* (*mef*) which is more suitable for analytic inference than the original design effect. The *mef* is aimed at measuring the effects of incorrect specification of both the sampling scheme and the considered model.^[20,24] The *mef* is defined in a similar way as the *deff* with the same numerator, but with the denominator consisting of the expectation of a variance estimator which ignores the complex design.

Let^[20,24] $var_0(\hat{\theta}) = var_{iid}(\hat{\theta})$ be a consistent estimator of the variance of $\hat{\theta}$, when we assume that the observations are independent and identically distributed (iid). The

effect of the complex sampling scheme on $var_0(\hat{\theta})$ can be evaluated if we examine its distribution. Skinner, Holt and Smith define the misspecification effect as

$$meff(\hat{\theta}, var_0) = \frac{Var_{true}(\hat{\theta})}{E_{true}[var_0(\hat{\theta})]}$$

which gives a measure of how much $var_0(\hat{\theta})$ over- or underestimates $Var_{true}(\hat{\theta})$, and may be estimated, for example by

$$\widehat{meff}(\hat{\theta}, var_0) = \frac{var_{true}(\hat{\theta})}{var_0(\hat{\theta})}$$

The *meff* defined above is thus a measure of relative bias of the variance estimator.

The misspecification effect is interpreted as follows:

$meff(\hat{\theta}, var_0)$	$Bias[var_0(\hat{\theta})]$	Interpretation
< 1	> 0	Overestimation of $Var_{true}(\hat{\theta})$
$= 1$	$= 0$	Correct estimation of $Var_{true}(\hat{\theta})$
> 1	< 0	Underestimation of $Var_{true}(\hat{\theta})$

The $meff(\hat{\theta}, var_0)$ is theoretically different from $deff_{Kish}(\hat{\theta})$ because it depends on the arguments $\hat{\theta}$ and var_0 whereas $deff_{Kish}(\hat{\theta})$ only depends on $\hat{\theta}$. However, their respective estimators, calculated from a unique sample may coincide in practice.^[20,23]

Section 1.3 below discusses meta-analysis and the need for combining surveys.

1.3 Meta-analysis and combining of surveys

It is sometimes necessary to combine and analyse two or more cross-sectional surveys to increase the sample size of sub-populations of interest, or to improve coverage given by individual surveys. Large surveys may not have enough sample size for some subgroups of interest to make any meaningful inference. For example, if one were interested in the prevalence of HIV among people living with albinism in South Africa, the prevalence of albinism in South Africa is very low. Hence several surveys would need to be combined before enough sample size to make meaningful estimation is attained. Combining surveys may increase the available sample size and allow inference to be done. Merkouris investigated combining regression estimators from multiple surveys for this purpose.^[25]

Combining surveys can provide an opportunity to investigate a broader range of research questions than the original individual surveys. It is important to make sure that similar variables in the surveys to be combined actually measure the same characteristic, considering the item wording in questionnaires and collection method on which the surveys were based.^[26,27] Telephone interviews, self-administered questionnaires and face to face interviews have different non-sampling errors that can be introduced in a survey. Language and privacy issues also need to be considered. The same item wording may have different interpretations across time periods and cultures.^[28] Kish emphasized that the definition of concepts, variables and populations; methods of measurement and the substantive analysis should be similar for comparisons or combining of estimates to be carried out. He defines substantive analysis as the part of the analysis that is informed by a specific field or science to differentiate it from statistical analysis.^[28] A key assumption in combining of surveys is that the target populations of the surveys are equivalent.^[29] Tests of homogeneity should be used where possible. Sample designs and sample sizes do not need to be

similar. In the case of multi-national surveys, different countries will have different resources at their disposal.^[28]

Combining surveys that have been designed separately has been likened to meta-analysis.^[27] In this situation, meta-analysis is defined as the statistical analysis of results from two or more independent studies for the purpose of integrating findings.^[27,29,30] Alternatively, meta-analysis can be defined as the analysis of results from a series of studies.^[31] It aims to improve the estimate of an effect size through pooling information across studies.

In this section, a distinction between combining surveys and classical meta-analysis is made. Rao et al. define meta-analysis as a method of analysing individual data combined from two or more studies or analysis of summary measures obtained from two or more reports.^[32] Kish points out a few dissimilarities between meta-analysis and combining surveys. These dissimilarities are linked to the differences in the way experimental studies and surveys are conducted. Combining experimental studies emphasizes experimental control through randomization of variables over subjects. On the other hand, combining surveys is based on probability sampling with randomized selection of subjects, not variables. In experimental studies, unexplained heterogeneity is mainly due to the underlying model whereas in surveys, additional heterogeneity is introduced by sampling, especially in the presence of clustering. Additionally, the population base for combining surveys is specified whereas the population base for combining experiments is usually not specified and cannot be specified.^[33] Fox refers to the meta-analysis of survey data and emphasizes that even though combining surveys appears to be a straightforward extension of meta-analysis, comprehensible methodology that has clearly defined limitations is required to avoid erroneously combining survey data.^[29] For example, some methodology could work well in the

presence of stratification but not work well in the presence of clustering. The schematic diagram in Figure 1 illustrates how a meta-analysis is performed.

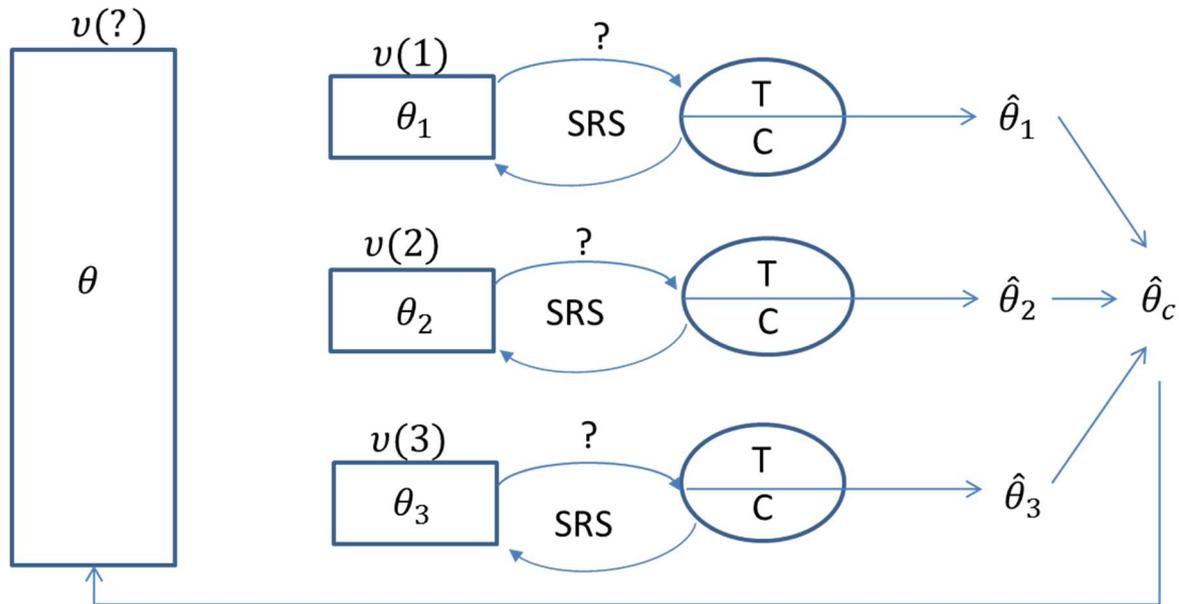


Figure 1: Schematic Diagram of a Meta-analysis

In this illustration, a special case with three Randomized Control Trials has been considered. More studies could be used, including some with different study designs. A simple random sample is taken from population $v(1)$ with parameter θ_1 . The sample is selected and observed to calculate an estimate $\hat{\theta}_1$, of θ_1 . The same procedure is used to obtain the estimates $\hat{\theta}_2$ and $\hat{\theta}_3$ of θ_2 and θ_3 , for populations $v(2)$ and $v(3)$. The three estimates are then combined to obtain $\hat{\theta}_c$ which is then used to estimate θ which is a parameter of some undefined population $v(?)$ that includes all three populations studied.

In combining surveys, we consider three alternative scenarios. The first scenario discussed here considers the case when different survey samples are selected independently from the same population as illustrated in schematic Figure 2. The surveys are sampled through a sampling design $P(S_d), d = 1, 2, \dots, D$, where D is the number of selected samples. Under this setup, the sampling design is the same for all samples, for simplicity

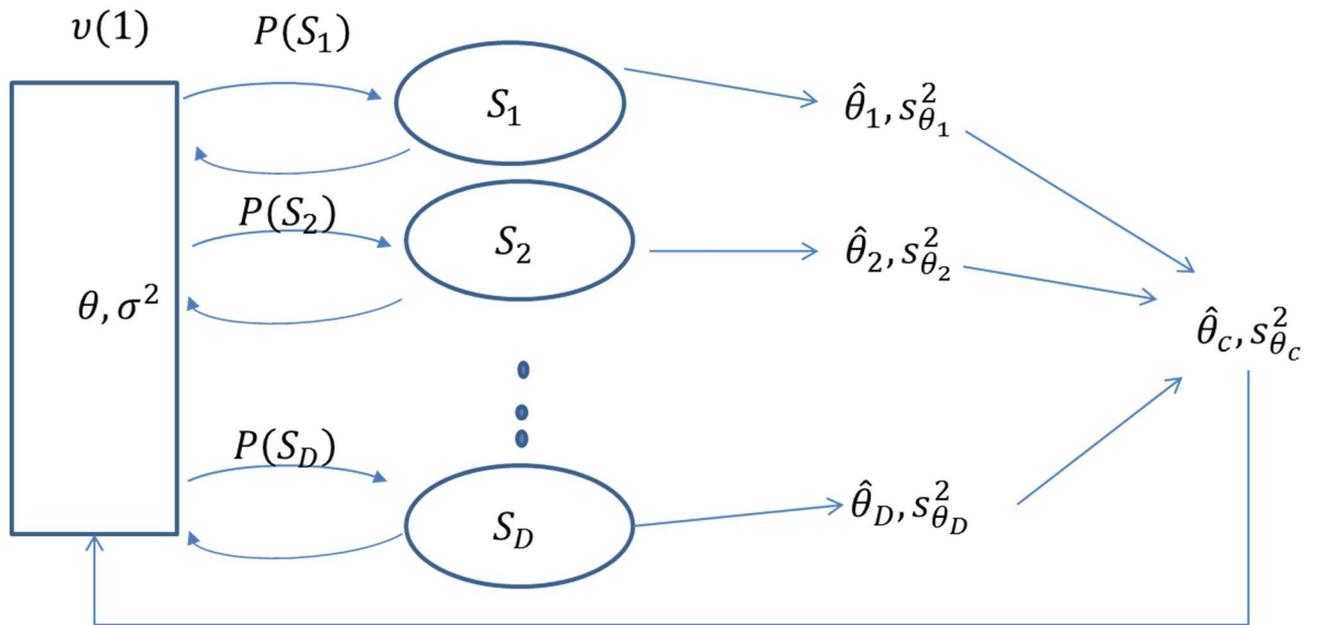


Figure 2: Combining surveys: Scenario 1

The sample surveys are used to calculate the estimates $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_D$ which are then combined to get $\hat{\theta}_c$. Note that $\hat{\theta}_c$ is an estimator for θ , the population parameter for the population $v(1)$. The main purpose of combining the surveys is to get a more precise estimate $\hat{\theta}_c$ for the population parameter θ compared to each of the $\hat{\theta}_d$, with $d = 1, \dots, D$. We assume that the surveys do not overlap in our discussions, which is realistic in most applications. This thesis focuses mainly on combining surveys sampled under this scenario.

The second scenario involves survey samples taken from complementary populations, through a sampling design $P(S_d), d = 1, 2$ as presented in Figure 3.

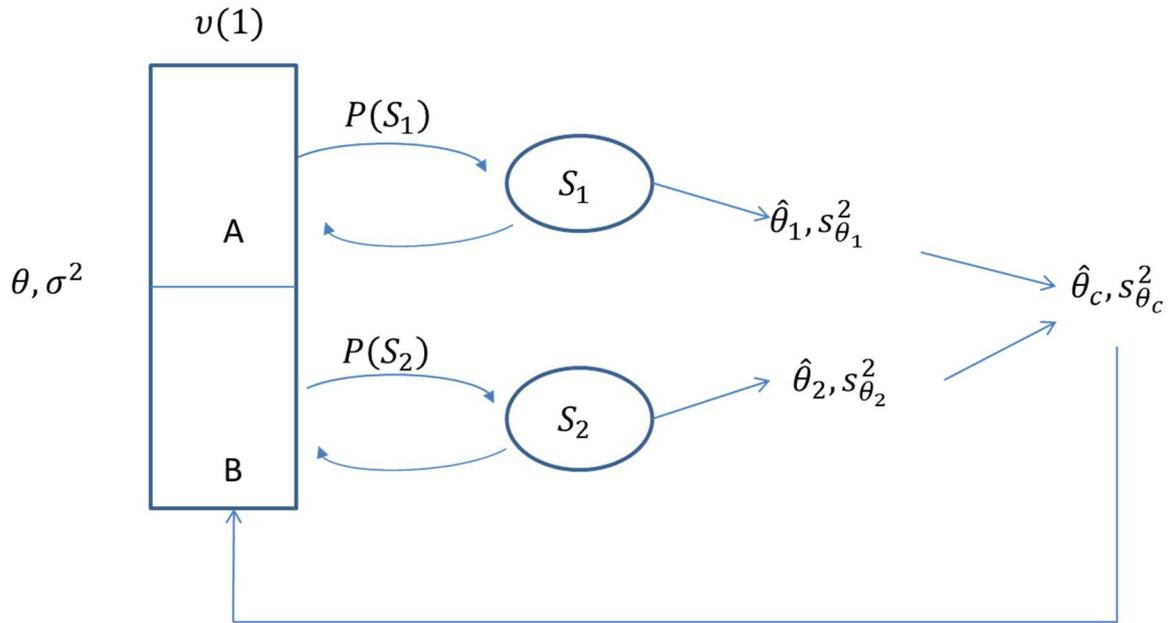


Figure 3: Combining Surveys: Scenario 2

Scenario 2 considers A and B to be complimentary parts of a population $v(1)$. Survey samples are taken from each of A and B and used to calculate parameter estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ which are then combined to get $\hat{\theta}_c$, an estimate of θ the combined population parameter. Schenker et al., point out that, when complementary surveys are combined, it is important to know to what extent the target populations of the complementary surveys are mutually exclusive and exhaustive components of the overall domain of interest^[34]. In that paper, the authors combined the National Health Interview Survey (NHIS)^[35] and the National Nursing Home Survey (NNHS)^[36] to obtain estimates of the prevalence of chronic conditions in the United States of America. The two surveys are complementary because the NHIS looked at the health of the non-institutionalized population of the USA, thus excluding people from the nursing homes who were then included in the NNHS. Considering the NNHS only would have overestimated the prevalence because more sickly people tend to reside in nursing homes compared to ordinary households.^[34] When combining mutually exclusive and exhaustive complimentary surveys, the target populations can be treated as strata in an overall target population encompassing the relevant target populations.^[34]

The third scenario is described by Kish as an accumulation of surveys over time.^{[33][37]} This is illustrated in Figure 4.

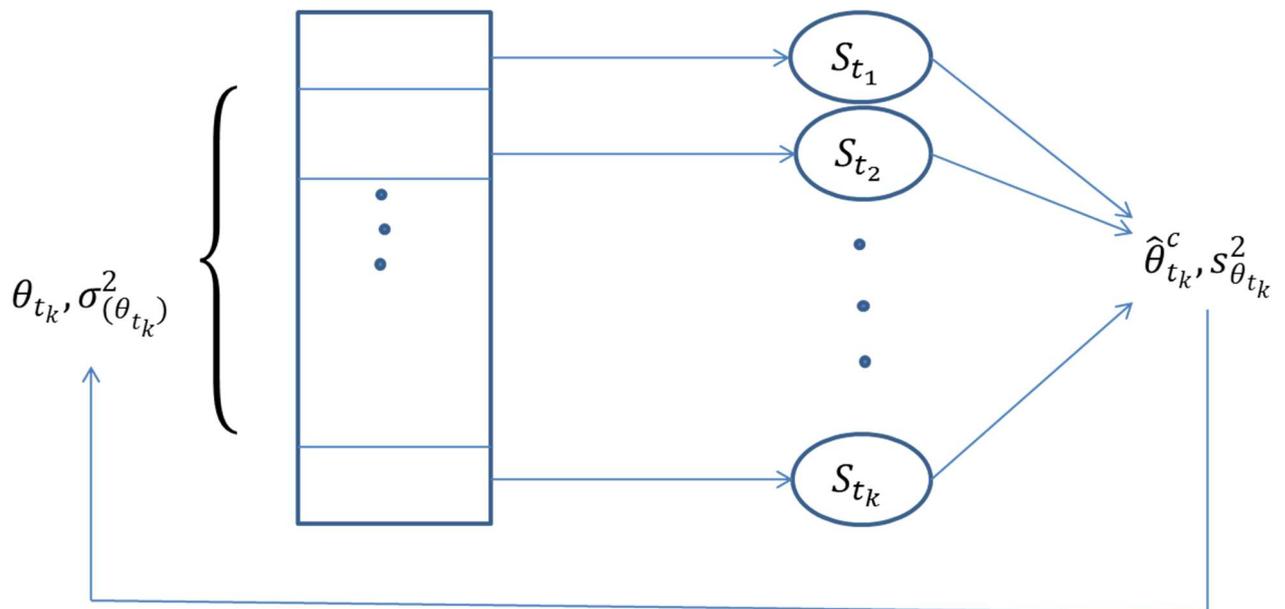


Figure 4: Combining Surveys: Scenario 3

This accumulation is particularly useful when research is focused on rare populations or when the analysis of a time series is required. Samples S_{t_k} are selected from population $\nu(1)$ at different time points t_1, t_2, \dots, t_k . The population parameter at time t_k , θ_{t_k} is then estimated by $\hat{\theta}_{ctk}$ from an accumulation of all the samples. Knowledge about sampling errors is also accumulated together with knowledge about cost factors and components of the total variance. Precision of variance estimates can also be improved by averaging over similar surveys. Information on response and coverage rates is also cumulated allowing for better control of the sample size. Reasons for non-response can be revealed and addressed in successive surveys.

When conducting an analysis where data have been cumulated, consideration should be given to any possible temporal changes, for example, the emergence of new techniques for diagnosing or treating a condition^[18,33].

The combined estimate can be calculated by either using the pooled or separate approach where data can be combined to get estimates or individual estimates are calculated and combined respectively. A more comprehensive discussion of the pooled and separate approaches is given in chapter 2.

1.4 Weighting implications in combining surveys and meta-analysis

When analysing data from multiple surveys, each with its own set of sampling weights and its own sampling design, the question arises over which set of weights to use for calculating point estimates and how to measure the precision of the estimates, addressing the design complexities of all the surveys adequately. This is especially important when data are pooled together to calculate an estimate from the larger combined data set. The weighting strategy has an important impact on the estimation of population parameters regardless of the circumstances under which the combination of surveys is occurring.

Instead of pooling the data, one could calculate estimates from individual surveys and then combine them to get a combined estimate. In this case, the simplest way to get a combined estimate from the multiple surveys would be to take an arithmetic mean of the individual survey estimates. This gives equal importance to all surveys, regardless of size or precision of estimates. All the three scenarios discussed in section 1.3 require careful consideration on how to calculate weights for each survey in the calculation of the combined estimates. That is, when combining surveys selected from the same population as in scenario 1, or combining surveys selected from complimentary parts of the same population as in scenario 2 or analysing an accumulation of surveys over time as in scenario 3.

In the analysis of individual surveys, the use of weights may cause an increase in the standard errors of the estimates.^[13] In an individual survey where data have been collected under simple random sampling, theoretically the probability of selection is equal for each element of the population, resulting in a self-weighting survey. However, with decreasing response rates, probability samples may become more like self-selected samples, that is samples where it is completely left to the participants to decide whether to take part in a survey or not^[38]. Hence, in practice, the need for sampling weights may arise due to nonresponse and missing values. The ideal solution would be to eliminate nonresponse, but this is often very costly and may even be impossible. Weights may be used to adjust for nonresponse. Additionally, population weights may be used to enable the production of

population figures. For example, under simple random sampling, the mean of a variable Y may be estimated as

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i,$$

where y_i is the observed value of variable Y on individual i in a sample of size n . The variance, assuming simple random sampling without replacement is estimated as

$$var(\bar{y}) = \left(1 - \frac{n}{N}\right) \frac{s^2}{n},$$

where $s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$ and N is the population size^[18,22].

Under unequal probability sampling, weights are introduced into the survey and the Hájek estimator for the mean is expressed as^[18,22]

$$\bar{y}_w = \frac{\sum_{i=1}^n w_i y_i}{\sum_{i=1}^n w_i},$$

where w_i is the sampling weight allocated to the i^{th} individual. For the expression of variance of the Hájek estimator, see Sarndal et al (1992).^[39]

When the variable of interest takes different mean values in different subgroups, the population estimate precision may be improved by collecting the data through stratified sampling. That is, if the population can be divided into homogeneous, mutually exclusive groups, (or strata), independent samples can be selected from each stratum. This is referred to as stratified random sampling. This sampling procedure ensures representation of each stratum in the sample. The mean is then estimated by^{[22][40]}

$$\bar{y} = \frac{\sum_{h=1}^H W_h \bar{y}_h}{\sum_{h=1}^H W_h},$$

with $\bar{y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{hi}$ (when simple random sampling is considered within each strata),

W_h the weight allocated to stratum h , y_{hi} the observed value of Y on individual i in stratum h and n_h the sample size, sampled from stratum h . The stratum weights $W_h = \frac{N_h}{N}$ adjust for the size of each stratum. The allocation of sample size to the strata may be proportional to the size of the strata in the population, or equal sample sizes may be allocated to all

strata. A decision may be made to oversample a specific subgroup where information is required in order to improve the precision of the estimates. Alternatively, an optimal allocation where stratum variability is used to allocate sample size may also be used, that is, the Neyman optimal allocation.^[22] The stratum with the greatest variability is allocated a larger sample size.

The variance of the estimate depends on the variability within the strata. It may be estimated using the following formula, when simple random sampling is considered within each strata^[22],

$$var(\bar{y}) = \sum_{h=1}^H \left(1 - \frac{n_h}{N_h}\right) \frac{N_h^2 s_h^2}{N^2 n_h},$$

with $s_h^2 = \frac{1}{n_h-1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2$

However, if there is need for weighting within the strata, for various reasons, the estimation for the mean becomes^[16]

$$\bar{y}_w = \frac{\sum_{h=1}^H \sum_{i=1}^{n_h} w_{hi} y_{hi}}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_{hi}},$$

where w_{hi} , is a composite weight incorporating the design weight and nonresponse adjustments. The variance may be estimated similarly to the estimation for stratification without weights, that is

$$var(\bar{y}_w) = \sum_{h=1}^H \left(1 - \frac{n_h}{N_h}\right) \frac{N_h^2 s_h^2}{N^2 n_h},$$

with $s_h^2 = \frac{1}{n_h-1} \frac{\sum_{i=1}^{n_h} w_{hi} (y_{hi} - \bar{y}_h)^2}{\sum_{i=1}^{n_h} w_{hi}}$.

When units are aggregated into mutually exclusive larger units (clusters), which are easily defined, these clusters may be used as a basis for sampling. That is, the cluster is used as the sampling unit. Cluster sampling may be done to reduce cost, especially where the population is widely distributed geographically or when a list of individuals is not available; but one is available for the clusters. This may be done in one stage, where everyone in the cluster becomes a part of the sample survey. Alternatively, the sampling may be done in

two stages where a sample of individuals is selected from the sampled clusters.^{[40][22]} Many surveys use a simple two stage cluster sampling approach. That is;

1. Select an SRS S of a primary sampling units (clusters or psus) from the population of A psus.
2. Select an SRS of secondary sampling units (ssus) from each selected psu. The SRS of m_a elements from the a^{th} psu is denoted by S_a .

Assuming equal cluster sizes, the mean may then be estimated as^[18]

$$\bar{y} = \frac{\sum_{k=1}^a \bar{y}_k}{a}$$

Where $\bar{y}_k = \frac{1}{m_k} \sum_{j=1}^{m_k} y_{kj}$ is the estimated mean for cluster k . The variance of the mean comprises of two components, that is, variance within clusters and variance between clusters. This may be estimated as^[18,29]

$$\text{var}(\bar{y}) = \left(1 - \frac{a}{A}\right) \frac{1}{a} s_1^2 + \frac{1}{am} s_2^2$$

Where m is the cluster sample size (assumed to be equal in this case), $s_1^2 = \frac{1}{a-1} \sum_{k=1}^a (\bar{y}_k - \bar{y})^2$ and $s_2^2 = \frac{1}{a(m-1)} \sum_{k=1}^a \sum_{j=1}^m (y_{kj} - \bar{y}_k)^2$.

When combining surveys, weighting strategies which lead to smaller mean square errors are preferable in the analysis of both individual and combined surveys.^[27] It is desirable to give the greatest weight to the survey with the most precise estimates to optimise precision in the combined estimate. In meta-analysis, a weight of one, sample size and inverse variance weighting have been compared and results showed that the optimal choice of weighting procedure depended on the statistic of interest.^[41]

When combining surveys, following the fixed effects model of meta-analysis, regardless of the sampling methods employed, the point estimate θ , which may be the mean, may be estimated as^[29]

$$\hat{\theta}_{mc} = \frac{1}{\sum_{d=1}^D w_d} \sum_{d=1}^D w_d \hat{\theta}_d$$

where w_d is the weight allocated to survey d , and is the inverse of the variance of the variable of interest in survey d . $\hat{\theta}_d$ is the estimate of the population parameter θ from survey d . The variance of the combined estimate may be estimated by

$$var(\hat{\theta}_{mc}) = \frac{1}{\sum_{d=1}^D w_d}$$

We consider this together with other variance estimation techniques in the context of combining surveys in chapter 5.

1.5 Motivation and aims of the thesis

Combining survey samples is expected to improve estimates for a parameter of interest because more information is used for estimation. It is expected that the overall increase in sample size would lead to a reduction in sampling errors. Periodic surveys may also be combined to allow for the estimation of change. In the case of multiple frame methods, combining surveys may be used to improve the coverage of individual surveys. Furthermore, many surveys are conducted by individual countries but estimates for the continent may be of interest to the researcher. Some problems may be global such as air pollution or some disease epidemics requiring estimates over different countries.^[28] In order to achieve all these aims, appropriate methods should be used to combine the surveys and produce the improved estimate. However, it is not always clear what the appropriate approach should be for combining multiple surveys.

The aim of this thesis is to evaluate, and possibly improve on, the methods currently available for combining multiple surveys and produce improved estimates. This was achieved through the following objectives:

Study objectives

1. Literature review of analysis of multiple and complex surveys.
2. To compare the weighting strategies available for combining multiple surveys.
3. To propose an alternative weighting strategy for combining multiple cross-sectional surveys.
4. To propose variance estimation methods for the proposed estimators.
5. Compare methods suggested in literature with the one proposed here using simulated data.
6. Use data to test recommended methods in a realistic public health application.

1.6 Outline of the thesis

In this Chapter, the different types of surveys usually used for collecting health related data have been discussed. The advantages and disadvantages of repeated cross-sectional and longitudinal surveys are also given together with some well-known examples of these. The importance of sample surveys to health professionals and policy makers is also discussed. Sampling techniques, oversampling and the need for weighting in complex survey data analysis are also considered. The concepts of the design effect and misspecification effect are introduced. In this chapter, consideration was also given to meta-analysis, contrasting it with combining surveys and discussing the different scenarios where surveys may be combined. The implications of weighting on meta-analysis and combining surveys are also discussed together with the motivation for this study.

Chapter 2 gives a comprehensive review of the different approaches to combining surveys, that is the separate approach and the pooled approach. Regression analysis of survey data is discussed, leading to regression analysis in the context of combined surveys. Some applications in health and social surveys are also examined.

Chapter 3 discusses methods of combining surveys under the separate approach, giving a review of the methods that have been used to date. The study's contributions to these methods under simple random sampling and complex survey sampling are given followed by a table summary of literature where surveys were combined.

Chapter 4 provides simulation studies that were carried out, where different weighting strategies have been used for combining surveys. This thesis investigates different weighting strategies for combining surveys. Previously the sample size, population size and the inverse of variance among many others have been used. In this thesis, the coefficient of variation, skewness, kurtosis and a measure combining skewness and kurtosis are considered. Only mutually exclusive and exhaustive samples are considered in this thesis. We do not consider the case where overlapping samples are combined. The performances of the weighting strategies

were compared through relative bias, mean square error (MSE), coefficient of variation (CV) and confidence intervals. Comparisons were done under the simple random sampling, stratified sampling and cluster sampling settings.

Chapter 5 provides a discussion of variance estimators considered under the separate approach. Traditional methods and replication methods were investigated. The variance estimator used in meta-analysis was investigated for use with the different weighting strategies discussed in chapter 4.

Chapter 6 discusses the simulation studies carried out to compare variance estimators under simple random sampling, stratified sampling and cluster sampling. Chapter 7 illustrates the performance of the weighting strategies and variance estimators with real survey data sets.

Chapter 8 outlines the conclusions from this thesis.

Details of the various simulation results, by distribution, are presented in the appendix. Appendices A to D present results of simulations under simple random sampling, stratified and cluster sampling. R code used in the simulations is given in appendices E and H.

Chapter 2

Methods for combining surveys and applications – A review

Methods used to calculate combined estimates, from multiple surveys, can generally be divided into methods that use separate analyses for the surveys and methods that combine the data and perform a pooled analysis. That is, methods which calculate estimates for each survey and then combine them; and methods which combine data from all the surveys to form one big survey data set before calculating a combined estimate. This chapter presents an introduction to both approaches, although the thesis focuses on the separate approach. We have chosen to focus on the separate approach because micro data are not always available.

2.1 Separate and pooled approaches

In the separate approach, estimates are obtained from each of the surveys separately.^[42,43] A combined estimate is then calculated as a function of the separate estimates, commonly based on a linear combination of the separate estimates. For example, in the case of two surveys, the combined estimate is given by

$$\hat{\theta}_c = \alpha \hat{\theta}_1 + (1 - \alpha) \hat{\theta}_2, \quad (2.1)$$

where $\hat{\theta}_d$, is the parameter estimator which considers data from survey d , with $d=1,2$ in expression (2.1), and α is a weight allocated to survey one. There are many different choices of α which can be used. These will be discussed in chapter 3. The separate approach estimator is unbiased for θ , regardless of the weight α , provided that both $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased estimates for θ .^[43] Latouche et al, for example, used the separate approach to produce cross-sectional estimates based on combining two longitudinal panels of the Survey of Labour and Income Dynamics (SLID).^[44] The survey consisted of two overlapping panels with duration of six years each. For a description of the SLID see Lavigne and Michaud.^[45]

In the pooled approach, individual records from the surveys are combined. Original survey weights may be modified, and estimates are calculated based on the new weights and the pooled sample. Different weight rescaling strategies can be used in either the separate or the pooled approaches. Schenker and Raghunathan combined the National Health Interview Survey (NHIS) and the National Nursing Home Survey (NNHS) to estimate the prevalence of chronic conditions among the elderly people of the United States of America.^[46] In their illustration of how surveys can be combined to extend coverage, they combined the data and treated the two surveys as strata for the combined population. The surveys were considered to be complementary as one targeted the non-institutionalised elderly people and the other referred to the elderly people living in nursing homes.

Roberts and Binder used the following diagrams adapted and presented in figures 5 and 6 to illustrate the difference between the pooled and separated approaches.^[43]

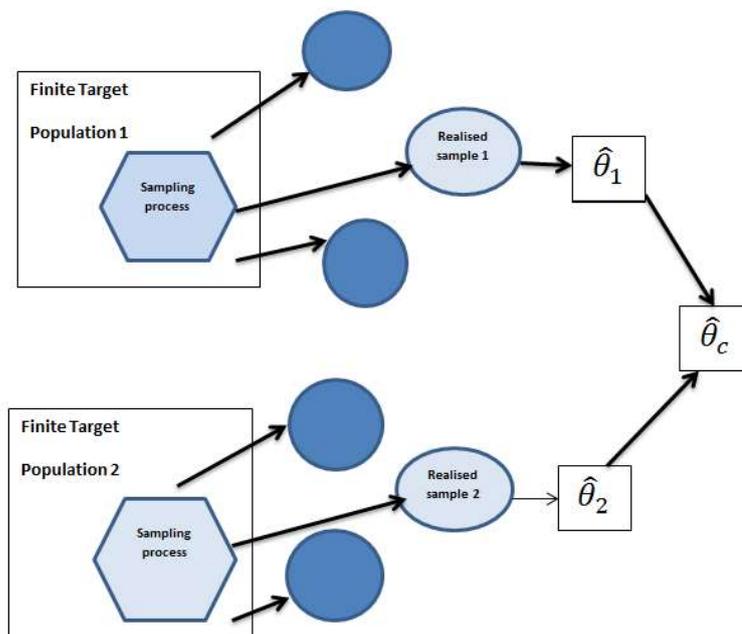


Figure 5: Separate Approach: Adapted from G. Roberts and D. Binder 2009

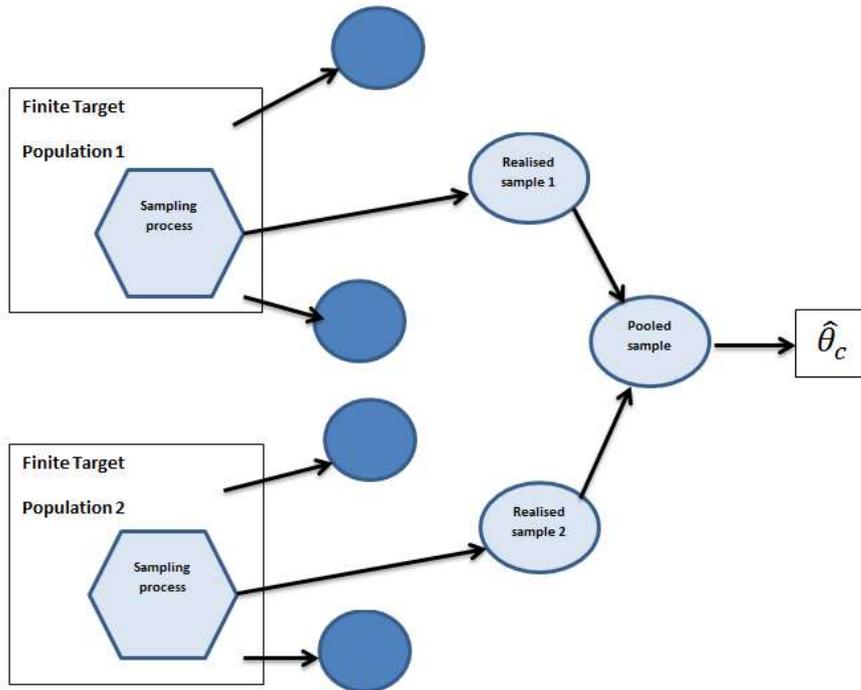


Figure 6: Pooled Analysis: Adapted from G. Roberts and D. Binder 2009.

Because in the separate approach the combined estimate is calculated from the individual survey estimates, it is not essential to have individual micro-data to calculate the combined estimate. In contrast, one can only perform a pooled analysis when micro-data from the individual surveys are available. Wu suggested using an empirical likelihood method to combine information from multiple surveys using the pooled approach.^[47] Suppose in a population of size N , there are two study variables y_1 and y_2 and vectors of auxiliary variables x_1 , x_2 and z . Two surveys are sampled independently such that (y_{1i}, x_{1i}, z_i) are collected in survey 1 and (y_{2i}, x_{2i}, z_i) are collected in survey 2, $i = 1, 2, \dots, n_d$ and $d = 1, 2$. That is, data on the common auxiliary variables z are collected in both surveys. Suppose the population means $\bar{X}_d = \frac{1}{N} \sum_{i=1}^N x_{di}$, $d = 1, 2$ are known but $\bar{Z} = \frac{1}{N} \sum_{i=1}^N z_i$ is unknown. In the pooled approach, \bar{Z} is then estimated from the pooled sample by maximum pseudo empirical likelihood. Wu suggested using the separate approach to avoid computational complications that could arise from trying to estimate the unknown population mean vector \bar{Z} from the pooled sample.^[47]

Thomas and Wannell discussed both the separate and pooled approaches to combining surveys and suggested that the pooled approach is more attractive because of the power of the increased sample size, even though both have their advantages and disadvantages.^[42] The advantages of the separate approach according to these authors are that the combined estimator is easier to interpret and the combined estimate can be calculated even when micro-data are not available. The authors caution that the separate approach can be cumbersome if the required estimates are not published or have been published without their variance estimates. One would have to calculate them separately for each survey. This can be time-consuming especially when many estimates are required. The pooled approach has an advantage in the power of the larger sample size. Moreover, once a combined estimate has been calculated, there will be no need to go back to the individual survey data sets. However, the pooled approach requires more technical expertise in manipulating data files.^[42]

O'Muircheartaigh and Pedlow compared estimating the combined parameter by creating sampling weights based on cumulating cases and deriving appropriate sampling weights (pooled approach) to using weights adjusted as a linear combination of the two individual survey weights (separate approach). They concluded that cumulating cases was better for the US National Longitudinal Survey of Youth 97.^[48]

Merkouris proposed a composite regression estimator for producing a combined estimate where there is a single common variable in two independent surveys, using the separate approach.^[25,49] The basic principles are similar to the composite estimation which will be further discussed in section 3.1 in a non-regression context. Merkouris advocated the use of the pooled approach for the general setting where there are q common variables and different auxiliary variables for the two survey samples. He suggested that the pooled approach might yield more efficient estimators for the non-common variables depending on their correlation with the common ones. The generalized regression procedures yield sampling weights that

incorporate auxiliary information from all the surveys, resulting in more efficient estimators. Information is therefore shared between the surveys.

Dong et al. used multiple imputation to combine information from multiple surveys sampled from the same finite population. They generated synthetic populations using a finite population Bayesian bootstrap which accounts for complex sample designs. Each synthetic population was analysed as if data were collected through simple random sampling. The point and variance estimates were combined using extensions of the combining rules for synthetic data. Data from the 2006 National Health Interview Survey (NHIS) and the 2006 Medical Expenditure Panel Survey (MEPS) were combined to illustrate the method and obtained more precise estimates than those obtained from the individual surveys^[50].

A measurement error model approach for combining information from two surveys was proposed by Park et al. One survey, considered to be without measurement error (gold standard), is used to estimate values for the other survey through parametric fractional imputation. Parameter estimates are calculated from the imputed survey and the gold standard. A composite estimator combining the two values is then used to estimate the overall parameter. They applied the method by combining The Food and Nutrition Technical Assistance III Project (Guatemala) baseline and interim surveys to obtain zone of influence estimates. The geography of the two surveys did not exactly coincide. However, there was substantial geographic overlap. The union of the geography covered by the two surveys represents the Feed The Future Zone of Influence.^[51]

2.2 Analysis of survey data in the context of combining surveys

Traditionally, analysis of survey data used to involve mainly producing estimates of population via descriptive statistics. However, regression on survey data has gradually become more prevalent. The main objective of regression is to establish the relationship between a response variable and one or more explanatory variables. That is, to find a scientifically interesting and parsimonious explanation of the systematic variation in the response variable. For example, socio-economic, demographic and behavioural characteristics may be used in a model studying the variation in health status and use of health services.^[52]

There are differences between estimation methods for classic linear regression and regression on complex survey data. Data from complex surveys are usually not independent and identically distributed (iid). Usually there are varying inclusion probabilities in the survey sample. An analysis which does not adjust for these varying inclusion probabilities may result in biased estimates for the regression parameters.^{[22][52]} Stratification should be considered when analysing complex surveys which use stratified random sampling. Stratified samples usually result in smaller variance estimates compared to a simple random sample of the same size. Therefore, variances are overestimated when stratification is not accounted for in the analysis. Additionally, if the survey uses cluster sampling, then intra-cluster correlation should be accounted for in the regression model. This may be done through multi-level hierarchical modelling or the use of weighted least squares or pseudo maximum likelihood estimation, for example.^{[52][13]}

The classic simple linear regression model is expressed as

$$y_i = \beta_0 + \beta_1 x_{1i} + \varepsilon_i,$$

where y_i is the dependent variable, x_{1i} is the explanatory variable, β_0 and β_1 are the model parameters and ε_i is the random error term which reflects the difference between the observed y_i and its conditional expectation under the model. In the case of multiple linear regression, there may be more than one explanatory variable, $1, \dots, p$, in the model. The regression model would be represented as

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi} + \varepsilon_i.$$

The finite population parameter values of β_p minimize the residual sum of squares (RSS)

$$RSS = \sum_{i=1}^N (y_i - \mathbf{x}_i' \boldsymbol{\beta})^2.$$

Where \mathbf{x}_i is the vector of explanatory variables and $\boldsymbol{\beta}$ is the model parameter vector. In the case of regression of complex survey data, a choice is made on whether to estimate the finite population parameters or parameters of the more universal superpopulation. The superpopulation is defined as the infinite population which is assumed to have generated the finite population. Weighted estimation of the regression parameters results in unbiased estimates for the finite population regression parameters \mathbf{B} . These parameters are expected to vary about the true superpopulation model parameters $\boldsymbol{\beta}$. When the finite population is large, unbiased estimates for \mathbf{B} can be used as unbiased estimates for $\boldsymbol{\beta}$. This may be estimated from a sample, incorporating survey weights as

$$\widehat{RSS} = \sum_{i=1}^n w_i (y_i - \mathbf{x}_i' \widehat{\boldsymbol{\beta}})^2$$

Where w_i is the survey sampling weight associated with individual i .^[13] When clustering and stratification are used,

$$\widehat{RSS} = \sum_{h=1}^H \sum_{k=1}^K \sum_{i=1}^{n_{hk}} w_{hki} (y_{hki} - \mathbf{x}_{hki}' \widehat{\boldsymbol{\beta}})^2,$$

where h and k are indexing strata and clusters respectively.

In the context of combined surveys, Rao and Kim discussed two scenarios in which regression parameters could be estimated by combining data from two independent surveys.^[53] In the first scenario, a large survey A only has observations of the explanatory variables, x and a smaller survey B has observations of both the dependent y and explanatory variables. The authors proposed to fit a regression model using survey B and use it to predict the y values for survey A. The predicted values in survey A are then used together with the explanatory variables for inference.

In the second scenario, the dependent and explanatory variables are observed in both survey A and survey B. The data set (y_1, x_1) is observed from survey A and (y_2, x_2) from survey B. Statistical matching is used to estimate \tilde{y}_1 values for survey B based on common explanatory variables. A regression of y_2 on y_1 and other explanatory variables of interest is then fitted on survey B.^[53] For statistical matching, they considered multiple imputation, and a method based on an instrumental variable, to generate \tilde{y}_1 values for survey B.

In the closely related field of dual frame surveys, Lu discussed a weighted average of independent regression coefficient estimates from two surveys.^[54] Regression methods will not be used in this thesis, as they will be considered in planned future work. Elkasabi discussed weighting procedures for dual frame telephone surveys using a case study in Egypt.^[10] He acknowledged that different weights may apply to different frames and this needed to be addressed in estimation procedures.

We present in the following section examples of how surveys have been combined for health and social surveys research.

2.3 Application in health and social surveys

Combining surveys can be very beneficial in the production of health and social statistics. Surveys have been combined to serve various purposes in health and social surveys research. When there is no single survey that has been designed to cover a population of interest, but there are two or more surveys covering subsets of the population of interest, these may be combined to extend coverage and get improved estimates for the population of interest. The 1985, 1995 and 1997 National Health Interview Survey (NHIS) and the National Nursing Home Survey (NNHS) were combined to estimate prevalence rates of chronic diseases among the elderly people of the United States of America (USA).^[34,46]

When there is more than one source of information, for example self-reporting of certain conditions and doctors' examination reports, integrating the information from the two sources may result in improved estimates. Schenker and Raghunathan combined two surveys conducted by the National Centre for Health Statistics to improve on analyses of self-reported data on health conditions in the USA. One of the surveys, the National Health and Nutrition Examination Survey, asked self-report questions on health conditions during face-to-face interviews, but it also obtained clinical measures based on physical examinations. The other survey, the NHIS relied on self-report questions on health conditions.^[46] Integrating the self-reported data with physical examinations enhanced the accuracy of the estimates.

The Behavioural Risk Factor Surveillance System and the NHIS have been combined to estimate prevalence of cancer risk factors and cancer screening in the United States of America at the county level^[46]. The United States of America revised their standards for classifying race and ethnicity in 1997 to allow individuals to classify themselves as more than one race group. The previous standards established in 1977 allowed for only one race group. The NHIS allows respondents to classify themselves as more than one race group but asks a follow up question where the respondent has to choose one race group that best describes them. In order to compare vital statistics from censuses pre and post 1997, the NHIS was

used to build bridging models which can be used to predict the single race groups respondents would have chosen had the 1977 standards been still in place.^[46]

Rao et al. used meta-analysis principles in health services research to combine surveys to obtain an estimate for the use of electronic health records.^[32] Kish also suggests that the problem of combining surveys has some fundamental features in common with the problem of combining information obtained from different experiments.^{[37][28]}

Roberts and Binder studied whether gay and bisexual men differ in their use of health care. They combined data from the Canadian Community Health Surveys of 2003 and 2005. The sample sizes of gay and bisexual men were small in both surveys hence the surveys were combined to increase sample size and improve accuracy of estimates.^[43]

Two panels of the Survey of Labour and Income Dynamics (SLID) were combined to obtain improved estimates on the labour activities of individuals in Canada.^[44]

Although one panel could have been used to obtain these estimates, better accuracy might have been attained by combining the two panels.

Three studies investigating compulsive internet use, one involving private schools, one involving the same private schools at a later time point, and one involving public schools, were combined to build a model predicting compulsive internet use among adolescents in India.^[3]

The National Longitudinal Survey of Youth (NLSY97) in the United States of America examines the issues surrounding youth entry into the work force and the subsequent movements in and out of the workforce. It consists of two samples, the cross-sectional sample, representing the population race groups in their correct proportions, and a supplemental sample where Hispanic and non-Hispanic Black youths are oversampled. These two samples were combined to obtain more accurate estimates for the minority Hispanic and non-Hispanic Black youth.^[48]

The Behavioural Risk Factor Surveillance System (BRFSS) and National Health Interview Survey (NHIS) were combined using propensity scores to obtain cancer risk factor prevalence in small areas.^[55] The NHIS was designed to produce annual estimates of a variety of health conditions and health related behaviours at national

level. The BRFSS was designed to produce state-specific estimates of risk behaviours. The NHIS had a high response rate and the BRFSS had larger sample sizes in each state. Combining the two surveys yielded more accurate estimates. In this thesis, we combine sample surveys selected from the South African Community Survey of 2016 under simple random sampling to obtain a more accurate estimate of age of the head of the household. Under stratified sampling, we combine sample surveys selected from the South African General Household Survey of 2016 to obtain a more accurate estimate of the age of the head of the household. There appears to be a large number of child-headed households because of the HIV/AIDS epidemic. These children would require special social support systems since they are unlikely to have developed enough skills to support other minors left in their care. A study in Iran found a significant association between the use of health services and the age of the head of the household. The study showed an increase in the demand for outpatient services with increase in age of the head of the household.^[56] Another study in Nigeria showed that there was a significant association between the use of improved drinking water sources and the age of the household head. Fewer households with younger heads of household, 15 years to 24 years, used improved drinking water sources compared to older age groups.^[57] It is therefore important to have an idea of the average age of household heads in the population.

Chapter 3

Methods for combining surveys: Point estimators under the separate approach

3.1 A review

Methods used to calculate combined estimates from multiple surveys can generally be divided into methods that use separate parameter estimates from the different surveys; and methods that combine the data and perform a pooled analysis. This thesis focuses on methods that use separate parameter estimates from the different surveys, loosely referred to as the separate approach, in a context when the individual survey data are not available together with design information.

Some authors have used a straightforward average of the descriptive statistics obtained from the individual surveys as the estimate from the combined surveys.^[42] The combined estimate in this case is calculated as

$$\hat{\theta}_c = \frac{1}{D} \sum_{d=1}^D \hat{\theta}_d,$$

where $\hat{\theta}_c$ is the combined estimate and $\hat{\theta}_d$ are the estimates obtained from the individual surveys $d = 1, \dots, D$.

For example, Thomas and Wannell used averages to combine point estimates from the different cycles of the Canadian Community Health Survey. The Canadian Community Health Survey (CCHS) consists of two cross-sectional sample surveys. The first cycle collects general health information from more than 120 health regions, while the second cycle focuses on specific health topics and collects data for estimation at the provincial level.^[42,58] Analysis of each survey accounts for the survey design in each case. Using averages to combine survey point estimates gives equal weight, and hence importance to all surveys, regardless of size or precision of estimates.

Different weighting strategies have been used in combining point estimates from different surveys. Sample size is often used for weighting when combining estimates from small samples.^[28] This method rests on the assumption that larger samples results in estimates with greater precision compared to smaller surveys.

However, if there is clustering, as it is often the case with complex surveys, this may not always be true. For example, if survey A has a sample size of 100 with ten clusters and survey B has a sample size of 300 with five clusters, assuming the same intra-cluster correlation for both surveys, it is likely that survey B will have a larger design effect and hence less precision despite its larger sample size. Using the effective sample size for weighting might be a way of handling this problem.^[58] Chu et al. suggested that using effective sample size in composite weighting, corrects for size and precision.^[59] The composite estimator for combining estimates from two surveys, for example, may be expressed as

$$\hat{\theta}_c = \alpha \hat{\theta}_1 + (1 - \alpha) \hat{\theta}_2$$

where $\hat{\theta}_1$ is estimated from survey 1 and $\hat{\theta}_2$ is estimated from survey 2. The weight α may be chosen to be the effective sample size of survey 1 as a fraction of the total effective sample size of the two surveys combined, for example.

In Biostatistics, effect size is defined as a quantification of an impact of an intervention, a quantification of the relationship between two variables or even the prevalence of a condition. The effective sample size in a complex survey is the sample size for a simple random sample selected with replacement which gives the same variance for an estimate as the variance obtained using data from the complex survey.^[18,43,58] The effective sample size is calculated as $n_d^* = \frac{n_d}{\overline{Deff}(\hat{\theta}_d)}$. Here n_d is the sample size for survey d and $\overline{Deff}(\hat{\theta}_d)$ is an average design effect. The design effects may vary across subgroups and it becomes difficult to calculate individual design effects for each subgroup. A compromise or average design effect is used instead. The effective sample size can be used to weight survey estimates when calculating a combined estimate when the surveys have different variances because of design effects.^[28,33,44,60] The combined estimate may be calculated as,

$$\hat{\theta}_c = \frac{\sum_d^D n_d^* \hat{\theta}_d}{\sum_d^D n_d^*}$$

The effective sample size is often used for combining national estimates.^[28] It is sometimes not possible to calculate the correct effective sample size. This may be because it is too complicated, too computationally expensive or because there is insufficient information for the effective sample size to be computed^[18]. An alternative formula developed by Kish is often used to estimate the effective sample size,

$$\hat{n}^* = \frac{(\sum_{i=1}^n w_i)^2}{\sum_{i=1}^n w_i^2}$$

where w_i is the weight of the i th respondent. This estimation is very inaccurate when the weight is correlated with the variables of interest in the analysis. It also ignores clustering^[18].

Survey weights can be allocated proportional to the size of the population from which the sample survey was taken.^[33] This method gives more weight to surveys representing larger populations.

$$\hat{\theta}_c = \frac{\sum_d^D N_d \hat{\theta}_d}{\sum_d^D N_d}$$

where N_d is the size of the population from which the d th sample survey was sampled. Using the population size for weighting may be justified when population differences seem to be more important than sampling precision.^[60] Population sizes are often used for large samples and combining domain statistics to get national statistics and have been shown to yield unbiased estimates.^[28]

Different studies will differ in their degrees of precision in estimating the effect size. Precision in this case describes the accuracy of the estimated effect size as an estimate of the true effect size.^[31]

A closely related field, meta-analysis, usually uses the inverse of variance of the estimator applied for each individual study to weight the different studies in the calculation of a combined effect size. In this way, less precise studies are given less

weight in the calculation of the combined estimates^[31] The individual estimates of the variances of the estimators are used to weight the sample surveys when a combined estimate is calculated,^{[60][29]}

$$\hat{\theta}_c = \frac{\sum_d^D \frac{1}{Var(\hat{\theta}_d)} \hat{\theta}_d}{\sum_d^D \frac{1}{Var(\hat{\theta}_d)}}$$

where $Var(\hat{\theta}_d)$ are the variance of the estimators, $\hat{\theta}_d$, in the individual surveys.

This weighting system could also be used to weight different surveys when multiple complex surveys are analysed. In this thesis, the author evaluates the use of the inverse of variance for weighting surveys to be combined, as with meta-analysis.

There are two statistical models usually used in meta-analysis, that is, the fixed effects model and the random effects model. Meta-analysis is commonly used in clinical studies where the aim is to find the effect of an intervention. The fixed effects model assumes that there is one true effect which all the studies are trying to estimate. All the observed differences are due to sampling error. The estimated effect for any study can be written as

$$\hat{\theta}_d = \theta + \varepsilon_d,$$

where $d = 1, 2, \dots, D$, ε_d is the random sampling error in study d , and θ is the common effect size. In contrast, the random effects model assumes that there may be different true effect sizes across studies. For example, an effect may be lower in a study where participants are older or less healthy than others. If an infinite number of studies, each with its own true effect size could be carried out, then the true study effect sizes would be distributed about some mean parameter, θ . The true effect sizes of the studies used in the meta-analysis could be viewed as a random sample from this distribution. The estimated parameter in any study could be written as

$$\hat{\theta}_d = \theta + \tau_d + \varepsilon_d$$

where θ is the mean of the different true study parameters, τ_d is the deviation of each true study effect size from the mean and ε_d is the random sampling error in study d . The expression for the fixed effects model does not have τ_d because it is assumed that all the studies being combined have one true effect size. The observed variation in the estimated statistics represents the sampling errors within studies and the variation of the true study effect sizes across studies. When carrying out a meta-analysis, it is of great importance to use study weights which minimize both sources of error.^[31]

In estimation, it is desirable to minimize variance, hence improving precision. In meta-analysis, the inverse variance weighting method has been shown to yield the most precise estimate. The weighting strategy penalizes surveys for imprecision when they are combined. However, in surveys, variability does not only originate from the model that generates the data or the sample size. Some variability originates from the sampling procedure itself.

One frequently used method in combining survey estimates that minimizes the variance of the combined survey estimate, uses the design effects.^[29] In this method, α is set as a function of the variability of weights for the two surveys. The survey with the greatest precision is given greater relative influence in the scaling of weights. The formula used for calculating α is:

$$\alpha = 1 - \frac{Def f(w_{1i})}{Def f(w_{1i}) + Def f(w_{2i})}$$

where $Def f(w_{1i})$ and $Def f(w_{2i})$ are the respective design effects due to unequal weighting for Survey 1 and Survey 2. It makes no difference which $Def f(w_{di})$ is used in the numerator. Calculation of the design effect due to unequal weighting for a survey is accomplished using the following formula^{[42][33]}:

$$Def f(w_{di}) = 1 + \left(\frac{cv(w_{di})}{100} \right)^2$$

where $cv(w_{1i})$ is the coefficient of variation for the weights in Survey d and

$$cv(w_{di})\% = \frac{\sqrt{Var(w_{di})}}{\mu_{w_{di}}} \times 100\%.$$

where $Var(w_{di})$ is the variance of the weights from survey one and $\mu_{w_{di}}$ is the mean of the weights from survey d .^[61] This calculation of design effects, however, does not capture the effects of clustering and stratification. Although this is a separate approach method, micro data are required in the calculation, losing the advantage which other separate approach methods have. Thomas et al also used design effects in the calculation of weights for combining cycles of the Canadian Community Health survey with the weight allocated to survey d expressed as $w_d = \frac{n_d}{Def f_d}$, where n_d is the sample size for survey d and $Def f_d$ is the design effect for the d^{th} survey.^{[43][62]}

Selected, frequently used weighting strategies were assessed through simulation in this thesis, for use in combining surveys. The variance estimators for the resulting estimators are discussed in Chapter 5. The properties of the estimators were studied through simulation, under different conditions, i.e. with and without clustering and stratification. The results of the simulation are presented in chapter 6.

3.2 Contributions under simple random sampling

3.2.1 Methods used to weight surveys

Fox compared the sample size, the inverse variance and straightforward average weighting strategies for combining surveys.^[29] Fox generated a finite population, selected simple random samples from the finite population, calculated point estimates, the sample mean, from the selected samples and combined the sample means, weighting each sample mean by either the inverse of the sample size, the inverse of the variance or a weight of 1. She then studied the behaviour of the combined mean estimate as the number of samples being combined increased. The combined estimate of the mean converged towards the finite population mean. In this thesis, the coefficient of variation (cv), skewness, kurtosis and the D'Agostino-Pearson test statistic were considered for use as the weighting strategies when combining surveys, under the separate approach. These were compared to the sample size and the inverse variance studied by Fox. Initial comparisons were made by combining simple random samples. All the scenarios considered, and results, are presented in Chapter 4.

The coefficient of variation is a relative measure of precision estimated as ^[63,64]

$$cv = \frac{s.e(\hat{\theta})}{\hat{\theta}} \times 100\%$$

where $s.e(\hat{\theta})$ is the standard error of the estimate $\hat{\theta}$. When combining surveys, it is desirable to give more weight to surveys with greater precision hence it is important to consider variability. In this thesis, we evaluate the use of the inverse of the coefficient of variation to weight surveys in calculating combined estimates, $\hat{\theta}_c$,

$$\hat{\theta}_c = \sum_{d=1}^D \left(\frac{\frac{1}{cv(\hat{\theta}_d)} \times \hat{\theta}_d}{\sum_{d=1}^D \frac{1}{cv(\hat{\theta}_d)}} \right). \quad (3.1)$$

We consider the use of the inverse of the coefficient of variation estimate to weight surveys, in addition to using the inverse of the variance as used in meta-analysis,

because the standard deviation (square root of the variance) must be understood in the context of the estimate; whereas the coefficient of variation is independent. To give an extreme example, a survey with a standard error of ten with an estimate of one hundred, $cv=10\%$ and a survey with a standard error of ten with an estimate of 1000, $cv=1\%$, do not speak to the same degree of precision. Although the value of the standard deviation is the same in both cases, there is greater variability in the survey with a $cv=10\%$. The interpretation of the standard deviation should be relative to the size of the mean whereas the cv is absolute. When working with different surveys, the coefficient of variation might offer a more reasonable measure of variation than the variance.^{[40][65]}

We also consider using the inverse of skewness as a weighting strategy for combining surveys. Skewness measures the extent to which the tails of the data may be more drawn out on one side than the other. A symmetric distribution will have a skewness equal to zero. Skewness may be calculated from simple random sample data as follows

$$skewness = \sqrt{b_1} = \frac{m_3}{m_2^{3/2}} = \frac{m_3}{s_3}$$

where $m_k = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^k$, the k^{th} moment about the mean of the variable of interest and $s_3 = (\sqrt{s^2})^3$ the cube of the standard deviation.^[66] This formula changes when complex sampling is used.

It is well known that estimators behave better with symmetric data and that data from most populations are skewed; hence it might be worthwhile to consider skewness when combining surveys,^[22] with the combined estimate expressed as

$$\hat{\theta}_c = \sum_{d=1}^D \left(\frac{\frac{1}{skewness(\hat{\theta}_d)} \times \hat{\theta}_d}{\sum_{d=1}^D \frac{1}{skewness(\hat{\theta}_d)}} \right), \quad (3.2)$$

where $skewness(\hat{\theta}_d)$ is the skewness estimated from survey d . Weighting the surveys by the inverse of skewness favours the less skewed sample surveys, giving larger weight to the more symmetric survey.

Kurtosis measures the degree of curvature for a distribution.^[67] If the distribution is flat-topped or too peaked, then it deviates from the bell shape of the normal distribution. The more data is like a normal distribution, the better behaved the estimators are according to established statistical theory. Weighting the surveys by the inverse of kurtosis gives greater weight to the data which resembles the normal distribution more. Kurtosis may be calculated as follows

$$kurtosis = b_2 = \frac{m_4}{m_2^2} = \frac{m_4}{s_4},$$

where $s_4 = (s^2)^2$, the square of the variance.^[66] Data which are normally distributed have kurtosis equal to three. Some statistical software, for example R, calculate kurtosis as

$$kurtosis = b_2 - 3 = \frac{m_4}{s_4} - 3$$

The combined estimate may be calculated as

$$\hat{\theta}_c = \sum_{d=1}^D \left(\frac{\frac{1}{kurtosis(\hat{\theta}_d)} \times \hat{\theta}_d}{\sum_{d=1}^D \frac{1}{kurtosis(\hat{\theta}_d)}} \right), \quad (3.3)$$

where $kurtosis(\hat{\theta}_d)$ is the kurtosis estimated from survey d . Skewness and kurtosis were considered for use in combining surveys because the efficiency of variance estimators is based on the properties of the point estimator as well as properties of the possible skewness or kurtosis of the variable of interest.^[68]

Skewness and kurtosis do not always occur in isolation in survey data. They often occur together, hence the D'Agostino-Pearson K^2 test statistic that combines skewness and kurtosis was considered as a possible weighting strategy when combining surveys. The inverse of the D'Agostino-Pearson statistic was assessed for weighting surveys in calculating the combined estimate. The test has good power properties over a broad range of non-normal distributions.^[67] The statistic is calculated as follows

$$K^2 = Z^2(\sqrt{b_1}) + Z^2(b_2),$$

where $Z(\sqrt{b_1})$ and $Z(b_2)$ are the normal approximations to $\sqrt{b_1}$ and b_2 respectively. We calculate the combined estimate

as

$$\hat{\theta}_c = \sum_{d=1}^D \left(\frac{\frac{1}{K^2(\hat{\theta}_d)} \times \hat{\theta}_d}{\sum_{d=1}^D \frac{1}{K^2(\hat{\theta}_d)}} \right). \quad (3.4)$$

3.3 Criteria for evaluating estimators

The mean square error (MSE), relative bias, confidence intervals and the coefficient of variation were calculated and used to compare the estimators under different weighting strategies. These measures have been used in the comparison of the performance of the estimators in a simulation study considering weighting methods to improve, for example, labour force estimates of immigrants in Ireland.^[69] We use these criteria to compare weighting strategies in this thesis, with the MSE carrying the most weight.

The bias of an estimate is defined as the difference between the expected value of an estimator and its true population parameter.^[40] An estimator is said to be unbiased if the bias is equal to zero. The bias is defined as

$$Bias(\hat{\theta}) = E(\hat{\theta}) - \theta,$$

where $\hat{\theta}$ is an estimator for θ . In this thesis, in the simulation studies, we consider combining $D=2, 3, 5, 10, 15$ or 20 samples with $K=1000$ replicates. The expected value of the combined estimate is estimated as

$$\hat{E}(\hat{\theta}_c) = \frac{1}{K} \sum_{k=1}^K \hat{\theta}_{c,k}$$

where $k=1, \dots, K$. Hence bias is estimated as

$$Bias(\hat{\theta}_c) = \hat{E}(\hat{\theta}_c) - \theta$$

The bias is sensitive to scale hence it may be more informative to calculate the relative bias, which is the ratio of the bias to the estimate,

$$Relative\ Bias = \frac{Bias(\hat{\theta})}{\hat{\theta}} \times 100\%$$

. For our simulation, we calculate relative bias as

$$Relative\ Bias(\hat{\theta}_c) = \frac{Bias(\hat{\theta}_c)}{\hat{E}(\hat{\theta}_c)} \times 100\% \quad (3.5)$$

The variance is often used as a measure of precision of an estimator when there is no bias or when the bias is small. The variance of the combined estimate may be estimated empirically through simulation as^[68]

$$var_{empirical}(\hat{\theta}_c) = \frac{1}{K-1} \sum_{k=1}^K (\hat{\theta}_{c,k} - \hat{E}(\hat{\theta}_c))^2 .$$

The standard error may be expressed as

$$s.e_{empirical}(\hat{\theta}_c) = \frac{\sqrt{var_{empirical}(\hat{\theta}_c)}}{\sqrt{K}}$$

The MSE measures the accuracy of an estimator and is equal to the average squared distance between each sample estimate and the corresponding true population percentage.^[69] It incorporates bias and precision into an overall measure of accuracy which can be used even in the presence of bias. The MSE for the combined estimate $\hat{\theta}_c$, may be calculated as^{[40][52]}

$$MSE(\hat{\theta}_c) = Var(\hat{\theta}_c) + [Bias(\hat{\theta}_c)]^2 \quad (3.6)$$

In this thesis, we use the MSE to compare the performance of the various weighting strategies under consideration, with the weighting strategy yielding the lowest MSE value being selected as the best weighting strategy for a given scenario.

The quality of confidence intervals may be used to compare estimators. It is desirable to have tighter confidence intervals which contain the population parameter value being estimated. The 95% simulation confidence intervals were calculated as

$$CI: \hat{\theta}_c \pm Z_{\frac{\alpha}{2}} \times s.e(\hat{\theta}_c) \quad (3.7)$$

In our simulations, the finite population mean could be calculated, hence we also checked whether the 95% confidence interval included the population mean.

The coefficient of variation has been considered as a weighting strategy because the cv measures the relative variability of an estimate. For the same reason, the coefficient of variation may be used to assess the performance of an estimator.

$$cv = \frac{s.d(\hat{\theta}_c)}{\hat{E}(\hat{\theta}_c)} \times 100\% \quad (3.8)$$

In this thesis, we calculated and compared the cv for the combined estimates though greater weight was accorded to the MSE values.

3.3 Contributions under complex sampling: stratification and clustering

All the weighting strategies considered under simple random sampling were studied under complex sampling, except the D'Agostino-Pearson statistic. The D'Agostino-Pearson statistic was excluded to avoid the computational difficulties associated with calculating the statistic under complex sampling conditions. The combined estimate could be expressed as

$$\hat{\theta}_c = \sum_{d=1}^D \left(\frac{w_d \times \hat{\theta}_d}{\sum_{d=1}^D w_d} \right)$$

The weight allocated to an individual survey, w_d could be 1, the inverse of the sample size, n , or the inverse of the variance, $\frac{1}{var(\hat{\theta})}$. A design-consistent variance, allowing for either stratification or clustering, was calculated using the R survey package. The inverse of the coefficient of variation was also evaluated for weighting complex surveys in the calculation of the combined estimate. A design consistent estimate of the variance was used to calculate a design consistent coefficient of variation.

For stratified samples, the sample moment of order k may be calculated as follows^{[70][71]}

$$mnt_k = \frac{\sum_{h=1}^L \sum_{i=1}^{n_h} w_{hi} y_{hi}^k}{\sum_{h=1}^L \sum_{i=1}^{n_h} w_{hi}}$$

Skewness may then be calculated as

$$skewness = \frac{mnt_3 - 3mnt_2 m_1 + 2mnt_1^3}{(mnt_2 - mnt_1^2)^{\frac{3}{2}}}$$

And kurtosis may be calculated as

$$kurtosis = \frac{mnt_4 - 2mnt_2 mnt_1^2 + mnt_1^4}{(mnt_2 - mnt_1^2)^2}$$

In general, for any weighted samples, (because of clustering, stratification and any other reason), the sample moment may be expressed as^{[72][73]}

$$mnt_k = \frac{\sum_{i=1}^n w_i y_i^k}{\sum_{i=1}^n w_i}$$

where w_i incorporates all the weights associated with individual i .

In this study, skewness and kurtosis for complex surveys were calculated using the R survey package. The package allows for study sampling design in the calculation of all the statistics.

Additionally, the inverse of the misspecification effect, which measures the impact of the sampling design, was evaluated as a possible weighting strategy when combining estimates under complex sampling. The combined estimate, using the inverse of the misspecification effect was calculated as

$$\hat{\theta}_c = \sum_{d=1}^D \left(\frac{1}{meff_d} \times \hat{\theta}_d \right), \quad (3.9)$$

where $meff_d$ is the misspecification effect estimated from survey d . Skinner, Holt and Smith describe the misspecification effect as a measure of how calculating the variance assuming that observations are independent and identically distributed under complex sampling, over- or underestimates the true variance.^[23] We evaluated a weighting strategy considering both the variance and the misspecification effect for combining estimates from complex samples. The weight combining misspecification effect and variance was calculated as

$$w_d = \frac{meff_d}{var(\hat{\theta}_d)}$$

where $meff_d$ is the misspecification effect in survey d . Under simple random sampling, $meff_d = 1$ hence the weight becomes $w_d = \frac{1}{var(\hat{\theta}_d)}$. The combined estimate may then be expressed as

$$\hat{\theta}_c = \sum_{d=1}^D \left(\frac{\frac{meff_d}{var(\hat{\theta}_d)} \times \hat{\theta}_d}{\sum_{d=1}^D \frac{meff_d}{var(\hat{\theta}_d)}} \right). \quad (3.10)$$

The same criteria used to assess estimators resulting from different weighting strategies under simple random sampling were used to assess estimators under complex sampling.

Table 3.1 gives a summary of studies where surveys were combined, together with the aims and statistical methods used.

Table 3. 1: Table Summary of the literature review where surveys were combined

Authors	Year of Publication	Journal	Aims of the paper	Surveys combined	Statistical methods used	Sampling considered? Yes/No	Software used	Area of application
Roberts G, Binder D	2009	JSM	Study whether gay and bisexual men differ in their use of health care.	Canadian community health surveys of 2003 and 2005	Pooled Unmodified weights Logistic regression, time indicator	Yes	Not specified	Health Sciences
Fox K	2010	JSM	Give a framework for the meta-analysis of complex survey data	simulations	Fixed model, inverse variance Random effects model Super population estimate	yes	SAS	simulations
Dhir A, Chen S, Nieminen M.	2015	Computers & Education	Predict compulsive internet use in adolescents	A. Private schools, B. same private schools later time point, C. public school	Separate analyses Factor analysis	NO	SPSS	Social/ Education/ computers

Graubard, BI Korn, EL	1999	Journal of the national cancer institute	Longitudinal associations of risk factors and development of cancer. Cross-sectional associations of risk factors and various outcomes	NHANES III	Design based analysis Naïve analysis	yes	SUDAAN	Social/Health
Lu Y	2014	Communication in Statistics	Estimation of regression coefficients in dual frame surveys	Simulated data	Regression Pseudo Maximum likelihood estimation Cross validation Linearization	yes	Not specified	simulations
Qi Dong, Michael R. Elliott and Trivellore E. Raghunathan	2014	Survey Methodology	Combining surveys to obtain more accurate estimates	NHIS, MEPS and BRFSS	Bayesian bootstrap to generate synthetic population. Finite population Bayesian bootstrap (FPBB) Multiple imputation combining rules	Yes	Not specified	Social/Health

Rao JNK, Kim JK	2013	Proceedings of the 59 th ISI conference, Hong Kong	Estimating totals and regression parameters from two independent surveys	Simulated data	Fractional imputation Instrumental variables, regression analysis Statistical matching	yes	Not specified	simulations
Qualité L, Tillé Y	2008	Survey methodology	Variance estimation of changes in repeated surveys and its application to the Swiss survey of value added	Swiss Survey of value added: 1999, 2000,2001	Robustification regression	yes	Not specified	Social/Health
Seho Park, Jae Kwang Kim, Diana Stukel	2017	Metron	Integrating mean estimates from two surveys	The Food and Nutrition Technical Assistance III Project (Guatemala) baseline and interim surveys (Overlap in geography)	Measurement error models Structural equation models	yes	Not specified	Social/Health

In chapter 4, simulation studies were carried out to compare the different weighting strategies under different data distributions, variances and sample sizes.

Chapter 4

Simulation study 1

The weighting strategy used when combining surveys has an important impact on the estimation of population parameters, regardless of the circumstances under which the combination of surveys occurs.^[43] In this chapter, results from repeated simulations are presented. A finite population of size 100 000 was generated, multiple samples were selected from this finite population without replacement, and sample estimates were then calculated using these samples. The sample estimates were combined using different weighting strategies which have been used in the literature; and some that are being proposed in this thesis.

The coefficient of variation (cv), skewness, kurtosis and the D'Agostino-Pearson test statistic, allowing for the sampling design, were considered for use in the weighting strategies when combining surveys in this study. To the best of my knowledge, these statistics have not been previously considered as weighting strategies for combining surveys. These were compared to the sample size and the inverse variance studied by Fox.^[29] The combined estimates were calculated as shown in equations 3.1 to 3.4.

The weighting strategies were compared with respect to their mean square errors (equation 3.6), coefficients of variation (equation 3.8) and relative biases (equation 3.5). The weighting strategy that resulted in the least mean square error was deemed to be performing better than the other weighting strategies. Ninety five percent confidence intervals (equation 3.7) were also calculated to check for the inclusion of the population parameter.

4.1 Simulation set up

In this simulation, we initially replicated the simulation done by Fox (2010) where convergence of combined estimates from a sequence of simple random samples using three different weighting strategies was assessed. We then proceeded to do simulations where the weighting strategies were compared using the mean square error and other properties of estimators discussed above.

4.1.1 Comparison of three weighting strategies

A finite population of size $N=100\,000$ was simulated from a super-population normal distribution model with mean 1 and variance 25. The mean of the simulated finite population was calculated. Initially, simple random sampling was used to select $d=500$ samples of size $n=500$. Estimates of the population mean were then calculated from each sample. The sample estimates were then combined using the three weighting strategies under consideration and convergence (as $d \rightarrow \infty$) was assessed numerically for the three weighting strategies. The weighting strategies compared in this study are; the inverse sample size, the inverse of the variance and weight of one, which is equivalent to a simple average of the estimates. The results are shown in Figure 7 below.

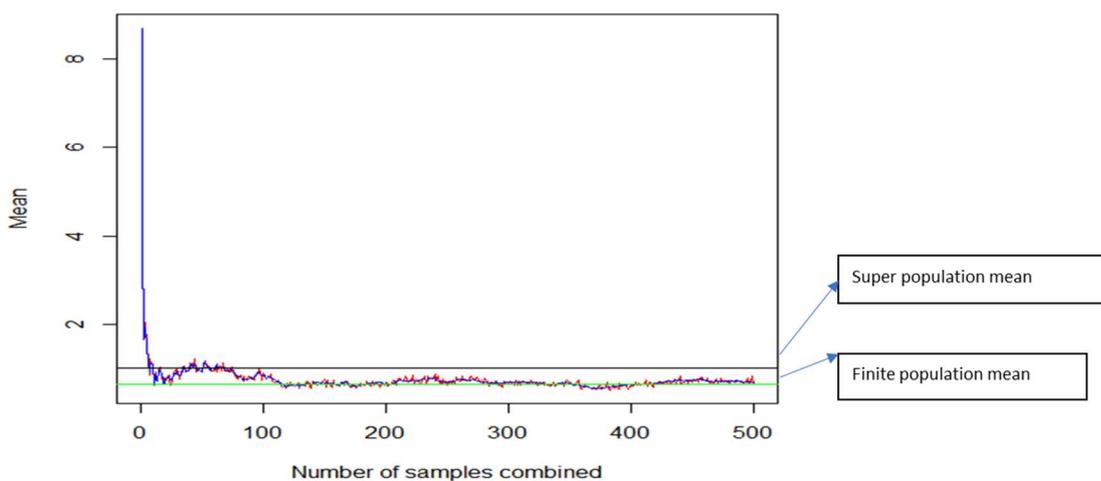


Figure 7: Convergence of combined estimates as number of samples increases

The y axis represents the mean of the variable of interest and the x axis represents the number of sample surveys combined to obtain the estimate. The graphs for the three strategies were indistinguishable. This may be because equal sample sizes were used, and the variances were very similar. The graphs showed convergence towards the finite population mean rather than the super population mean as the number of samples combined increased. These results are comparable to the results obtained by Fox and theoretical expectations. The combined means converge towards the finite population mean as the number of samples combined increases. In practise, only a few surveys are available to be combined. We take this into consideration in subsequent simulations, only combining two, three, five, ten, fifteen or 20 samples.

4.1.2 Comparison of other weighting strategies

Other weighting strategies were then considered for various reasons. Firstly, the finite population was generated from a normal distribution super-population model. It was considered that a finite population could emanate from other superpopulation models hence, finite populations generated from other super-population models have also been considered. Such other models were considered for generating skewed or peaked finite populations and hence the inverse of skewness and the inverse of kurtosis were considered for weighting the samples. The combined estimates were calculated as in equations 3.2 and 3.3 respectively. A weighting strategy which combines skewness and kurtosis was also considered in the form of the inverse of the D'Agostino-Pearson K^2 statistic, with the combined estimate being calculated as in equation 3.4. The inverse of the coefficient of variation was also considered as a weighting strategy. The combined estimate was calculated as in equation 3.1. The simulation set up is summarized in the schematic diagram below.

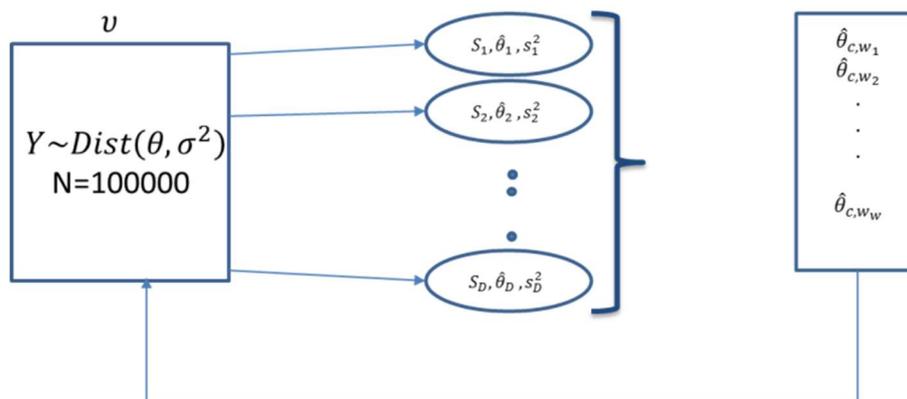


Figure 8: Comparing weighting strategies for combining surveys under simple random sampling.

In the diagram, θ and σ^2 are the mean and variance of the population, S_d represents the sample $d, d = 1, \dots, D$, $\hat{\theta}_d$ and s_d^2 are the point and variance estimates calculated from sample d . The combined estimate estimated using weighting strategy $w, w = 1, \dots, w$ is denoted $\hat{\theta}_{c,w_w}$. This simulation process is repeated $K=1000$ times.

It was also considered that the finite population size N and the sample size n , used in the simulation might influence the behaviour of the weighting strategy used to combine samples. Therefore, three finite population sizes and three sample sizes were investigated in the simulations. The population sizes investigated were 100 000, 1 000 000 and 5 000 000 and the sample sizes were 1000, 2500 and 5000.

Various super-population models were considered for generating the finite populations, that is, the normal distribution, the t-distribution with three degrees of freedom, the t-distribution with four degrees of freedom and their skewed counterparts. A mean of 100 and three different values for the variance were considered, that is, $\sigma^2 = 25$, $\sigma^2 = 49$ and $\sigma^2 = 81$. Different types of sampling were considered for selecting the samples to be combined. That is, simple random sampling, simple stratified sampling and simple cluster random sampling. Essentially, a finite population of size $N = 100\ 000$, $1\ 000\ 000$ or $5\ 000\ 000$ was simulated from a super-population with a given distribution. Initially, simple random sampling was used to select $D = 2, 3, 5, 10, 15$ or 20 samples of size $n = 1000$, 5000 or 10000 . Estimates of the mean were then calculated from each sample. The sample estimates of the mean were then combined using the weighting strategies under consideration.

To compare the weighting strategies, the mean square error (MSE), equation 3.6, relative bias, equation 3.5, confidence intervals, equation 3.7 and coefficient of variation, equation 3.8, were calculated. The weighting strategy resulting in the smallest MSE was chosen as the best strategy for the scenario. Preliminary simulations showed that using a weight of one, inverse sample size and sample size yielded identical results, when the same sample size was used in all the samples. Hence, only results of using the inverse of the sample size as a weighting strategy were presented here. Fox does not specifically touch on this but the graphs illustrating convergence as the number of samples combined increases were indistinguishable when she used a weight of one and the inverse of sample size.

4.2 Simulation results under simple random sampling

Results of the comparisons of weighting strategies were presented for each superpopulation model in turn. Tables of comparison were presented for the different weighting strategies, sample sizes and number of samples combined.

4.2.1 Simulation results for a finite population generated using the normal distribution, under simple random sampling

The normal distribution was used to generate the finite population. The results of comparing the weighting strategies where $D=2,3,5,10,15,20$ samples were combined for population size $N=1\ 000\ 000$ are presented below. Samples of sizes 1000, 5000 and 10 000 selected from a superpopulation distributed as $N(100,25)$, are presented here, with the rest of the scenarios presented in Appendix A1.

The weighting strategies were compared, using a sample size $n=1000$ in table 4.1.

Table 4. 1: Combining D samples of size n=1000, from a population of size N=1000 000, distributed as N(100,25), $\theta=99.996$.

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	$SE(\hat{\theta}_c)$ $\times 10^{-1}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-3}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-2}$
						LB	UB	
1		99,997	1,300	1,590	5,029	99,987	100,007	2,529
2	1/n	99,996	0,413	1,128	3,568	99,989	100,003	1,273
	$(Var(\hat{\theta}_a))^{-1}$	99,996	0,284	1,130	3,572	99,989	100,003	1,276
	$(CV(\hat{\theta}_a))^{-1}$	99,996	0,468	1,129	3,569	99,989	100,003	1,274
	$(skewness(s_{dc}))^{-1}$	99,996	0,188	1,259	3,980	99,988	100,003	1,584
	$(kurtosis(k_{dc}))^{-1}$	99,996	0,343	1,132	3,580	99,989	100,003	1,281
	$(K^2)^{-1}$	99,996	0,800	1,291	4,081	99,988	100,004	1,666
3	1/n	99,995	-0,024	0,927	2,930	99,990	100,001	0,859
	$(Var(\hat{\theta}_a))^{-1}$	99,995	-0,123	0,928	2,934	99,990	100,001	0,861
	$(CV(\hat{\theta}_a))^{-1}$	99,996	0,083	0,927	2,931	99,990	100,001	0,859
	$(skewness(s_{dc}))^{-1}$	99,995	-0,706	1,141	3,607	99,988	100,002	1,301
	$(kurtosis(k_{dc}))^{-1}$	99,995	-0,076	0,928	2,934	99,990	100,001	0,861
	$(K^2)^{-1}$	99,994	-1,203	1,141	3,609	99,987	100,001	1,302
5	1/n	99,996	0,320	0,721	2,280	99,991	100,000	0,520
	$(Var(\hat{\theta}_a))^{-1}$	99,996	0,234	0,722	2,283	99,991	100,000	0,521
	$(CV(\hat{\theta}_a))^{-1}$	99,996	0,469	0,721	2,280	99,991	100,000	0,520
	$(skewness(s_{dc}))^{-1}$	99,995	-0,978	0,989	3,126	99,988	100,001	0,977
	$(kurtosis(k_{dc}))^{-1}$	99,996	0,319	0,722	2,284	99,991	100,000	0,522
	$(K^2)^{-1}$	99,993	-2,911	0,976	3,086	99,987	99,999	0,953
10	1/n	99,998	2,187	0,506	1,602	99,995	100,001	0,257
	$(Var(\hat{\theta}_a))^{-1}$	99,998	2,137	0,506	1,601	99,994	100,001	0,257
	$(CV(\hat{\theta}_a))^{-1}$	99,998	2,383	0,506	1,601	99,995	100,001	0,257
	$(skewness(s_{dc}))^{-1}$	100,001	6,015	0,868	2,744	99,996	100,007	0,756
	$(kurtosis(k_{dc}))^{-1}$	99,998	2,133	0,506	1,602	99,994	100,001	0,257
	$(K^2)^{-1}$	99,998	2,703	0,816	2,581	99,993	100,003	0,667
15	1/n	99,996	0,443	0,402	1,270	99,993	99,998	0,161
	$(Var(\hat{\theta}_a))^{-1}$	99,996	0,430	0,402	1,270	99,993	99,998	0,161
	$(CV(\hat{\theta}_a))^{-1}$	99,996	0,673	0,402	1,270	99,994	99,999	0,161
	$(skewness(s_{dc}))^{-1}$	99,999	3,612	0,824	2,607	99,994	100,004	0,681
	$(kurtosis(k_{dc}))^{-1}$	99,996	0,394	0,402	1,272	99,993	99,998	0,162
	$(K^2)^{-1}$	99,995	-0,844	0,746	2,359	99,990	99,999	0,556
20	1/n	99,995	-0,464	0,346	1,095	99,993	99,997	0,120
	$(Var(\hat{\theta}_a))^{-1}$	99,995	-0,474	0,347	1,097	99,993	99,997	0,120
	$(CV(\hat{\theta}_a))^{-1}$	99,995	-0,229	0,347	1,096	99,993	99,997	0,120
	$(skewness(s_{dc}))^{-1}$	100,000	4,025	0,763	2,412	99,995	100,004	0,583
	$(kurtosis(k_{dc}))^{-1}$	99,995	-0,556	0,347	1,096	99,993	99,997	0,120
	$(K^2)^{-1}$	99,995	-0,711	0,687	2,172	99,991	99,999	0,472

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, $SE(\hat{\theta}_c)$ =standard error of the combined estimate, $CV(\hat{\theta}_c)$ (%) = coefficient of variation, $CI_{95\%}(\hat{\theta}_c)$ = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

The MSE value was higher for the estimate calculated using one sample compared to combining samples for all the weighting strategies considered. The MSE decreased as the number of samples combined was increased. The weighting strategies resulted in MSE values that were close to each other except the inverse of skewness and the inverse of the D'Agostino-Pearson statistic.

Table 4.2 presents a comparison of the weighting strategies when the sample size was increased to $n=5000$.

Table 4. 2: Combining D samples of size n=5000, from a population of size N=1000 000, distributed as N(100,25), $\theta=99.996$.

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-4}$	$SE(\hat{\theta}_c)$ $\times 10^{-2}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-3}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-3}$
						LB	UB	
1		99,997	-3,678	7,284	2,303	99,992	100,001	5,305
2	1/n	99,997	-2,069	5,162	1,632	99,994	100,000	2,664
	$(Var(\hat{\theta}_a))^{-1}$	99,997	-2,123	5,162	1,632	99,994	100,000	2,664
	$(CV(\hat{\theta}_a))^{-1}$	99,997	-1,845	5,162	1,632	99,994	100,000	2,664
	$(skewness(s_{dc}))^{-1}$	99,997	-2,854	5,837	1,846	99,993	100,000	3,407
	$(kurtosis(k_{dc}))^{-1}$	99,997	-2,153	5,164	1,633	99,994	100,000	2,667
3	1/n	99,997	1,068	4,170	1,319	99,995	100,000	1,739
	$(Var(\hat{\theta}_a))^{-1}$	99,997	0,859	4,169	1,318	99,995	100,000	1,738
	$(CV(\hat{\theta}_a))^{-1}$	99,997	1,313	4,169	1,318	99,995	100,000	1,738
	$(skewness(s_{dc}))^{-1}$	99,998	3,651	5,120	1,619	99,994	100,001	2,622
	$(kurtosis(k_{dc}))^{-1}$	99,997	1,271	4,175	1,320	99,995	100,000	1,743
5	1/n	99,998	5,052	3,096	0,979	99,996	100,000	0,959
	$(Var(\hat{\theta}_a))^{-1}$	99,998	5,026	3,095	0,979	99,996	100,000	0,958
	$(CV(\hat{\theta}_a))^{-1}$	99,998	5,441	3,095	0,979	99,996	100,000	0,958
	$(skewness(s_{dc}))^{-1}$	99,998	8,539	4,406	1,393	99,995	100,001	1,942
	$(kurtosis(k_{dc}))^{-1}$	99,998	5,175	3,099	0,980	99,996	100,000	0,961
10	1/n	99,997	-0,158	2,171	0,686	99,996	99,998	0,471
	$(Var(\hat{\theta}_a))^{-1}$	99,997	0,060	2,171	0,687	99,996	99,999	0,471
	$(CV(\hat{\theta}_a))^{-1}$	99,997	0,399	2,171	0,686	99,996	99,999	0,471
	$(skewness(s_{dc}))^{-1}$	99,997	-6,351	3,698	1,169	99,994	99,999	1,368
	$(kurtosis(k_{dc}))^{-1}$	99,997	-0,116	2,173	0,687	99,996	99,998	0,472
15	1/n	99,998	5,049	1,797	0,568	99,997	99,999	0,323
	$(Var(\hat{\theta}_a))^{-1}$	99,998	5,217	1,797	0,568	99,997	99,999	0,323
	$(CV(\hat{\theta}_a))^{-1}$	99,998	5,593	1,797	0,568	99,997	99,999	0,323
	$(skewness(s_{dc}))^{-1}$	99,997	-3,834	3,413	1,079	99,995	99,999	1,165
	$(kurtosis(k_{dc}))^{-1}$	99,998	5,065	1,799	0,569	99,997	99,999	0,324
20	1/n	99,998	4,169	1,531	0,484	99,997	99,999	0,235
	$(Var(\hat{\theta}_a))^{-1}$	99,998	4,333	1,530	0,484	99,997	99,999	0,234
	$(CV(\hat{\theta}_a))^{-1}$	99,998	4,719	1,531	0,484	99,997	99,999	0,234
	$(skewness(s_{dc}))^{-1}$	99,997	0,135	3,310	1,047	99,995	99,999	1,096
	$(kurtosis(k_{dc}))^{-1}$	99,998	4,186	1,533	0,485	99,997	99,999	0,235

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, $SE(\hat{\theta}_c)$ =standard error of the combined estimate, $CV(\hat{\theta}_c)$ (%) = coefficient of variation, $CI_{95\%}(\hat{\theta}_c)$ = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Combining samples resulted in much better estimates compared to using one sample. Weighting strategies generally performed similarly with MSE values close together. Using the inverse of skewness as a weighting strategy resulted in higher

MSE values. The MSE values became smaller as the number of samples combined increased.

Table 4. 3: Combining D samples of size n=10000, from a population of size N=1000 000, distributed as N(100,25), $\theta=99.996$.

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-3}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		99,995	-2,420	4,978	1,574	99,992	99,998	2,484
2	1/n	99,996	-0,920	3,609	1,141	99,994	99,998	1,303
	$(Var(\hat{\theta}_a))^{-1}$	99,996	-0,909	3,610	1,142	99,994	99,998	1,304
	$(CV(\hat{\theta}_a))^{-1}$	99,996	-0,902	3,610	1,141	99,994	99,998	1,304
	$(skewness(s_{dc}))^{-1}$	99,996	-0,734	3,946	1,248	99,994	99,999	1,557
	$(kurtosis(k_{dc}))^{-1}$	99,996	-0,911	3,611	1,142	99,994	99,998	1,304
3	1/n	99,997	-0,344	2,902	0,918	99,995	99,999	0,842
	$(Var(\hat{\theta}_a))^{-1}$	99,997	-0,327	2,902	0,918	99,995	99,999	0,842
	$(CV(\hat{\theta}_a))^{-1}$	99,997	-0,319	2,902	0,918	99,995	99,999	0,842
	$(skewness(s_{dc}))^{-1}$	99,996	-0,982	3,440	1,088	99,994	99,998	1,184
	$(kurtosis(k_{dc}))^{-1}$	99,997	-0,324	2,902	0,918	99,995	99,999	0,842
5	1/n	99,997	0,165	2,276	0,720	99,996	99,999	0,518
	$(Var(\hat{\theta}_a))^{-1}$	99,997	0,180	2,274	0,719	99,996	99,999	0,517
	$(CV(\hat{\theta}_a))^{-1}$	99,997	0,192	2,275	0,719	99,996	99,999	0,518
	$(skewness(s_{dc}))^{-1}$	99,997	-0,382	3,085	0,976	99,995	99,999	0,952
	$(kurtosis(k_{dc}))^{-1}$	99,997	0,171	2,275	0,719	99,996	99,999	0,518
10	1/n	99,997	0,037	1,555	0,492	99,996	99,998	0,242
	$(Var(\hat{\theta}_a))^{-1}$	99,997	0,040	1,555	0,492	99,996	99,998	0,242
	$(CV(\hat{\theta}_a))^{-1}$	99,997	0,060	1,555	0,492	99,996	99,998	0,242
	$(skewness(s_{dc}))^{-1}$	99,997	-0,409	2,633	0,833	99,995	99,998	0,693
	$(kurtosis(k_{dc}))^{-1}$	99,997	0,046	1,555	0,492	99,996	99,998	0,242
15	1/n	99,997	0,194	1,275	0,403	99,997	99,998	0,163
	$(Var(\hat{\theta}_a))^{-1}$	99,997	0,198	1,275	0,403	99,997	99,998	0,163
	$(CV(\hat{\theta}_a))^{-1}$	99,997	0,219	1,275	0,403	99,997	99,998	0,163
	$(skewness(s_{dc}))^{-1}$	99,997	0,064	2,546	0,805	99,996	99,999	0,648
	$(kurtosis(k_{dc}))^{-1}$	99,997	0,201	1,276	0,403	99,997	99,998	0,163
20	1/n	99,997	0,104	1,118	0,353	99,997	99,998	0,125
	$(Var(\hat{\theta}_a))^{-1}$	99,997	0,107	1,117	0,353	99,997	99,998	0,125
	$(CV(\hat{\theta}_a))^{-1}$	99,997	0,129	1,117	0,353	99,997	99,998	0,125
	$(skewness(s_{dc}))^{-1}$	99,997	-0,239	2,451	0,775	99,995	99,998	0,601
	$(kurtosis(k_{dc}))^{-1}$	99,997	0,111	1,118	0,354	99,997	99,998	0,125

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

The weighting strategies performed similarly, except the inverse of skewness which mostly resulted in MSE values that were larger than the other weighting strategies for the same number of samples combined. This was expected since the data were not skewed. Using only one sample resulted in a much larger MSE value compared to combining the samples using any of the weighting strategies considered.

Weighting strategies were compared across superpopulation variances and sizes. The results of the comparison for the normal distribution with sample size $n=1000$, where two samples were combined are presented in table 4.4 below.

Table 4. 4: MSE comparing weighting strategies for a normal distribution superpopulation, $D=2$ and $n=1000$ across variances and population sizes

Method	variance	N=100000 $\times 10^{-2}$	N=1000000 $\times 10^{-2}$	N=5000000 $\times 10^{-2}$
$1/n$	25	1,297	1,273	1,317
$(Var(\hat{\theta}_d))^{-1}$		1,300	1,276	1,324
$(CV(\hat{\theta}_d))^{-1}$		1,297	1,273	1,317
$(skewness(s_{dc}))^{-1}$		1,300	1,274	1,320
$(kurtosis(k_{dc}))^{-1}$		1,602	1,584	1,735
$(K^2)^{-1}$		1,298	1,281	1,315
$1/n$	49	2,343	2,655	2,581
$(Var(\hat{\theta}_d))^{-1}$		2,336	2,645	2,595
$(CV(\hat{\theta}_d))^{-1}$		2,343	2,655	2,581
$(skewness(s_{dc}))^{-1}$		2,339	2,650	2,588
$(kurtosis(k_{dc}))^{-1}$		3,094	3,274	3,400
$(K^2)^{-1}$		2,356	2,645	2,576
$1/n$	81	3,873	4,389	4,267
$(Var(\hat{\theta}_d))^{-1}$		3,862	4,372	4,290
$(CV(\hat{\theta}_d))^{-1}$		3,873	4,389	4,267
$(skewness(s_{dc}))^{-1}$		3,866	4,380	4,279
$(kurtosis(k_{dc}))^{-1}$		5,114	5,412	5,621
$(K^2)^{-1}$		3,895	4,372	4,259

$\hat{\theta}_d$ = estimate calculated from each individual survey

The MSE values increased as the variance of the superpopulation was increased. However, increasing the superpopulation variance appeared not to influence the choice of the weighting strategy, that is, the weighting strategy with the lowest MSE value, as illustrated in table 4.4. This phenomenon was observed for all sample sizes

considered as well as all numbers of samples combined. The choice of weighting strategy was the same for $N=100\ 000$ and $1\ 000\ 000$.

4.2.2 Simulation results for a finite population generated using the skewed normal distribution, under simple random sampling

The skewed normal distribution was used to generate the finite population. A location parameter of -1.165, scale parameter of 1.53 and an asymmetry parameter 3 were used to generate strongly asymmetric data with mean 100 and variance 25, 49 and 81. A population size of one million was used with sample sizes 1 000, 5 000 and 10 000. The sampling process was repeated 1 000 times. The results of comparing the weighting strategies are presented below in tables 4.5 to 4.7.

Table 4. 5 Scenario 1: Combining D samples of size n=1000, from a population of size N=1000 000, distributed as the skewed normal distribution SN(100,25), $\theta=100.082$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		100,074	-7,974	5,131	1,621	100,064	100,084	2,639
2	1/n	100,078	-4,495	3,651	1,154	100,070	100,085	1,335
	$(Var(\hat{\theta}_a))^{-1}$	100,076	-6,478	3,657	1,156	100,068	100,083	1,342
	$(CV(\hat{\theta}_a))^{-1}$	100,077	-5,353	3,653	1,154	100,070	100,084	1,337
	$(skewness(s_{dc}))^{-1}$	100,078	-3,771	3,671	1,160	100,071	100,086	1,349
	$(kurtosis(k_{dc}))^{-1}$	100,079	-3,631	3,663	1,157	100,071	100,086	1,343
	$(K^2)^{-1}$	100,080	-2,507	3,732	1,179	100,072	100,087	1,393
3	1/n	100,080	-2,054	2,963	0,936	100,074	100,086	0,878
	$(Var(\hat{\theta}_a))^{-1}$	100,078	-4,461	2,974	0,940	100,072	100,084	0,887
	$(CV(\hat{\theta}_a))^{-1}$	100,079	-3,086	2,967	0,938	100,073	100,085	0,882
	$(skewness(s_{dc}))^{-1}$	100,081	-1,319	2,984	0,943	100,075	100,087	0,891
	$(kurtosis(k_{dc}))^{-1}$	100,081	-1,173	2,974	0,940	100,075	100,087	0,885
	$(K^2)^{-1}$	100,082	-0,140	3,047	0,963	100,076	100,088	0,928
5	1/n	100,080	-1,680	2,305	0,728	100,076	100,085	0,531
	$(Var(\hat{\theta}_a))^{-1}$	100,078	-4,440	2,310	0,730	100,073	100,082	0,536
	$(CV(\hat{\theta}_a))^{-1}$	100,079	-2,857	2,307	0,729	100,075	100,084	0,533
	$(skewness(s_{dc}))^{-1}$	100,081	-0,712	2,326	0,735	100,077	100,086	0,541
	$(kurtosis(k_{dc}))^{-1}$	100,082	-0,622	2,314	0,731	100,077	100,086	0,535
	$(K^2)^{-1}$	100,083	0,713	2,387	0,754	100,078	100,088	0,570
10	1/n	100,081	-1,325	1,600	0,506	100,078	100,084	0,256
	$(Var(\hat{\theta}_a))^{-1}$	100,078	-4,447	1,602	0,506	100,075	100,081	0,258
	$(CV(\hat{\theta}_a))^{-1}$	100,079	-2,659	1,600	0,506	100,076	100,083	0,257
	$(skewness(s_{dc}))^{-1}$	100,082	-0,320	1,607	0,508	100,079	100,085	0,258
	$(kurtosis(k_{dc}))^{-1}$	100,082	-0,329	1,604	0,507	100,079	100,085	0,257
	$(K^2)^{-1}$	100,083	1,221	1,649	0,521	100,080	100,087	0,272
15	1/n	100,080	-1,648	1,278	0,404	100,078	100,083	0,164
	$(Var(\hat{\theta}_a))^{-1}$	100,077	-4,913	1,275	0,403	100,075	100,080	0,165
	$(CV(\hat{\theta}_a))^{-1}$	100,079	-3,043	1,276	0,403	100,077	100,082	0,164
	$(skewness(s_{dc}))^{-1}$	100,082	-0,572	1,281	0,405	100,079	100,084	0,164
	$(kurtosis(k_{dc}))^{-1}$	100,082	-0,594	1,279	0,404	100,079	100,084	0,164
	$(K^2)^{-1}$	100,083	1,055	1,316	0,416	100,081	100,086	0,173
20	1/n	100,081	-1,085	1,102	0,348	100,079	100,083	0,122
	$(Var(\hat{\theta}_a))^{-1}$	100,078	-4,387	1,100	0,348	100,076	100,080	0,123
	$(CV(\hat{\theta}_a))^{-1}$	100,080	-2,495	1,101	0,348	100,077	100,082	0,122
	$(skewness(s_{dc}))^{-1}$	100,082	0,110	1,109	0,351	100,080	100,084	0,123
	$(kurtosis(k_{dc}))^{-1}$	100,082	-0,007	1,105	0,349	100,080	100,084	0,122
	$(K^2)^{-1}$	100,084	1,859	1,141	0,360	100,082	100,086	0,131

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Combining samples resulted in much smaller MSE values compared to using only one sample. Using the inverse of the D'Agostino-Pearson statistic resulted in slightly higher MSE values for all the numbers of samples combined. The weighting strategies resulted in similar MSE values which got smaller as the number of samples combined increased. All the weighting strategies appeared to underestimate the population parameter most of the time as shown by the negative values of the relative bias, though the population parameter was included in the confidence interval.

Table 4. 6: Scenario 2: Combining D samples of size n=5000, from a population of size N=1000 000, distributed as the skewed normal distribution SN(100,25), $\theta=100.082$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		100,081	-1,239	2,241	7,081	100,077	100,085	5,024
2	1/n	100,081	-1,145	1,596	5,042	100,078	100,084	2,547
	$(Var(\hat{\theta}_a))^{-1}$	100,081	-1,479	1,595	5,040	100,078	100,084	2,547
	$(CV(\hat{\theta}_a))^{-1}$	100,081	-1,286	1,595	5,041	100,078	100,084	2,547
	$(skewness(s_{dc}))^{-1}$	100,081	-1,077	1,599	5,052	100,078	100,084	2,557
	$(kurtosis(k_{dc}))^{-1}$	100,081	-1,114	1,597	5,047	100,078	100,084	2,552
3	1/n	100,081	-1,431	1,291	4,080	100,078	100,083	1,669
	$(Var(\hat{\theta}_a))^{-1}$	100,080	-1,879	1,289	4,074	100,078	100,083	1,666
	$(CV(\hat{\theta}_a))^{-1}$	100,081	-1,620	1,290	4,077	100,078	100,083	1,667
	$(skewness(s_{dc}))^{-1}$	100,081	-1,248	1,295	4,091	100,078	100,083	1,678
	$(kurtosis(k_{dc}))^{-1}$	100,081	-1,310	1,293	4,086	100,078	100,083	1,674
5	1/n	100,081	-1,046	1,026	3,242	100,079	100,083	1,054
	$(Var(\hat{\theta}_a))^{-1}$	100,081	-1,575	1,025	3,240	100,079	100,083	1,054
	$(CV(\hat{\theta}_a))^{-1}$	100,081	-1,270	1,026	3,241	100,079	100,083	1,054
	$(skewness(s_{dc}))^{-1}$	100,081	-0,871	1,028	3,247	100,079	100,083	1,057
	$(kurtosis(k_{dc}))^{-1}$	100,081	-0,919	1,027	3,244	100,079	100,083	1,055
10	1/n	100,081	-0,826	0,736	2,325	100,080	100,083	0,542
	$(Var(\hat{\theta}_a))^{-1}$	100,081	-1,431	0,735	2,324	100,079	100,082	0,543
	$(CV(\hat{\theta}_a))^{-1}$	100,081	-1,084	0,736	2,324	100,080	100,082	0,542
	$(skewness(s_{dc}))^{-1}$	100,082	-0,585	0,738	2,333	100,080	100,083	0,545
	$(kurtosis(k_{dc}))^{-1}$	100,082	-0,625	0,737	2,330	100,080	100,083	0,544
15	1/n	100,082	-0,460	0,583	1,843	100,081	100,083	0,341
	$(Var(\hat{\theta}_a))^{-1}$	100,081	-1,074	0,584	1,844	100,080	100,082	0,342
	$(CV(\hat{\theta}_a))^{-1}$	100,081	-0,721	0,583	1,844	100,080	100,083	0,341
	$(skewness(s_{dc}))^{-1}$	100,082	-0,218	0,586	1,851	100,081	100,083	0,343
	$(kurtosis(k_{dc}))^{-1}$	100,082	-0,258	0,585	1,847	100,081	100,083	0,342
20	1/n	100,082	-0,441	0,500	1,580	100,081	100,083	0,250
	$(Var(\hat{\theta}_a))^{-1}$	100,081	-1,064	0,500	1,581	100,080	100,082	0,251
	$(CV(\hat{\theta}_a))^{-1}$	100,081	-0,705	0,500	1,580	100,080	100,082	0,251
	$(skewness(s_{dc}))^{-1}$	100,082	-0,178	0,502	1,587	100,081	100,083	0,252
	$(kurtosis(k_{dc}))^{-1}$	100,082	-0,224	0,501	1,584	100,081	100,083	0,251

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

When a sample size of 5000 was used, all the weighting strategies performed similarly, resulting in MSE values which were of comparable magnitude. In all the cases, combining samples resulted in lower MSE values compared to using

estimates calculated using one sample. The resulting MSE values decreased in magnitude as more samples were combined.

Table 4. 7: Scenario 3: Combining D samples of size n=10 000, from a population of size N=1000 000, distributed as the skewed normal distribution, SN(100,25), $\theta=100.082$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-4}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		100,082	-3,639	1,595	5,040	100,079	100,085	2,545
2	1/n	100,082	-3,914	1,143	3,612	100,080	100,084	1,307
	$(Var(\hat{\theta}_a))^{-1}$	100,082	-5,544	1,143	3,611	100,079	100,084	1,306
	$(CV(\hat{\theta}_a))^{-1}$	100,082	-4,603	1,143	3,611	100,079	100,084	1,306
	$(skewness(s_{dc}))^{-1}$	100,082	-3,734	1,144	3,615	100,080	100,084	1,309
	$(kurtosis(k_{dc}))^{-1}$	100,082	-3,678	1,143	3,613	100,080	100,084	1,307
3	1/n	100,083	4,985	0,953	3,010	100,081	100,085	0,908
	$(Var(\hat{\theta}_a))^{-1}$	100,082	2,930	0,952	3,008	100,081	100,084	0,907
	$(CV(\hat{\theta}_a))^{-1}$	100,083	4,122	0,952	3,009	100,081	100,084	0,907
	$(skewness(s_{dc}))^{-1}$	100,083	5,440	0,953	3,012	100,081	100,085	0,909
	$(kurtosis(k_{dc}))^{-1}$	100,083	5,408	0,953	3,011	100,081	100,085	0,908
5	1/n	100,082	-0,073	0,726	2,295	100,081	100,084	0,528
	$(Var(\hat{\theta}_a))^{-1}$	100,082	-2,698	0,726	2,295	100,080	100,083	0,528
	$(CV(\hat{\theta}_a))^{-1}$	100,082	-1,191	0,726	2,295	100,081	100,083	0,528
	$(skewness(s_{dc}))^{-1}$	100,082	0,550	0,727	2,299	100,081	100,084	0,529
	$(kurtosis(k_{dc}))^{-1}$	100,082	0,558	0,727	2,297	100,081	100,084	0,528
10	1/n	100,083	5,640	0,498	1,574	100,082	100,084	0,249
	$(Var(\hat{\theta}_a))^{-1}$	100,082	2,747	0,498	1,575	100,081	100,083	0,248
	$(CV(\hat{\theta}_a))^{-1}$	100,083	4,419	0,498	1,574	100,082	100,084	0,249
	$(skewness(s_{dc}))^{-1}$	100,083	6,852	0,499	1,575	100,082	100,084	0,249
	$(kurtosis(k_{dc}))^{-1}$	100,083	6,684	0,498	1,575	100,082	100,084	0,249
15	1/n	100,082	3,277	0,405	1,280	100,082	100,083	0,164
	$(Var(\hat{\theta}_a))^{-1}$	100,082	0,253	0,405	1,281	100,081	100,083	0,164
	$(CV(\hat{\theta}_a))^{-1}$	100,082	1,999	0,405	1,281	100,082	100,083	0,164
	$(skewness(s_{dc}))^{-1}$	100,083	4,423	0,406	1,282	100,082	100,083	0,165
	$(kurtosis(k_{dc}))^{-1}$	100,083	4,249	0,405	1,281	100,082	100,083	0,165
20	1/n	100,082	2,232	0,355	1,122	100,082	100,083	0,126
	$(Var(\hat{\theta}_a))^{-1}$	100,082	-0,889	0,355	1,122	100,081	100,083	0,126
	$(CV(\hat{\theta}_a))^{-1}$	100,082	0,910	0,355	1,122	100,082	100,083	0,126
	$(skewness(s_{dc}))^{-1}$	100,082	3,306	0,355	1,122	100,082	100,083	0,126
	$(kurtosis(k_{dc}))^{-1}$	100,082	3,182	0,355	1,121	100,082	100,083	0,126

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

In general, the weighting strategies performed similarly. The MSE values decreased as more samples were combined. All the confidence intervals included the population parameter value.

MSE values were compared across sample sizes and number of samples combined in table 4.8 below.

Table 4. 8: MSE for the skewed Normal distribution, mean=100, N=1 000 000 across sample sizes n and number of samples combined D.

D	Method	n=1000 $\times 10^{-2}$	n=5000 $\times 10^{-3}$	n=10000 $\times 10^{-3}$
1		2,639	5,024	2,545
2	$1/n$	1,335	2,547	1,307
	$(Var(\hat{\theta}_a))^{-1}$	1,342	2,547	1,306
	$(CV(\hat{\theta}_a))^{-1}$	1,337	2,547	1,306
	$(skewness(s_{dc}))^{-1}$	1,349	2,557	1,309
	$(kurtosis(s_{dc}))^{-1}$	1,343	2,552	1,307
3	$1/n$	1,393	1,669	0,908
	$(Var(\hat{\theta}_a))^{-1}$	0,878	1,666	0,907
	$(CV(\hat{\theta}_a))^{-1}$	0,887	1,667	0,907
	$(skewness(s_{dc}))^{-1}$	0,882	1,678	0,909
	$(kurtosis(s_{dc}))^{-1}$	0,891	1,674	0,908
5	$1/n$	0,885	1,054	0,528
	$(Var(\hat{\theta}_a))^{-1}$	0,928	1,054	0,528
	$(CV(\hat{\theta}_a))^{-1}$	0,531	1,054	0,528
	$(skewness(s_{dc}))^{-1}$	0,536	1,057	0,529
	$(kurtosis(s_{dc}))^{-1}$	0,533	1,055	0,528
10	$1/n$	0,541	0,542	0,249
	$(Var(\hat{\theta}_a))^{-1}$	0,535	0,543	0,248
	$(CV(\hat{\theta}_a))^{-1}$	0,570	0,542	0,249
	$(skewness(s_{dc}))^{-1}$	0,256	0,545	0,249
	$(kurtosis(s_{dc}))^{-1}$	0,258	0,544	0,249
15	$1/n$	0,257	0,341	0,164
	$(Var(\hat{\theta}_a))^{-1}$	0,258	0,342	0,164
	$(CV(\hat{\theta}_a))^{-1}$	0,257	0,341	0,164
	$(skewness(s_{dc}))^{-1}$	0,272	0,343	0,165
	$(kurtosis(s_{dc}))^{-1}$	0,164	0,342	0,165
20	$1/n$	0,165	0,250	0,126
	$(Var(\hat{\theta}_a))^{-1}$	0,164	0,251	0,126
	$(CV(\hat{\theta}_a))^{-1}$	0,164	0,251	0,126
	$(skewness(s_{dc}))^{-1}$	0,164	0,252	0,126
	$(kurtosis(s_{dc}))^{-1}$	0,173	0,251	0,126

method=weighting strategy, D = number of samples combined.

It was clear that combining more samples resulted in better estimates, with smaller MSE values. Increasing the sample size also improved MSE values.

4.2.3 Simulation results for a finite population generated using the t- distribution with three degrees of freedom, under simple random sampling

The t- distribution with three degrees of freedom was used to generate the finite population. The results of comparing the weighting strategies for the first, second and third scenarios (sample size 1000, 5000 and 10000 respectively), are presented in tables 4.9, 4.10 and 4.11 below.

Table 4. 9: Scenario 1: Combining D samples of size n=1000, from a population of size N=1000 000, distributed as the t-distribution with three degrees of freedom, $t(3)(100,25)$, $\theta=99.998$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,999	0,687	4,992	1,579	99,989	100,009	2,493
2	1/n	99,997	-1,533	3,503	1,108	99,990	100,004	1,228
	$(Var(\hat{\theta}_a))^{-1}$	99,997	-1,422	3,443	1,089	99,990	100,004	1,185
	$(CV(\hat{\theta}_a))^{-1}$	99,997	-1,533	3,503	1,108	99,990	100,004	1,228
	$(skewness(s_{dc}))^{-1}$	99,997	-1,279	3,457	1,093	99,990	100,004	1,195
	$(kurtosis(k_{dc}))^{-1}$	99,997	-1,251	3,694	1,168	99,990	100,004	1,364
	$(K^2)^{-1}$	99,997	-1,577	3,625	1,146	99,990	100,004	1,314
3	1/n	99,997	-0,980	3,538	1,119	99,990	100,004	1,252
	$(Var(\hat{\theta}_a))^{-1}$	99,998	-0,460	2,877	0,910	99,992	100,003	0,828
	$(CV(\hat{\theta}_a))^{-1}$	99,998	-0,064	2,811	0,889	99,993	100,004	0,790
	$(skewness(s_{dc}))^{-1}$	99,998	-0,460	2,877	0,910	99,992	100,003	0,828
	$(kurtosis(k_{dc}))^{-1}$	99,998	0,019	2,826	0,894	99,993	100,004	0,799
	$(K^2)^{-1}$	99,999	0,414	3,167	1,001	99,992	100,005	1,003
5	1/n	99,998	-0,223	2,989	0,945	99,992	100,004	0,894
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,257	2,896	0,916	99,993	100,004	0,839
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,952	2,249	0,711	99,995	100,004	0,506
	$(skewness(s_{dc}))^{-1}$	100,000	1,526	2,189	0,692	99,996	100,004	0,479
	$(kurtosis(k_{dc}))^{-1}$	99,999	0,952	2,249	0,711	99,995	100,004	0,506
	$(K^2)^{-1}$	100,000	1,591	2,204	0,697	99,996	100,004	0,486
10	1/n	100,000	1,641	2,696	0,853	99,995	100,005	0,727
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,876	2,306	0,729	99,995	100,004	0,532
	$(CV(\hat{\theta}_a))^{-1}$	99,999	1,088	2,243	0,709	99,995	100,004	0,503
	$(skewness(s_{dc}))^{-1}$	99,997	-1,174	1,528	0,483	99,994	100,000	0,234
	$(kurtosis(k_{dc}))^{-1}$	99,998	-0,222	1,490	0,471	99,995	100,001	0,222
	$(K^2)^{-1}$	99,997	-1,174	1,528	0,483	99,994	100,000	0,234
15	1/n	99,998	-0,293	1,496	0,473	99,995	100,001	0,224
	$(Var(\hat{\theta}_a))^{-1}$	99,998	0,037	2,369	0,749	99,994	100,003	0,561
	$(CV(\hat{\theta}_a))^{-1}$	99,998	-0,084	1,621	0,513	99,995	100,001	0,263
	$(skewness(s_{dc}))^{-1}$	99,998	-0,040	1,585	0,501	99,995	100,001	0,251
	$(kurtosis(k_{dc}))^{-1}$	99,998	-0,385	1,243	0,393	99,995	100,000	0,155
	$(K^2)^{-1}$	99,999	0,678	1,208	0,382	99,997	100,001	0,146
20	1/n	99,998	-0,385	1,243	0,393	99,995	100,000	0,155
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,564	1,214	0,384	99,996	100,001	0,147
	$(CV(\hat{\theta}_a))^{-1}$	100,001	2,614	2,139	0,677	99,997	100,005	0,458
	$(skewness(s_{dc}))^{-1}$	100,000	1,285	1,330	0,421	99,997	100,002	0,177
	$(kurtosis(k_{dc}))^{-1}$	100,000	1,447	1,307	0,413	99,997	100,002	0,171
	$(K^2)^{-1}$	99,999	0,284	1,094	0,346	99,996	100,001	0,120

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

None of the weighting strategies consistently stood out as the number of samples combined was increased. The performance of the inverse of the D'Agostino-Pearson statistic improved as the number of samples combined was increased. It resulted in the lowest MSE value when fifteen samples were combined and when twenty samples were combined.

Table 4. 10: Scenario 2: Combining D samples of size n=5 000, from a population of size N=1000 000, distributed as the t-distribution with three degrees of freedom, $t(3)(100,25)$, $\theta=99.998$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		100,000	2,001	2,236	7,071	99,996	100,005	5,004
2	1/n	100,000	2,149	1,619	5,119	99,997	100,004	2,625
	$(Var(\hat{\theta}_a))^{-1}$	100,001	2,493	1,603	5,069	99,998	100,004	2,576
	$(CV(\hat{\theta}_a))^{-1}$	100,001	2,352	1,608	5,086	99,997	100,004	2,593
	$(skewness(s_{dc}))^{-1}$	100,001	3,080	1,759	5,564	99,998	100,005	3,105
	$(kurtosis(k_{dc}))^{-1}$	100,001	2,840	1,687	5,334	99,998	100,004	2,854
3	1/n	100,000	1,674	1,303	4,121	99,997	100,003	1,701
	$(Var(\hat{\theta}_a))^{-1}$	100,001	2,301	1,296	4,099	99,998	100,003	1,686
	$(CV(\hat{\theta}_a))^{-1}$	100,000	2,053	1,297	4,101	99,998	100,003	1,686
	$(skewness(s_{dc}))^{-1}$	100,001	3,008	1,571	4,969	99,998	100,004	2,478
	$(kurtosis(k_{dc}))^{-1}$	100,001	2,713	1,426	4,511	99,998	100,004	2,042
5	1/n	99,999	1,087	0,994	3,142	99,997	100,001	0,988
	$(Var(\hat{\theta}_a))^{-1}$	100,000	1,904	0,984	3,111	99,998	100,002	0,972
	$(CV(\hat{\theta}_a))^{-1}$	100,000	1,582	0,986	3,118	99,998	100,002	0,975
	$(skewness(s_{dc}))^{-1}$	100,001	3,094	1,344	4,252	99,999	100,004	1,817
	$(kurtosis(k_{dc}))^{-1}$	100,001	3,098	1,093	3,457	99,999	100,004	1,204
10	1/n	99,999	0,256	0,661	2,091	99,997	100,000	0,437
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,954	0,656	2,075	99,998	100,001	0,432
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,705	0,657	2,077	99,998	100,000	0,432
	$(skewness(s_{dc}))^{-1}$	100,000	1,997	1,125	3,558	99,998	100,002	1,270
	$(kurtosis(k_{dc}))^{-1}$	100,000	1,651	0,734	2,322	99,998	100,001	0,542
15	1/n	99,999	0,344	0,555	1,754	99,998	100,000	0,308
	$(Var(\hat{\theta}_a))^{-1}$	99,999	1,001	0,548	1,733	99,998	100,000	0,301
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,772	0,550	1,738	99,998	100,000	0,303
	$(skewness(s_{dc}))^{-1}$	100,000	1,471	1,047	3,310	99,998	100,002	1,098
	$(kurtosis(k_{dc}))^{-1}$	100,000	1,506	0,605	1,912	99,999	100,001	0,368
20	1/n	99,998	0,087	0,491	1,552	99,997	99,999	0,241
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,698	0,484	1,529	99,998	100,000	0,234
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,492	0,486	1,536	99,998	100,000	0,236
	$(skewness(s_{dc}))^{-1}$	99,999	1,061	1,015	3,209	99,997	100,001	1,031
	$(kurtosis(k_{dc}))^{-1}$	99,999	1,184	0,528	1,668	99,998	100,001	0,280

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Combining samples using any of the weighting strategies under consideration resulted in lower MSE values compared to using estimates from only one sample. The MSE value decreased as the number of samples combined was increased. The

weighting strategies' performance were comparable with the exception of the inverse of skewness which resulted in notably higher MSE values relative to the others weighting strategies. This was expected because the superpopulation distribution is not skewed.

Table 4. 11: Scenario 3: Combining D samples of size n=10 000, from a population of size N=1000 000, distributed as the t-distribution with three degrees of freedom, t(3)(100,25), $\theta=99.998$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		99,999	0,624	1,546	4,890	99,996	100,002	2,392
2	1/n	99,998	-0,522	1,093	3,457	99,996	100,000	1,195
	$(Var(\hat{\theta}_a))^{-1}$	99,998	-0,095	1,087	3,437	99,996	100,000	1,181
	$(CV(\hat{\theta}_a))^{-1}$	99,998	-0,292	1,089	3,443	99,996	100,000	1,185
	$(skewness(s_{dc}))^{-1}$	100,000	1,729	1,213	3,835	99,998	100,002	1,474
	$(kurtosis(k_{dc}))^{-1}$	100,000	1,811	1,169	3,698	99,998	100,002	1,371
3	1/n	99,998	0,146	0,887	2,804	99,997	100,000	0,786
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,539	0,881	2,786	99,997	100,001	0,777
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,369	0,882	2,791	99,997	100,000	0,779
	$(skewness(s_{dc}))^{-1}$	100,001	2,324	1,086	3,434	99,998	100,003	1,184
	$(kurtosis(k_{dc}))^{-1}$	100,000	1,956	0,979	3,097	99,998	100,002	0,963
5	1/n	99,999	0,416	0,665	2,104	99,997	100,000	0,443
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,826	0,660	2,087	99,998	100,000	0,436
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,657	0,661	2,092	99,998	100,000	0,438
	$(skewness(s_{dc}))^{-1}$	100,000	1,317	0,917	2,900	99,998	100,001	0,843
	$(kurtosis(k_{dc}))^{-1}$	100,000	1,927	0,743	2,348	99,999	100,002	0,555
10	1/n	99,998	0,162	0,482	1,523	99,998	99,999	0,232
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,639	0,480	1,518	99,998	100,000	0,231
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,443	0,480	1,518	99,998	100,000	0,231
	$(skewness(s_{dc}))^{-1}$	100,000	1,390	0,785	2,481	99,998	100,001	0,617
	$(kurtosis(k_{dc}))^{-1}$	100,000	1,868	0,543	1,716	99,999	100,001	0,298
15	1/n	99,998	0,179	0,401	1,267	99,998	99,999	0,160
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,639	0,399	1,261	99,998	100,000	0,159
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,453	0,399	1,262	99,998	100,000	0,159
	$(skewness(s_{dc}))^{-1}$	100,000	1,219	0,714	2,257	99,998	100,001	0,511
	$(kurtosis(k_{dc}))^{-1}$	100,000	1,623	0,443	1,402	99,999	100,001	0,199
20	1/n	99,998	0,043	0,347	1,098	99,998	99,999	0,121
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,489	0,345	1,092	99,998	99,999	0,119
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,312	0,346	1,093	99,998	99,999	0,120
	$(skewness(s_{dc}))^{-1}$	99,999	0,919	0,673	2,129	99,998	100,001	0,454
	$(kurtosis(k_{dc}))^{-1}$	100,000	1,318	0,377	1,193	99,999	100,000	0,144

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

The MSE values resulting from all the weighting strategies were comparable except the inverse of skewness which resulted in higher MSE values.

Table 4.12 presents a comparison of MSE values across sample sizes.

Table 4. 12: MSE for the t distribution with three degrees of freedom, mean=100, N=1 000 000 across sample sizes n and number of samples combined D.

D	Method	n=1000 $\times 10^{-2}$	n=5000 $\times 10^{-3}$	n=10000 $\times 10^{-3}$
1		2,493	5,004	2,392
2	$1/n$	1,228	2,625	1,195
	$(Var(\hat{\theta}_d))^{-1}$	1,185	2,576	1,181
	$(CV(\hat{\theta}_d))^{-1}$	1,228	2,593	1,185
	$(skewness(s_{dc}))^{-1}$	1,195	3,105	1,474
	$(kurtosis(s_{dc}))^{-1}$	1,364	2,854	1,371
3	$1/n$	1,314	1,701	0,786
	$(Var(\hat{\theta}_d))^{-1}$	1,252	1,686	0,777
	$(CV(\hat{\theta}_d))^{-1}$	0,828	1,686	0,779
	$(skewness(s_{dc}))^{-1}$	0,790	2,478	1,184
	$(kurtosis(s_{dc}))^{-1}$	0,828	2,042	0,963
5	$1/n$	0,799	0,988	0,443
	$(Var(\hat{\theta}_d))^{-1}$	1,003	0,972	0,436
	$(CV(\hat{\theta}_d))^{-1}$	0,894	0,975	0,438
	$(skewness(s_{dc}))^{-1}$	0,839	1,817	0,843
	$(kurtosis(s_{dc}))^{-1}$	0,506	1,204	0,555
10	$1/n$	0,479	0,437	0,232
	$(Var(\hat{\theta}_d))^{-1}$	0,506	0,432	0,231
	$(CV(\hat{\theta}_d))^{-1}$	0,486	0,432	0,231
	$(skewness(s_{dc}))^{-1}$	0,727	1,270	0,617
	$(kurtosis(s_{dc}))^{-1}$	0,532	0,542	0,298
15	$1/n$	0,503	0,308	0,160
	$(Var(\hat{\theta}_d))^{-1}$	0,234	0,301	0,159
	$(CV(\hat{\theta}_d))^{-1}$	0,222	0,303	0,159
	$(skewness(s_{dc}))^{-1}$	0,234	1,098	0,511
	$(kurtosis(s_{dc}))^{-1}$	0,224	0,368	0,199
20	$1/n$	0,561	0,241	0,121
	$(Var(\hat{\theta}_d))^{-1}$	0,263	0,234	0,119
	$(CV(\hat{\theta}_d))^{-1}$	0,251	0,236	0,120
	$(skewness(s_{dc}))^{-1}$	0,155	1,031	0,454
	$(kurtosis(s_{dc}))^{-1}$	0,146	0,280	0,144

method=weighting strategy, D = number of samples combined.

In most cases, the same choice of weighting strategy was made for the sample sizes $n=5000$ and $n=10000$ when the same number of samples were combined.

4.2.4 Simulation results for a finite population generated using the skewed t- distribution with three degrees of freedom, under simple random sampling

The skewed t-distribution with three degrees of freedom was used to generate the finite population. The results of comparing the weighting strategies are presented in Appendix D1. A location parameter of -0.7891, a scale parameter of 0.746 and an asymmetry parameter of 3 were used to generate strongly asymmetric data with mean 100 and variance 25, 49 and 81. The inverse of the sample size was selected as the best weighting strategy for all the scenarios considered. We compared estimates calculated from one sample to estimates resulting from combining two or more samples. In all the cases, combining samples resulted in lower MSE values. Almost all the confidence intervals excluded the finite population parameter value. Results for scenarios 1, 2 and 3 are presented below.

Table 4. 13: Scenario 1: Combining D samples of size n=1000, from a population of size N=1000 000, distributed as the skewed t-distribution with three degrees of freedom, at(3)(100,25), $\theta=99.955$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,956	0,111	4,988	1,578	99,947	99,966	2,488
2	1/n	99,956	0,112	3,532	1,117	99,949	99,963	1,247
	$(Var(\hat{\theta}_a))^{-1}$	99,942	-1,314	3,433	1,086	99,935	99,949	1,196
	$(CV(\hat{\theta}_a))^{-1}$	99,949	-0,616	3,447	1,091	99,942	99,956	1,192
	$(skewness(s_{dc}))^{-1}$	99,942	-1,318	3,546	1,122	99,935	99,949	1,274
	$(kurtosis(k_{dc}))^{-1}$	99,937	-1,842	3,704	1,172	99,930	99,944	1,406
	$(K^2)^{-1}$	99,945	-1,040	3,539	1,120	99,938	99,952	1,263
3	1/n	99,955	0,008	2,935	0,929	99,950	99,961	0,862
	$(Var(\hat{\theta}_a))^{-1}$	99,938	-1,747	2,873	0,909	99,932	99,943	0,856
	$(CV(\hat{\theta}_a))^{-1}$	99,946	-0,892	2,871	0,908	99,941	99,952	0,832
	$(skewness(s_{dc}))^{-1}$	99,937	-1,790	2,968	0,939	99,932	99,943	0,913
	$(kurtosis(k_{dc}))^{-1}$	99,930	-2,496	3,168	1,003	99,924	99,937	1,066
	$(K^2)^{-1}$	99,941	-1,472	2,966	0,939	99,935	99,946	0,902
5	1/n	99,956	0,126	2,238	0,708	99,952	99,961	0,501
	$(Var(\hat{\theta}_a))^{-1}$	99,936	-1,933	2,165	0,685	99,932	99,940	0,506
	$(CV(\hat{\theta}_a))^{-1}$	99,946	-0,939	2,171	0,687	99,942	99,950	0,480
	$(skewness(s_{dc}))^{-1}$	99,935	-1,976	2,252	0,713	99,931	99,940	0,546
	$(kurtosis(k_{dc}))^{-1}$	99,927	-2,803	2,441	0,773	99,922	99,932	0,674
	$(K^2)^{-1}$	99,939	-1,626	2,248	0,711	99,935	99,943	0,532
10	1/n	99,954	-0,095	1,525	0,483	99,951	99,957	0,233
	$(Var(\hat{\theta}_a))^{-1}$	99,932	-2,365	1,495	0,473	99,929	99,935	0,279
	$(CV(\hat{\theta}_a))^{-1}$	99,943	-1,273	1,488	0,471	99,940	99,945	0,238
	$(skewness(s_{dc}))^{-1}$	99,931	-2,388	1,570	0,497	99,928	99,934	0,303
	$(kurtosis(k_{dc}))^{-1}$	99,922	-3,305	1,735	0,549	99,919	99,926	0,410
	$(K^2)^{-1}$	99,935	-2,028	1,565	0,495	99,932	99,938	0,286
15	1/n	99,956	0,029	1,256	0,397	99,953	99,958	0,158
	$(Var(\hat{\theta}_a))^{-1}$	99,932	-2,298	1,251	0,396	99,930	99,935	0,209
	$(CV(\hat{\theta}_a))^{-1}$	99,943	-1,178	1,238	0,392	99,941	99,946	0,167
	$(skewness(s_{dc}))^{-1}$	99,932	-2,323	1,322	0,418	99,929	99,935	0,229
	$(kurtosis(k_{dc}))^{-1}$	99,923	-3,265	1,454	0,460	99,920	99,925	0,318
	$(K^2)^{-1}$	99,936	-1,958	1,312	0,415	99,933	99,938	0,210
20	1/n	99,956	0,091	1,117	0,353	99,954	99,958	0,125
	$(Var(\hat{\theta}_a))^{-1}$	99,932	-2,275	1,111	0,352	99,930	99,935	0,175
	$(CV(\hat{\theta}_a))^{-1}$	99,944	-1,136	1,099	0,348	99,942	99,946	0,134
	$(skewness(s_{dc}))^{-1}$	99,932	-2,287	1,171	0,371	99,930	99,935	0,189
	$(kurtosis(k_{dc}))^{-1}$	99,923	-3,224	1,286	0,407	99,921	99,926	0,269
	$(K^2)^{-1}$	99,936	-1,916	1,162	0,368	99,934	99,938	0,172

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Weighting strategies performed comparably. However, the inverse of kurtosis resulted in higher MSE values.

Table 4. 14: Scenario 2: Combining D samples of size n=5 000, from a population of size N=1000 000, distributed as the skewed t-distribution with three degrees of freedom, at(3)(100,25), $\theta=99.955$

	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	$SE(\hat{\theta}_c)$ $\times 10^{-3}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-2}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-3}$
						LB	UB	
1		99,956	0,055	2,310	7,307	99,951	99,960	5,335
2	1/n	99,958	0,317	1,648	5,213	99,955	99,962	2,725
	$(Var(\hat{\theta}_a))^{-1}$	99,956	0,030	1,642	5,194	99,952	99,959	2,696
	$(CV(\hat{\theta}_a))^{-1}$	99,957	0,174	1,641	5,192	99,954	99,960	2,697
	$(skewness(s_{dc}))^{-1}$	99,955	-0,069	1,686	5,333	99,951	99,958	2,842
	$(kurtosis(k_{dc}))^{-1}$	99,953	-0,253	1,774	5,613	99,949	99,956	3,154
3	1/n	99,957	0,192	1,364	4,315	99,954	99,960	1,864
	$(Var(\hat{\theta}_a))^{-1}$	99,953	-0,193	1,363	4,311	99,951	99,956	1,861
	$(CV(\hat{\theta}_a))^{-1}$	99,955	-0,001	1,360	4,302	99,953	99,958	1,849
	$(skewness(s_{dc}))^{-1}$	99,952	-0,319	1,408	4,455	99,949	99,955	1,993
	$(kurtosis(k_{dc}))^{-1}$	99,949	-0,575	1,511	4,781	99,947	99,952	2,317
5	1/n	99,957	0,130	1,062	3,361	99,954	99,959	1,130
	$(Var(\hat{\theta}_a))^{-1}$	99,952	-0,338	1,048	3,315	99,950	99,954	1,109
	$(CV(\hat{\theta}_a))^{-1}$	99,954	-0,104	1,052	3,328	99,952	99,956	1,107
	$(skewness(s_{dc}))^{-1}$	99,950	-0,485	1,076	3,404	99,948	99,952	1,181
	$(kurtosis(k_{dc}))^{-1}$	99,947	-0,815	1,158	3,664	99,945	99,949	1,408
10	1/n	99,956	0,099	0,724	2,291	99,955	99,958	0,525
	$(Var(\hat{\theta}_a))^{-1}$	99,951	-0,431	0,716	2,265	99,950	99,952	0,531
	$(CV(\hat{\theta}_a))^{-1}$	99,954	-0,167	0,718	2,270	99,952	99,955	0,518
	$(skewness(s_{dc}))^{-1}$	99,949	-0,616	0,735	2,325	99,948	99,951	0,578
	$(kurtosis(k_{dc}))^{-1}$	99,945	-1,036	0,802	2,539	99,943	99,946	0,751
15	1/n	99,956	0,087	0,593	1,875	99,955	99,957	0,352
	$(Var(\hat{\theta}_a))^{-1}$	99,951	-0,468	0,588	1,862	99,949	99,952	0,368
	$(CV(\hat{\theta}_a))^{-1}$	99,953	-0,191	0,589	1,862	99,952	99,954	0,350
	$(skewness(s_{dc}))^{-1}$	99,949	-0,672	0,606	1,917	99,947	99,950	0,412
	$(kurtosis(k_{dc}))^{-1}$	99,944	-1,124	0,664	2,101	99,943	99,945	0,567
20	1/n	99,955	0,024	0,506	1,602	99,954	99,956	0,256
	$(Var(\hat{\theta}_a))^{-1}$	99,950	-0,540	0,505	1,599	99,949	99,951	0,285
	$(CV(\hat{\theta}_a))^{-1}$	99,953	-0,259	0,504	1,596	99,952	99,954	0,261
	$(skewness(s_{dc}))^{-1}$	99,948	-0,750	0,522	1,650	99,947	99,949	0,328
	$(kurtosis(k_{dc}))^{-1}$	99,943	-1,216	0,571	1,808	99,942	99,944	0,474

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, $SE(\hat{\theta}_c)$ =standard error of the combined estimate, $CV(\hat{\theta}_c)$ (%) = coefficient of variation, $CI_{95\%}(\hat{\theta}_c)$ = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

The weighting strategies resulted in comparable MSE values barring the inverse of kurtosis and to a lesser extent the inverse of skewness.

Table 4. 15: Scenario 3: Combining D samples of size n=10 000, from a population of size N=1000 000, distributed as the skewed t-distribution with three degrees of freedom, at(3)(100,25), $\theta=99.955$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		99,955	0,068	1,573	4,976	99,952	99,958	2,474
2	1/n	99,955	-0,312	1,095	3,464	99,953	99,957	1,199
	$(Var(\hat{\theta}_a))^{-1}$	99,953	-1,893	1,091	3,451	99,951	99,955	1,193
	$(CV(\hat{\theta}_a))^{-1}$	99,954	-1,094	1,092	3,454	99,952	99,956	1,193
	$(skewness(s_{dc}))^{-1}$	99,952	-2,952	1,114	3,526	99,950	99,954	1,251
	$(kurtosis(k_{dc}))^{-1}$	99,951	-4,528	1,184	3,745	99,948	99,953	1,422
3	1/n	99,955	0,101	0,905	2,862	99,954	99,957	0,818
	$(Var(\hat{\theta}_a))^{-1}$	99,953	-1,987	0,903	2,857	99,951	99,955	0,819
	$(CV(\hat{\theta}_a))^{-1}$	99,954	-0,934	0,903	2,855	99,953	99,956	0,815
	$(skewness(s_{dc}))^{-1}$	99,952	-3,398	0,937	2,965	99,950	99,954	0,890
	$(kurtosis(k_{dc}))^{-1}$	99,950	-5,562	1,026	3,247	99,948	99,952	1,084
5	1/n	99,956	0,580	0,716	2,264	99,954	99,957	0,512
	$(Var(\hat{\theta}_a))^{-1}$	99,953	-1,854	0,717	2,268	99,952	99,955	0,517
	$(CV(\hat{\theta}_a))^{-1}$	99,955	-0,628	0,715	2,263	99,953	99,956	0,512
	$(skewness(s_{dc}))^{-1}$	99,952	-3,382	0,745	2,358	99,950	99,953	0,567
	$(kurtosis(k_{dc}))^{-1}$	99,949	-6,033	0,821	2,597	99,948	99,951	0,710
10	1/n	99,956	0,339	0,510	1,614	99,955	99,957	0,260
	$(Var(\hat{\theta}_a))^{-1}$	99,953	-2,465	0,512	1,620	99,952	99,954	0,268
	$(CV(\hat{\theta}_a))^{-1}$	99,954	-1,057	0,510	1,615	99,953	99,955	0,262
	$(skewness(s_{dc}))^{-1}$	99,951	-4,304	0,535	1,693	99,950	99,952	0,305
	$(kurtosis(k_{dc}))^{-1}$	99,948	-7,487	0,593	1,876	99,947	99,949	0,408
15	1/n	99,956	0,597	0,411	1,302	99,955	99,957	0,170
	$(Var(\hat{\theta}_a))^{-1}$	99,953	-2,308	0,411	1,299	99,952	99,954	0,174
	$(CV(\hat{\theta}_a))^{-1}$	99,954	-0,850	0,410	1,298	99,954	99,955	0,169
	$(skewness(s_{dc}))^{-1}$	99,951	-4,147	0,426	1,348	99,950	99,952	0,199
	$(kurtosis(k_{dc}))^{-1}$	99,948	-7,439	0,469	1,485	99,947	99,949	0,276
20	1/n	99,956	0,648	0,364	1,150	99,955	99,957	0,133
	$(Var(\hat{\theta}_a))^{-1}$	99,953	-2,256	0,362	1,146	99,952	99,954	0,136
	$(CV(\hat{\theta}_a))^{-1}$	99,954	-0,798	0,362	1,146	99,954	99,955	0,132
	$(skewness(s_{dc}))^{-1}$	99,951	-4,046	0,375	1,187	99,950	99,952	0,157
	$(kurtosis(k_{dc}))^{-1}$	99,948	-7,326	0,413	1,307	99,947	99,949	0,224

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

The inverse of kurtosis and to a smaller degree, the inverse of skewness, resulted in higher MSE values relative to all the other weighting strategies. As in the other superpopulation distributions, MSE values were also compared across sample sizes and presented in table 4.16.

Table 4. 16: MSE for the skewed t distribution with three degrees of freedom, mean=100, N=1 000 000 across sample sizes n and number of samples combined D.

D	Method	n=1000 $\times 10^{-2}$	n=5000 $\times 10^{-3}$	n=10000 $\times 10^{-3}$
1		2,488	5,335	2,474
2	$1/n$	1,247	2,725	1,199
	$(Var(\hat{\theta}_d))^{-1}$	1,196	2,696	1,193
	$(CV(\hat{\theta}_d))^{-1}$	1,192	2,697	1,193
	$(skewness(s_{dc}))^{-1}$	1,274	2,842	1,251
	$(kurtosis(s_{dc}))^{-1}$	1,406	3,154	1,422
3	$1/n$	1,263	1,864	0,818
	$(Var(\hat{\theta}_d))^{-1}$	0,862	1,861	0,819
	$(CV(\hat{\theta}_d))^{-1}$	0,856	1,849	0,815
	$(skewness(s_{dc}))^{-1}$	0,832	1,993	0,890
	$(kurtosis(s_{dc}))^{-1}$	0,913	2,317	1,084
5	$1/n$	1,066	1,130	0,512
	$(Var(\hat{\theta}_d))^{-1}$	0,902	1,109	0,517
	$(CV(\hat{\theta}_d))^{-1}$	0,501	1,107	0,512
	$(skewness(s_{dc}))^{-1}$	0,506	1,181	0,567
	$(kurtosis(s_{dc}))^{-1}$	0,480	1,408	0,710
10	$1/n$	0,546	0,525	0,260
	$(Var(\hat{\theta}_d))^{-1}$	0,674	0,531	0,268
	$(CV(\hat{\theta}_d))^{-1}$	0,532	0,518	0,262
	$(skewness(s_{dc}))^{-1}$	0,233	0,578	0,305
	$(kurtosis(s_{dc}))^{-1}$	0,279	0,751	0,408
15	$1/n$	0,238	0,352	0,170
	$(Var(\hat{\theta}_d))^{-1}$	0,303	0,368	0,174
	$(CV(\hat{\theta}_d))^{-1}$	0,410	0,350	0,169
	$(skewness(s_{dc}))^{-1}$	0,286	0,412	0,199
	$(kurtosis(s_{dc}))^{-1}$	0,158	0,567	0,276
20	$1/n$	0,209	0,256	0,133
	$(Var(\hat{\theta}_d))^{-1}$	0,167	0,285	0,136
	$(CV(\hat{\theta}_d))^{-1}$	0,229	0,261	0,132
	$(skewness(s_{dc}))^{-1}$	0,318	0,328	0,157
	$(kurtosis(s_{dc}))^{-1}$	0,210	0,474	0,224

method=weighting strategy, D = number of samples combined.

There was no one weighting strategy which consistently stood out as better than the others. The weighting strategies performed comparably well.

4.2.7 Summary of simulation results under simple random sampling

Increasing the sample size appeared to improve the MSE values for all the superpopulations. In general, increasing the number of samples combined led to a decrease in MSE values. Under simple random sampling, the variance of the superpopulation did not influence the choice of the weighting strategy. However, increasing the variance resulted in an increase in the MSE values. The skewness of the superpopulation appeared to affect the choice of the weighting strategy. In general, all the weighting strategies except the inverse of skewness and the inverse of the D'Agostino-Pearson statistic appeared to work well with the symmetric superpopulations. This is expected as skewness is not expected to play an important role when the data are symmetrical. The size of the finite population also appeared to influence the choice of the weighting strategy.

Stratified random sampling and cluster random sampling were studied in sections 4.3 and 4.4, respectively.

4.3 Simulation results under stratified sampling

Comparison of the weighting strategies when combining survey data collected through stratified random sampling was carried out as described in section 4.1. A stratified finite population of size $N=1000\ 000$ was simulated from a super-population with a normal distribution, mean 100 and variance 25. The population was generated with five strata of equal size. The distributions within the strata were chosen in such a way that the combined population would have a normal distribution with mean 100 and variance 25. Population means and variances for the strata were as follows; ((130, 36), (115, 25), (100, 25), (85, 25), (70, 36)).

Sampling fractions were chosen to yield the required sample sizes of 1 000, 5 000 and 10 000 with higher sampling fractions in the strata with greater variation. The sampling fractions used were (0.00144, 0.00072, 0.00072, 0.00072, 0.00144), (0.00716, 0.00358, 0.00358, 0.00358, 0.00716) and (0.01432, 0.00714, 0.00714, 0.00714, 0.01432) for the respective required sample sizes. Within each stratum, the sample was selected by simple random sampling without replacement.

As with simple random sampling, the weighting strategies were compared when estimates from 2,3,5,10,15 or 20 samples were combined.

Misspecification effects were calculated for the estimated means from each sample selected. The average misspecification effects observed for each superpopulation distribution and sample size are presented in table 4.3.1 below.

Table 4. 17 : Misspecification effects for all the distributions and sample sizes considered under stratified sampling

Distribution	Sample size	Misspecification effect
Normal	1000	0,402
	5000	0,402
	10000	0,406
Skewed normal	1000	0,409
	5000	0,405
	10000	0,406
t, 3 d.f	1000	0,408
	5000	0,424
	10000	0,399
skewed t, 3 d.f	1000	0,391
	5000	0,412
	10000	0,422

The misspecification effects were approximately the same across sample sizes and data distributions. They are all less than one. Ignoring stratification in the analysis would lead to overestimation of the variance, $Var(\hat{\theta})$. In section 1.2, we discussed design effects and misspecification effects. In this thesis, we consider two weighting strategies which use the misspecification effect. That is, $\frac{1}{m_{eff}}$ and $\frac{m_{eff}}{var(y)}$.

Similarly, to simple random sampling, weighting by the inverse of the sample size resulted in identical results to using a weight of one. This could be because all the strata used in this study are of equal size. When strata are of different sizes, as is often the case in practical situations, these weighting strategies may not yield identical results. Only results for weighting by the inverse of the sample size will be presented in subsections 4.3.1 to 4.3.6.

4.3.1 Simulation results for a finite population generated using the normal distribution, under stratified sampling

The normal distribution was used to generate the finite population from which different sample surveys were then selected. Further results of comparing the weighting strategies for combining the sample surveys are presented in Appendix A2. The tables below only present the results obtained when design consistent weighting strategies were used. Only three scenarios were considered for this simulation. Simulations were done for a superpopulation distributed as $N(100,25)$ with sample sizes 1000, 5000 and 10000.

Table 4. 18 : Scenario 1: Combining D samples of size n=1000, from a population of size N=1000 000, distributed as the normal distribution N(100,25), $\theta=99.999$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,996	-3,524	3,471	1,098	99,989	100,002	1,206
2	1/n	99,997	-2,406	2,432	0,769	99,992	100,001	0,592
	$(Var(\hat{\theta}_a))^{-1}$	99,997	-1,839	2,457	0,777	99,992	100,002	0,604
	$(CV(\hat{\theta}_a))^{-1}$	99,997	-1,824	2,454	0,776	99,992	100,002	0,603
	$(skewness(s_{dc}))^{-1}$	99,998	-1,424	2,761	0,873	99,992	100,003	0,762
	$(kurtosis(k_{dc}))^{-1}$	99,997	-1,899	2,456	0,777	99,992	100,002	0,604
	$meff/Var(y)$	99,997	-1,984	2,454	0,776	99,992	100,002	0,603
	$(meff)^{-1}$	99,997	-1,792	2,456	0,777	99,992	100,002	0,604
3	1/n	100,000	0,982	1,912	0,605	99,996	100,004	0,366
	$(Var(\hat{\theta}_a))^{-1}$	100,001	1,559	1,924	0,608	99,997	100,004	0,370
	$(CV(\hat{\theta}_a))^{-1}$	100,001	1,606	1,921	0,608	99,997	100,004	0,369
	$(skewness(s_{dc}))^{-1}$	100,001	1,611	2,422	0,766	99,996	100,005	0,587
	$(kurtosis(k_{dc}))^{-1}$	100,001	1,489	1,919	0,607	99,997	100,004	0,368
	$meff/Var(y)$	100,000	1,374	1,926	0,609	99,997	100,004	0,371
	$(meff)^{-1}$	100,001	1,608	1,925	0,609	99,997	100,004	0,371
5	1/n	100,000	1,088	1,553	0,491	99,997	100,003	0,241
	$(Var(\hat{\theta}_a))^{-1}$	100,000	1,213	1,555	0,492	99,997	100,003	0,242
	$(CV(\hat{\theta}_a))^{-1}$	100,000	1,245	1,554	0,491	99,997	100,003	0,241
	$(skewness(s_{dc}))^{-1}$	100,000	0,516	2,177	0,688	99,995	100,004	0,474
	$(kurtosis(k_{dc}))^{-1}$	100,000	1,102	1,556	0,492	99,997	100,003	0,242
	$meff/Var(y)$	100,000	1,027	1,555	0,492	99,997	100,003	0,242
	$(meff)^{-1}$	100,000	1,075	1,561	0,494	99,997	100,003	0,244
10	1/n	100,001	2,338	1,110	0,351	99,999	100,004	0,124
	$(Var(\hat{\theta}_a))^{-1}$	100,001	2,382	1,112	0,351	99,999	100,004	0,124
	$(CV(\hat{\theta}_a))^{-1}$	100,002	2,467	1,111	0,351	99,999	100,004	0,124
	$(skewness(s_{dc}))^{-1}$	100,001	2,101	1,853	0,586	99,998	100,005	0,344
	$(kurtosis(k_{dc}))^{-1}$	100,001	2,380	1,111	0,351	99,999	100,004	0,124
	$meff/Var(y)$	100,001	2,225	1,111	0,351	99,999	100,003	0,124
	$(meff)^{-1}$	100,001	2,372	1,114	0,352	99,999	100,004	0,125
15	1/n	100,000	0,870	0,919	0,291	99,998	100,002	0,085
	$(Var(\hat{\theta}_a))^{-1}$	100,000	0,891	0,918	0,290	99,998	100,002	0,084
	$(CV(\hat{\theta}_a))^{-1}$	100,000	0,993	0,918	0,290	99,998	100,002	0,084
	$(skewness(s_{dc}))^{-1}$	99,999	0,134	1,764	0,558	99,996	100,003	0,311
	$(kurtosis(k_{dc}))^{-1}$	100,000	0,898	0,919	0,291	99,998	100,002	0,085
	$meff/Var(y)$	100,000	0,896	0,917	0,290	99,998	100,002	0,084
	$(meff)^{-1}$	100,000	0,950	0,920	0,291	99,998	100,002	0,085

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.18 continued: Scenario 1: Combining D samples of size n=1000, from a population of size N=1000 000, distributed as the normal distribution N(100,25), $\theta=99.999$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	1/n	99,999	0,347	0,791	0,250	99,998	100,001	0,063
	$(Var(\hat{\theta}_d))^{-1}$	99,999	0,347	0,793	0,251	99,998	100,001	0,063
	$(CV(\hat{\theta}_d))^{-1}$	100,000	0,461	0,792	0,250	99,998	100,001	0,063
	$(skewness(s_{dc}))^{-1}$	99,998	-0,625	1,673	0,529	99,995	100,002	0,280
	$(kurtosis(k_{dc}))^{-1}$	99,999	0,366	0,791	0,250	99,998	100,001	0,063
	$meff/Var(y)$	99,999	0,360	0,790	0,250	99,998	100,001	0,062
	$(meff)^{-1}$	99,999	0,325	0,794	0,251	99,998	100,001	0,063

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Using only one sample resulted in a much larger MSE value compared to combining samples using any one of the weighting strategies considered. The weighting strategies performed similarly with very close MSE values. Weighting the samples by the inverse of skewness resulted in higher MSE values in relation to the all the other weighting strategies. This was expected since the superpopulation distribution is not skewed.

Table 4. 19: Scenario 2: Combining D samples of size n=5000, from a population of size N=1000 000, distributed as the normal distribution N(100,25), $\theta=99.999$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		100,001	2,212	1,591	5,032	99,998	100,004	2,537
2	1/n	100,001	1,666	1,094	3,461	99,999	100,003	1,200
	$(Var(\hat{\theta}_a))^{-1}$	100,001	1,848	1,100	3,478	99,999	100,003	1,213
	$(CV(\hat{\theta}_a))^{-1}$	100,001	1,860	1,100	3,477	99,999	100,003	1,212
	$(skewness(s_{dc}))^{-1}$	100,001	1,798	1,199	3,792	99,999	100,003	1,441
	$(kurtosis(k_{dc}))^{-1}$	100,001	1,830	1,099	3,475	99,999	100,003	1,211
	$meff/Var(y)$	100,001	1,835	1,100	3,477	99,999	100,003	1,213
	$(meff)^{-1}$	100,001	1,861	1,100	3,477	99,999	100,003	1,213
3	1/n	100,000	1,208	0,911	2,882	99,999	100,002	0,832
	$(Var(\hat{\theta}_a))^{-1}$	100,000	1,341	0,916	2,896	99,999	100,002	0,841
	$(CV(\hat{\theta}_a))^{-1}$	100,000	1,354	0,916	2,895	99,999	100,002	0,840
	$(skewness(s_{dc}))^{-1}$	100,001	1,736	1,098	3,471	99,999	100,003	1,208
	$(kurtosis(k_{dc}))^{-1}$	100,000	1,324	0,915	2,893	99,999	100,002	0,839
	$meff/Var(y)$	100,000	1,341	0,915	2,894	99,999	100,002	0,839
	$(meff)^{-1}$	100,000	1,345	0,916	2,898	99,999	100,002	0,841
5	1/n	100,000	0,815	0,703	2,223	99,999	100,001	0,495
	$(Var(\hat{\theta}_a))^{-1}$	100,000	0,811	0,703	2,223	99,999	100,001	0,495
	$(CV(\hat{\theta}_a))^{-1}$	100,000	0,832	0,703	2,223	99,999	100,001	0,495
	$(skewness(s_{dc}))^{-1}$	100,001	1,637	0,978	3,092	99,999	100,003	0,959
	$(kurtosis(k_{dc}))^{-1}$	100,000	0,797	0,703	2,222	99,998	100,001	0,495
	$meff/Var(y)$	100,000	0,831	0,703	2,222	99,999	100,001	0,494
	$(meff)^{-1}$	100,000	0,831	0,704	2,225	99,999	100,001	0,496
10	1/n	99,999	-0,195	0,496	1,567	99,998	100,000	0,246
	$(Var(\hat{\theta}_a))^{-1}$	99,999	-0,204	0,496	1,567	99,998	100,000	0,246
	$(CV(\hat{\theta}_a))^{-1}$	99,999	-0,178	0,496	1,567	99,998	100,000	0,246
	$(skewness(s_{dc}))^{-1}$	100,000	0,850	0,867	2,741	99,998	100,002	0,752
	$(kurtosis(k_{dc}))^{-1}$	99,999	-0,193	0,496	1,568	99,998	100,000	0,246
	$meff/Var(y)$	99,999	-0,220	0,496	1,569	99,998	100,000	0,246
	$(meff)^{-1}$	99,999	-0,190	0,496	1,567	99,998	100,000	0,246
15	1/n	99,999	-0,297	0,400	1,266	99,998	100,000	0,160
	$(Var(\hat{\theta}_a))^{-1}$	99,999	-0,308	0,401	1,267	99,998	100,000	0,161
	$(CV(\hat{\theta}_a))^{-1}$	99,999	-0,280	0,400	1,266	99,998	100,000	0,160
	$(skewness(s_{dc}))^{-1}$	100,000	0,918	0,816	2,580	99,998	100,002	0,667
	$(kurtosis(k_{dc}))^{-1}$	99,999	-0,297	0,400	1,266	99,998	100,000	0,160
	$meff/Var(y)$	99,999	-0,293	0,400	1,266	99,998	100,000	0,160
	$(meff)^{-1}$	99,999	-0,310	0,401	1,267	99,998	100,000	0,161

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.19 continued: Scenario 2: Combining D samples of size n=5000, from a population of size N=1000 000, distributed as the normal distribution N(100,25), $\theta=99.999$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	1/n	99,999	-0,136	0,352	1,113	99,998	100,000	0,124
	$(Var(\hat{\theta}_d))^{-1}$	99,999	-0,149	0,352	1,114	99,998	100,000	0,124
	$(CV(\hat{\theta}_d))^{-1}$	99,999	-0,120	0,352	1,113	99,998	100,000	0,124
	$(skewness(s_{dc}))^{-1}$	100,000	0,837	0,789	2,494	99,998	100,001	0,623
	$(kurtosis(k_{dc}))^{-1}$	99,999	-0,138	0,352	1,113	99,998	100,000	0,124
	$meff/Var(y)$	99,999	-0,192	0,353	1,117	99,998	100,000	0,125
	$(meff)^{-1}$	99,999	-0,169	0,352	1,112	99,998	100,000	0,124

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Combining samples resulted in lower MSE values compared to using one sample. The weighting strategies resulted in MSE values that were close in magnitude. As expected, using the inverse of skewness as a weighting strategy resulted in notably higher MSE values compared to the other weighting strategies studied.

Table 4. 20: Scenario 3: Combining D samples of size n=10000, from a population of size N=1000 000, distributed as the normal distribution N(100,25), $\theta=99.999$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		99,999	-0,251	1,086	3,434	99,997	100,001	1,179
2	1/n	100,000	1,047	0,778	2,459	99,999	100,002	0,606
	$(Var(\hat{\theta}_a))^{-1}$	100,000	1,057	0,782	2,474	99,999	100,002	0,613
	$(CV(\hat{\theta}_a))^{-1}$	100,000	1,074	0,782	2,473	99,999	100,002	0,613
	$(skewness(s_{dc}))^{-1}$	100,001	1,560	0,881	2,786	99,999	100,002	0,779
	$(kurtosis(k_{dc}))^{-1}$	100,000	1,082	0,782	2,472	99,999	100,002	0,612
	$meff/Var(y)$	100,000	1,069	0,782	2,473	99,999	100,002	0,613
	$(meff)^{-1}$	100,000	1,068	0,782	2,474	99,999	100,002	0,613
3	1/n	99,999	0,044	0,646	2,042	99,998	100,000	0,417
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,204	0,648	2,049	99,998	100,001	0,420
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,217	0,648	2,048	99,998	100,001	0,419
	$(skewness(s_{dc}))^{-1}$	100,000	0,689	0,778	2,459	99,998	100,001	0,605
	$(kurtosis(k_{dc}))^{-1}$	99,999	0,211	0,647	2,047	99,998	100,001	0,419
	$meff/Var(y)$	99,999	0,209	0,647	2,047	99,998	100,001	0,419
	$(meff)^{-1}$	99,999	0,208	0,648	2,049	99,998	100,001	0,420
5	1/n	100,000	0,637	0,489	1,546	99,999	100,001	0,239
	$(Var(\hat{\theta}_a))^{-1}$	100,000	0,635	0,489	1,548	99,999	100,001	0,240
	$(CV(\hat{\theta}_a))^{-1}$	100,000	0,646	0,489	1,547	99,999	100,001	0,240
	$(skewness(s_{dc}))^{-1}$	100,000	1,341	0,673	2,129	99,999	100,002	0,455
	$(kurtosis(k_{dc}))^{-1}$	100,000	0,642	0,489	1,545	99,999	100,001	0,239
	$meff/Var(y)$	100,000	0,622	0,489	1,548	99,999	100,001	0,240
	$(meff)^{-1}$	100,000	0,627	0,490	1,549	99,999	100,001	0,240
10	1/n	100,000	0,499	0,348	1,100	99,999	100,000	0,121
	$(Var(\hat{\theta}_a))^{-1}$	100,000	0,496	0,348	1,102	99,999	100,000	0,122
	$(CV(\hat{\theta}_a))^{-1}$	100,000	0,508	0,348	1,101	99,999	100,000	0,122
	$(skewness(s_{dc}))^{-1}$	100,000	0,959	0,595	1,882	99,999	100,001	0,355
	$(kurtosis(k_{dc}))^{-1}$	100,000	0,507	0,348	1,101	99,999	100,000	0,121
	$meff/Var(y)$	100,000	0,474	0,350	1,106	99,999	100,000	0,122
	$(meff)^{-1}$	100,000	0,497	0,349	1,102	99,999	100,000	0,122
15	1/n	99,999	0,405	0,282	0,893	99,999	100,000	0,080
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,405	0,283	0,894	99,999	100,000	0,080
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,416	0,283	0,894	99,999	100,000	0,080
	$(skewness(s_{dc}))^{-1}$	100,000	0,579	0,550	1,740	99,999	100,001	0,303
	$(kurtosis(k_{dc}))^{-1}$	99,999	0,411	0,283	0,894	99,999	100,000	0,080
	$meff/Var(y)$	99,999	0,389	0,283	0,894	99,999	100,000	0,080
	$(meff)^{-1}$	99,999	0,417	0,283	0,894	99,999	100,000	0,080

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.20 continued: Scenario 3: Combining D samples of size n=10000, from a population of size N=1000 000, distributed as the normal distribution N(100,25), $\theta=99.999$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	1/n	100,000	0,512	0,243	0,769	99,999	100,000	0,059
	$(Var(\hat{\theta}_d))^{-1}$	100,000	0,513	0,243	0,769	99,999	100,000	0,059
	$(CV(\hat{\theta}_d))^{-1}$	100,000	0,524	0,243	0,769	99,999	100,000	0,059
	$(skewness(s_{dc}))^{-1}$	100,000	0,517	0,525	1,659	99,999	100,001	0,275
	$(kurtosis(k_{dc}))^{-1}$	100,000	0,518	0,243	0,769	99,999	100,000	0,059
	$meff/Var(y)$	100,000	0,464	0,246	0,779	99,999	100,000	0,061
	$(meff)^{-1}$	100,000	0,475	0,245	0,776	99,999	100,000	0,060

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

The weighting strategies performed similarly with the exception of the inverse of skewness which mostly resulted in higher MSE values as the number of samples combined was varied.

Summary

In general, combining samples improved the MSE values. The MSE values tended to decrease as more samples were combined and when the sample size was increased. The inverse of skewness did work not well with the normal distribution superpopulation resulting in higher MSE values. The superpopulation distribution was not skewed hence this was expected. All the other weighting strategies performed similarly and could be used for weighting.

4.3.2 Simulation results for a finite population generated using the skewed normal distribution, under stratified sampling

In this simulation, the finite population was generated by the skewed normal distribution. A location parameter of -1.15, scale parameter of 1.541 and an asymmetry parameter 3 were used to generate strongly asymmetric data with mean 100 and variance 25.

The results of comparing the design conscious weighting strategies are presented in tables 4.21 to 4.22 below and in Appendix B1, where the naïve weights are also included for completeness.

Table 4. 21: Scenario 1: Comparing weighting strategies for a finite population generated using the skewed normal distribution, SN(100,25), N=1000 000 and n=1000 $\theta=100.056$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		100,052	-4,893	3,479	1,099	100,045	100,058	1,213
2	1/n	100,052	-3,949	2,503	0,791	100,048	100,057	0,628
	$(Var(\hat{\theta}_a))^{-1}$	100,051	-5,485	2,521	0,797	100,046	100,056	0,638
	$(CV(\hat{\theta}_a))^{-1}$	100,052	-4,799	2,518	0,796	100,047	100,057	0,637
	$(skewness(s_{dc}))^{-1}$	100,051	-5,335	2,575	0,814	100,046	100,056	0,666
	$(kurtosis(k_{dc}))^{-1}$	100,052	-4,150	2,518	0,796	100,047	100,057	0,636
	$mef/Var(y)$	100,052	-4,735	2,517	0,795	100,047	100,057	0,636
	$(mef)^{-1}$	100,051	-4,977	2,522	0,797	100,047	100,056	0,638
3	1/n	100,053	-3,311	2,037	0,644	100,049	100,057	0,416
	$(Var(\hat{\theta}_a))^{-1}$	100,051	-5,567	2,051	0,648	100,047	100,055	0,424
	$(CV(\hat{\theta}_a))^{-1}$	100,052	-4,682	2,051	0,648	100,048	100,056	0,423
	$(skewness(s_{dc}))^{-1}$	100,052	-4,330	2,113	0,668	100,048	100,056	0,448
	$(kurtosis(k_{dc}))^{-1}$	100,053	-3,880	2,051	0,648	100,049	100,057	0,422
	$mef/Var(y)$	100,052	-4,616	2,050	0,648	100,048	100,056	0,423
	$(mef)^{-1}$	100,052	-4,918	2,052	0,649	100,047	100,056	0,424
5	1/n	100,054	-1,992	1,541	0,487	100,051	100,057	0,238
	$(Var(\hat{\theta}_a))^{-1}$	100,053	-3,902	1,537	0,486	100,050	100,056	0,238
	$(CV(\hat{\theta}_a))^{-1}$	100,054	-2,851	1,539	0,486	100,051	100,057	0,238
	$(skewness(s_{dc}))^{-1}$	100,054	-2,399	1,630	0,515	100,051	100,057	0,266
	$(kurtosis(k_{dc}))^{-1}$	100,054	-1,944	1,543	0,488	100,051	100,058	0,238
	$mef/Var(y)$	100,054	-2,785	1,540	0,487	100,051	100,057	0,238
	$(mef)^{-1}$	100,053	-3,181	1,540	0,487	100,050	100,056	0,238
10	1/n	100,055	-1,312	1,079	0,341	100,053	100,057	0,117
	$(Var(\hat{\theta}_a))^{-1}$	100,053	-3,474	1,080	0,341	100,051	100,055	0,118
	$(CV(\hat{\theta}_a))^{-1}$	100,054	-2,285	1,079	0,341	100,052	100,056	0,117
	$(skewness(s_{dc}))^{-1}$	100,055	-1,198	1,186	0,375	100,053	100,058	0,141
	$(kurtosis(k_{dc}))^{-1}$	100,055	-1,307	1,080	0,341	100,053	100,057	0,117
	$mef/Var(y)$	100,054	-2,137	1,081	0,342	100,052	100,056	0,117
	$(mef)^{-1}$	100,054	-2,660	1,079	0,341	100,052	100,056	0,117
15	1/n	100,055	-1,722	0,887	0,280	100,053	100,056	0,079
	$(Var(\hat{\theta}_a))^{-1}$	100,052	-3,972	0,886	0,280	100,051	100,054	0,080
	$(CV(\hat{\theta}_a))^{-1}$	100,054	-2,735	0,886	0,280	100,052	100,055	0,079
	$(skewness(s_{dc}))^{-1}$	100,055	-1,384	0,998	0,315	100,053	100,057	0,100
	$(kurtosis(k_{dc}))^{-1}$	100,055	-1,781	0,889	0,281	100,053	100,056	0,079
	$mef/Var(y)$	100,054	-2,698	0,884	0,280	100,052	100,055	0,079
	$(mef)^{-1}$	100,053	-3,228	0,885	0,280	100,051	100,055	0,079

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.21 continued: Scenario 1: Comparing weighting strategies for a finite population generated using the skewed normal distribution, SN(100,25), N=1000 000 and n=1000 $\theta=100.056$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	$1/n$	100,055	-1,731	0,766	0,242	100,053	100,056	0,059
	$(Var(\hat{\theta}_d))^{-1}$	100,052	-4,046	0,765	0,242	100,051	100,054	0,060
	$(CV(\hat{\theta}_d))^{-1}$	100,054	-2,774	0,766	0,242	100,052	100,055	0,059
	$(skewness(s_{dc}))^{-1}$	100,056	-0,683	0,921	0,291	100,054	100,058	0,085
	$(kurtosis(k_{dc}))^{-1}$	100,055	-1,751	0,768	0,243	100,053	100,056	0,059
	$mef\bar{f}/Var(y)$	100,054	-2,676	0,771	0,244	100,052	100,055	0,060
	$(mef\bar{f})^{-1}$	100,053	-3,276	0,782	0,247	100,052	100,055	0,062

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

The weighting strategies mostly performed similarly, with all the MSE values close in magnitude for the same number of samples combined. Combining samples resulted in lower MSE values compared to using estimates from one sample, regardless of the weighting strategy used.

Table 4. 22: Scenario 2: Comparing weighting strategies for a finite population generated using the skewed normal distribution, SN(100,25), N=1000 000 and n=5000 $\theta=100.056$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-4}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		100,057	9,543	1,571	4,966	100,054	100,060	2,470
2	1/n	100,056	-2,360	1,088	3,438	100,054	100,058	1,183
	$(Var(\hat{\theta}_a))^{-1}$	100,057	1,791	1,090	3,446	100,054	100,059	1,189
	$(CV(\hat{\theta}_a))^{-1}$	100,057	3,083	1,090	3,445	100,055	100,059	1,188
	$(skewness(s_{dc}))^{-1}$	100,057	3,000	1,099	3,472	100,055	100,059	1,207
	$(kurtosis(k_{dc}))^{-1}$	100,057	4,092	1,091	3,448	100,055	100,059	1,190
	$meff/Var(y)$	100,057	3,159	1,090	3,445	100,055	100,059	1,188
	$(meff)^{-1}$	100,057	2,762	1,090	3,446	100,055	100,059	1,189
3	1/n	100,057	1,432	0,884	2,793	100,055	100,058	0,781
	$(Var(\hat{\theta}_a))^{-1}$	100,057	3,312	0,885	2,797	100,055	100,058	0,783
	$(CV(\hat{\theta}_a))^{-1}$	100,057	5,188	0,885	2,796	100,055	100,059	0,783
	$(skewness(s_{dc}))^{-1}$	100,057	7,177	0,894	2,824	100,055	100,059	0,799
	$(kurtosis(k_{dc}))^{-1}$	100,057	6,669	0,886	2,799	100,055	100,059	0,785
	$meff/Var(y)$	100,057	5,592	0,886	2,800	100,055	100,059	0,785
	$(meff)^{-1}$	100,057	4,624	0,885	2,797	100,055	100,059	0,784
5	1/n	100,057	1,621	0,698	2,205	100,055	100,058	0,487
	$(Var(\hat{\theta}_a))^{-1}$	100,056	-2,415	0,698	2,206	100,055	100,058	0,487
	$(CV(\hat{\theta}_a))^{-1}$	100,056	-0,194	0,698	2,205	100,055	100,058	0,487
	$(skewness(s_{dc}))^{-1}$	100,057	2,751	0,704	2,226	100,055	100,058	0,496
	$(kurtosis(k_{dc}))^{-1}$	100,057	1,735	0,698	2,206	100,055	100,058	0,487
	$meff/Var(y)$	100,056	-0,124	0,698	2,206	100,055	100,058	0,487
	$(meff)^{-1}$	100,056	-0,937	0,698	2,206	100,055	100,058	0,487
10	1/n	100,056	0,242	0,506	1,598	100,055	100,057	0,256
	$(Var(\hat{\theta}_a))^{-1}$	100,056	-4,181	0,506	1,600	100,055	100,057	0,256
	$(CV(\hat{\theta}_a))^{-1}$	100,056	-1,748	0,506	1,599	100,055	100,057	0,256
	$(skewness(s_{dc}))^{-1}$	100,057	0,745	0,512	1,617	100,055	100,058	0,262
	$(kurtosis(k_{dc}))^{-1}$	100,056	0,349	0,506	1,599	100,055	100,057	0,256
	$meff/Var(y)$	100,056	-1,084	0,508	1,605	100,055	100,057	0,258
	$(meff)^{-1}$	100,056	-2,316	0,507	1,601	100,055	100,057	0,257
15	1/n	100,057	0,928	0,414	1,308	100,056	100,057	0,171
	$(Var(\hat{\theta}_a))^{-1}$	100,056	-3,671	0,414	1,309	100,055	100,057	0,172
	$(CV(\hat{\theta}_a))^{-1}$	100,056	-1,145	0,414	1,308	100,056	100,057	0,171
	$(skewness(s_{dc}))^{-1}$	100,057	1,907	0,416	1,316	100,056	100,057	0,173
	$(kurtosis(k_{dc}))^{-1}$	100,057	1,046	0,413	1,306	100,056	100,057	0,171
	$meff/Var(y)$	100,056	-0,892	0,415	1,311	100,056	100,057	0,172
	$(meff)^{-1}$	100,056	-2,416	0,417	1,318	100,055	100,057	0,174

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.22 continued: Scenario 2: Comparing weighting strategies for a finite population generated using the skewed normal distribution, SN(100,25), N=1000 000 and n=5000 $\theta=100.056$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	$1/n$	100,057	3,099	0,354	1,118	100,056	100,057	0,125
	$(Var(\hat{\theta}_d))^{-1}$	100,056	-1,580	0,354	1,119	100,056	100,057	0,125
	$(CV(\hat{\theta}_d))^{-1}$	100,057	0,989	0,354	1,119	100,056	100,057	0,125
	$(skewness(s_{dc}))^{-1}$	100,057	3,510	0,355	1,122	100,056	100,057	0,126
	$(kurtosis(k_{dc}))^{-1}$	100,057	3,165	0,354	1,118	100,056	100,057	0,125
	$meff/Var(y)$	100,057	1,319	0,354	1,118	100,056	100,057	0,125
	$(meff)^{-1}$	100,056	0,435	0,355	1,121	100,056	100,057	0,126

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Combining samples using any of the weighting strategies considered resulted in smaller MSE values compared to using only one sample. Increasing the number of samples combined tended to improve the MSE values. No weighting strategy stood out from the others, they all resulted in similar MSE values.

Table 4. 23: Scenario 3: Comparing weighting strategies for a finite population generated using the skewed normal distribution, SN(100,25), N=1000 000 and n=10000 $\theta=100.056$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		100,058	1,602	1,086	3,432	100,056	100,060	1,182
2	1/n	100,056	-0,383	0,777	2,457	100,055	100,058	0,604
	$(\text{var}(\hat{\theta}_a))^{-1}$	100,056	-0,406	0,783	2,474	100,054	100,058	0,613
	$(\text{CV}(\hat{\theta}_a))^{-1}$	100,056	-0,338	0,783	2,473	100,055	100,058	0,612
	$(\text{skewness}(s_{dc}))^{-1}$	100,056	-0,175	0,783	2,474	100,055	100,058	0,613
	$(\text{kurtosis}(k_{dc}))^{-1}$	100,056	-0,272	0,782	2,472	100,055	100,058	0,612
	$\text{meff}/\text{Var}(y)$	100,056	-0,331	0,782	2,472	100,055	100,058	0,612
	$(\text{meff})^{-1}$	100,056	-0,358	0,783	2,474	100,055	100,058	0,613
3	1/n	100,056	-0,383	0,635	2,006	100,055	100,057	0,403
	$(\text{var}(\hat{\theta}_a))^{-1}$	100,056	-0,435	0,638	2,017	100,055	100,057	0,407
	$(\text{CV}(\hat{\theta}_a))^{-1}$	100,056	-0,349	0,638	2,017	100,055	100,057	0,407
	$(\text{skewness}(s_{dc}))^{-1}$	100,056	-0,214	0,640	2,022	100,055	100,057	0,409
	$(\text{kurtosis}(k_{dc}))^{-1}$	100,056	-0,272	0,638	2,017	100,055	100,057	0,407
	$\text{meff}/\text{Var}(y)$	100,056	-0,332	0,638	2,017	100,055	100,057	0,407
	$(\text{meff})^{-1}$	100,056	-0,373	0,638	2,018	100,055	100,057	0,408
5	1/n	100,056	-0,452	0,475	1,503	100,055	100,057	0,226
	$(\text{var}(\hat{\theta}_a))^{-1}$	100,056	-0,639	0,475	1,503	100,055	100,057	0,226
	$(\text{CV}(\hat{\theta}_a))^{-1}$	100,056	-0,536	0,475	1,503	100,055	100,057	0,226
	$(\text{skewness}(s_{dc}))^{-1}$	100,056	-0,396	0,476	1,505	100,055	100,057	0,227
	$(\text{kurtosis}(k_{dc}))^{-1}$	100,056	-0,447	0,475	1,503	100,055	100,057	0,226
	$\text{meff}/\text{Var}(y)$	100,056	-0,537	0,475	1,501	100,055	100,057	0,226
	$(\text{meff})^{-1}$	100,056	-0,553	0,476	1,505	100,055	100,057	0,227
10	1/n	100,057	0,301	0,342	1,081	100,056	100,057	0,117
	$(\text{var}(\hat{\theta}_a))^{-1}$	100,057	0,086	0,342	1,081	100,056	100,057	0,117
	$(\text{CV}(\hat{\theta}_a))^{-1}$	100,057	0,204	0,342	1,081	100,056	100,057	0,117
	$(\text{skewness}(s_{dc}))^{-1}$	100,057	0,310	0,342	1,081	100,056	100,057	0,117
	$(\text{kurtosis}(k_{dc}))^{-1}$	100,057	0,300	0,342	1,081	100,056	100,057	0,117
	$\text{meff}/\text{Var}(y)$	100,057	0,209	0,342	1,081	100,056	100,057	0,117
	$(\text{meff})^{-1}$	100,057	0,162	0,342	1,081	100,056	100,057	0,117
15	1/n	100,057	0,201	0,280	0,885	100,056	100,057	0,078
	$(\text{var}(\hat{\theta}_a))^{-1}$	100,056	-0,024	0,280	0,884	100,056	100,057	0,078
	$(\text{CV}(\hat{\theta}_a))^{-1}$	100,057	0,100	0,280	0,885	100,056	100,057	0,078
	$(\text{skewness}(s_{dc}))^{-1}$	100,057	0,211	0,281	0,887	100,056	100,057	0,079
	$(\text{kurtosis}(k_{dc}))^{-1}$	100,057	0,204	0,280	0,885	100,056	100,057	0,078
	$\text{meff}/\text{Var}(y)$	100,057	0,094	0,281	0,888	100,056	100,057	0,079
	$(\text{meff})^{-1}$	100,056	0,043	0,281	0,888	100,056	100,057	0,079

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy

Table 4.23 continued: Scenario 3: Comparing weighting strategies for a finite population generated using the skewed normal distribution, SN(100,25), N=1000 000 and n=10000 $\theta=100.056$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	1/n	100,057	0,329	0,244	0,771	100,056	100,057	0,060
	$(Var(\hat{\theta}_d))^{-1}$	100,057	0,098	0,244	0,771	100,056	100,057	0,059
	$(CV(\hat{\theta}_d))^{-1}$	100,057	0,225	0,244	0,771	100,056	100,057	0,060
	$(skewness(s_{dc}))^{-1}$	100,057	0,329	0,245	0,774	100,056	100,057	0,060
	$(kurtosis(k_{dc}))^{-1}$	100,057	0,330	0,244	0,772	100,056	100,057	0,060
	$mef f / Var(y)$	100,057	0,257	0,245	0,775	100,056	100,057	0,060
	$(mef f)^{-1}$	100,057	0,176	0,244	0,770	100,056	100,057	0,059

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Increasing the number of samples combined led to lower MSE values. Combining samples resulted in smaller MSE values compared to using estimates obtained from one sample. The weighting strategies resulted in very similar MSE values.

Summary

Mainly, the weighting strategies performed comparably well. The MSE values were close for the same number of samples combined, across sample sizes.

4.3.3 Simulation results for a finite population generated using the t- distribution with three degrees of freedom, under stratified sampling

The finite population was generated by a t- distribution with three degrees of freedom for this simulation. The results of comparing the weighting strategies for combining the surveys under this distribution, using design consistent weighting strategies are presented below. Appendix C1 presents further results, inclusive of naïve weighting strategies.

Table 4. 24: Scenario 1: Comparing weighting strategies for a finite population generated using the t distribution with three degrees of freedom, $t(3)(100,25)$, $N=1000\ 000$ and $n=1000$ $\theta=99.998$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		100,008	9,871	3,316	1,048	100,002	100,015	1,109
2	1/n	100,003	5,130	2,350	0,743	99,999	100,008	0,555
	$(\text{Var}(\hat{\theta}_a))^{-1}$	100,005	6,581	2,346	0,742	100,000	100,009	0,555
	$(\text{CV}(\hat{\theta}_a))^{-1}$	100,005	6,240	2,339	0,740	100,000	100,009	0,551
	$(\text{skewness}(s_{dc}))^{-1}$	100,003	4,996	2,640	0,835	99,998	100,008	0,699
	$(\text{kurtosis}(k_{dc}))^{-1}$	100,004	6,045	2,488	0,787	99,999	100,009	0,623
	$m_{eff}/\text{Var}(y)$	100,004	6,163	2,351	0,743	100,000	100,009	0,556
	$(m_{eff})^{-1}$	100,004	6,174	2,335	0,738	100,000	100,009	0,549
3	1/n	100,000	2,155	1,979	0,626	99,997	100,004	0,392
	$(\text{Var}(\hat{\theta}_a))^{-1}$	100,002	3,418	1,966	0,622	99,998	100,006	0,388
	$(\text{CV}(\hat{\theta}_a))^{-1}$	100,001	3,114	1,967	0,622	99,998	100,005	0,388
	$(\text{skewness}(s_{dc}))^{-1}$	100,001	2,456	2,376	0,751	99,996	100,005	0,565
	$(\text{kurtosis}(k_{dc}))^{-1}$	100,001	2,918	2,115	0,669	99,997	100,005	0,448
	$m_{eff}/\text{Var}(y)$	100,001	2,994	1,980	0,626	99,997	100,005	0,393
	$(m_{eff})^{-1}$	100,001	3,106	1,961	0,620	99,998	100,005	0,385
5	1/n	100,001	2,256	1,537	0,486	99,998	100,004	0,237
	$(\text{Var}(\hat{\theta}_a))^{-1}$	100,001	2,547	1,499	0,474	99,998	100,004	0,225
	$(\text{CV}(\hat{\theta}_a))^{-1}$	100,001	2,568	1,507	0,477	99,998	100,004	0,228
	$(\text{skewness}(s_{dc}))^{-1}$	100,000	1,188	2,051	0,649	99,995	100,004	0,421
	$(\text{kurtosis}(k_{dc}))^{-1}$	100,000	1,823	1,575	0,498	99,997	100,003	0,248
	$m_{eff}/\text{Var}(y)$	100,001	2,496	1,514	0,479	99,998	100,004	0,230
	$(m_{eff})^{-1}$	100,001	2,551	1,510	0,477	99,998	100,004	0,229
10	1/n	100,000	1,272	1,084	0,343	99,997	100,002	0,118
	$(\text{Var}(\hat{\theta}_a))^{-1}$	100,000	1,964	1,050	0,332	99,998	100,002	0,111
	$(\text{CV}(\hat{\theta}_a))^{-1}$	100,000	1,827	1,058	0,335	99,998	100,002	0,112
	$(\text{skewness}(s_{dc}))^{-1}$	100,002	4,164	1,699	0,537	99,999	100,006	0,290
	$(\text{kurtosis}(k_{dc}))^{-1}$	100,000	1,496	1,103	0,349	99,998	100,002	0,122
	$m_{eff}/\text{Var}(y)$	100,000	1,603	1,062	0,336	99,998	100,002	0,113
	$(m_{eff})^{-1}$	100,000	1,818	1,061	0,336	99,998	100,002	0,113
15	1/n	99,999	1,135	0,887	0,281	99,998	100,001	0,079
	$(\text{Var}(\hat{\theta}_a))^{-1}$	100,000	1,806	0,860	0,272	99,998	100,002	0,074
	$(\text{CV}(\hat{\theta}_a))^{-1}$	100,000	1,697	0,866	0,274	99,998	100,002	0,075
	$(\text{skewness}(s_{dc}))^{-1}$	100,003	5,127	1,564	0,495	100,000	100,007	0,247
	$(\text{kurtosis}(k_{dc}))^{-1}$	100,000	1,668	0,914	0,289	99,998	100,002	0,084
	$m_{eff}/\text{Var}(y)$	100,000	1,522	0,867	0,274	99,998	100,002	0,075
	$(m_{eff})^{-1}$	100,000	1,626	0,870	0,275	99,998	100,002	0,076

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.24 continued: Scenario 1: Comparing weighting strategies for a finite population generated using the t distribution with three degrees of freedom, $t(3)(100,25)$, $N=1000\ 000$ and $n=1000$ $\theta=99.998$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	$1/n$	100,000	1,221	0,772	0,244	99,998	100,001	0,060
	$(Var(\hat{\theta}_d))^{-1}$	100,000	1,811	0,749	0,237	99,999	100,002	0,056
	$(CV(\hat{\theta}_d))^{-1}$	100,000	1,758	0,754	0,238	99,999	100,002	0,057
	$(skewness(s_{dc}))^{-1}$	100,002	3,237	1,509	0,477	99,999	100,005	0,229
	$(kurtosis(k_{dc}))^{-1}$	100,000	1,677	0,788	0,249	99,998	100,002	0,062
	$meff/Var(y)$	100,000	1,693	0,759	0,240	99,999	100,002	0,058
	$(meff)^{-1}$	100,000	1,665	0,759	0,240	99,998	100,001	0,058

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Using estimates from only one sample resulted in higher MSE values compared to combining samples with all the weighting strategies considered. The MSE values tended to decrease as more samples were combined. The weighting strategies resulted in similar MSE values with the exception of the inverse of skewness which resulted in remarkably higher MSE values.

Table 4. 25: Scenario 2: Comparing weighting strategies for a finite population generated using the t distribution with three degrees of freedom, $t(3)(100,25)$, $N=1000\ 000$ and $n=5000$ $\theta=99.998$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		99,997	-1,094	1,604	5,074	99,994	100,000	2,575
2	1/n	99,997	-0,851	1,120	3,543	99,995	100,000	1,256
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,208	1,123	3,551	99,996	100,001	1,261
	$(CV(\hat{\theta}_a))^{-1}$	99,998	-0,005	1,123	3,553	99,996	100,001	1,262
	$(skewness(s_{dc}))^{-1}$	100,000	1,372	1,273	4,025	99,997	100,002	1,622
	$(kurtosis(k_{dc}))^{-1}$	99,999	0,979	1,213	3,837	99,997	100,002	1,473
	$meff/Var(y)$	99,998	-0,064	1,126	3,559	99,996	100,000	1,267
	$(meff)^{-1}$	99,998	0,027	1,123	3,550	99,996	100,001	1,261
3	1/n	99,996	-2,045	0,919	2,907	99,994	99,998	0,849
	$(Var(\hat{\theta}_a))^{-1}$	99,997	-0,985	0,928	2,934	99,996	99,999	0,862
	$(CV(\hat{\theta}_a))^{-1}$	99,997	-1,216	0,926	2,930	99,995	99,999	0,860
	$(skewness(s_{dc}))^{-1}$	99,998	-0,630	1,134	3,587	99,995	100,000	1,287
	$(kurtosis(k_{dc}))^{-1}$	99,998	-0,231	1,016	3,214	99,996	100,000	1,033
	$meff/Var(y)$	99,997	-1,291	0,927	2,931	99,995	99,999	0,861
	$(meff)^{-1}$	99,997	-1,193	0,927	2,932	99,995	99,999	0,861
5	1/n	99,997	-1,054	0,709	2,241	99,996	99,999	0,503
	$(Var(\hat{\theta}_a))^{-1}$	99,998	-0,532	0,706	2,233	99,996	99,999	0,499
	$(CV(\hat{\theta}_a))^{-1}$	99,998	-0,738	0,705	2,231	99,996	99,999	0,498
	$(skewness(s_{dc}))^{-1}$	99,999	0,241	0,980	3,100	99,997	100,000	0,961
	$(kurtosis(k_{dc}))^{-1}$	99,998	0,168	0,783	2,476	99,997	100,000	0,613
	$meff/Var(y)$	99,998	-0,782	0,708	2,238	99,996	99,999	0,501
	$(meff)^{-1}$	99,998	-0,703	0,706	2,233	99,996	99,999	0,499
10	1/n	99,998	-0,045	0,499	1,577	99,997	99,999	0,249
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,386	0,493	1,560	99,998	100,000	0,243
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,229	0,494	1,563	99,998	100,000	0,244
	$(skewness(s_{dc}))^{-1}$	99,999	0,677	0,792	2,506	99,997	100,001	0,628
	$(kurtosis(k_{dc}))^{-1}$	99,999	0,667	0,542	1,713	99,998	100,000	0,294
	$meff/Var(y)$	99,998	0,172	0,495	1,564	99,998	99,999	0,245
	$(meff)^{-1}$	99,999	0,269	0,494	1,563	99,998	100,000	0,244
15	1/n	99,998	0,033	0,407	1,287	99,998	99,999	0,166
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,438	0,402	1,270	99,998	100,000	0,161
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,300	0,403	1,274	99,998	99,999	0,162
	$(skewness(s_{dc}))^{-1}$	99,999	0,721	0,744	2,352	99,998	100,000	0,554
	$(kurtosis(k_{dc}))^{-1}$	99,999	0,591	0,433	1,369	99,998	100,000	0,188
	$meff/Var(y)$	99,999	0,242	0,404	1,277	99,998	99,999	0,163
	$(meff)^{-1}$	99,999	0,295	0,403	1,275	99,998	99,999	0,163

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.25 continued: Scenario 2: Comparing weighting strategies for a finite population generated using the t distribution with three degrees of freedom, $t(3)(100,25)$, $N=1000\ 000$ and $n=5000$ $\theta=99.998$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	$1/n$	99,998	-0,395	0,356	1,127	99,997	99,999	0,127
	$(Var(\hat{\theta}_d))^{-1}$	99,998	0,018	0,352	1,113	99,998	99,999	0,124
	$(CV(\hat{\theta}_d))^{-1}$	99,998	-0,122	0,353	1,116	99,998	99,999	0,125
	$(skewness(s_{dc}))^{-1}$	99,999	0,954	0,697	2,205	99,998	100,001	0,487
	$(kurtosis(k_{dc}))^{-1}$	99,998	0,160	0,380	1,202	99,998	99,999	0,145
	$meff/Var(y)$	99,998	-0,178	0,353	1,118	99,997	99,999	0,125
	$(meff)^{-1}$	99,998	-0,131	0,353	1,117	99,997	99,999	0,125

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Weighting the samples using the inverse of skewness resulted in conspicuously larger MSE values compared to the other weighting strategies. This is not surprising because the superpopulation distribution is not skewed.

Table 4. 26: Scenario 3: Comparing weighting strategies for a finite population generated using the t distribution with three degrees of freedom, $t(3)(100,25)$, $N=1000\ 000$ and $n=10000$ $\theta=99.998$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		100,000	1,207	1,093	3,456	99,997	100,002	1,196
2	1/n	99,999	0,622	0,783	2,477	99,997	100,000	0,614
	$(Var(\hat{\theta}_a))^{-1}$	100,000	1,354	0,789	2,496	99,998	100,001	0,625
	$(CV(\hat{\theta}_a))^{-1}$	100,000	1,207	0,788	2,490	99,998	100,001	0,622
	$(skewness(s_{dc}))^{-1}$	100,000	2,095	0,887	2,804	99,999	100,002	0,791
	$(kurtosis(k_{dc}))^{-1}$	100,000	2,179	0,860	2,720	99,999	100,002	0,744
	$meff/Var(y)$	99,999	1,173	0,787	2,489	99,998	100,001	0,621
	$(meff)^{-1}$	100,000	1,228	0,788	2,493	99,998	100,001	0,623
3	1/n	99,999	1,149	0,635	2,009	99,998	100,001	0,405
	$(Var(\hat{\theta}_a))^{-1}$	100,000	1,723	0,635	2,009	99,999	100,001	0,406
	$(CV(\hat{\theta}_a))^{-1}$	100,000	1,608	0,635	2,007	99,999	100,001	0,405
	$(skewness(s_{dc}))^{-1}$	100,001	2,296	0,776	2,453	99,999	100,002	0,607
	$(kurtosis(k_{dc}))^{-1}$	100,001	2,317	0,703	2,222	99,999	100,002	0,499
	$meff/Var(y)$	100,000	1,600	0,635	2,007	99,999	100,001	0,405
	$(meff)^{-1}$	100,000	1,616	0,635	2,009	99,999	100,001	0,406
5	1/n	99,999	0,683	0,479	1,516	99,998	100,000	0,230
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,996	0,476	1,505	99,998	100,000	0,228
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,862	0,477	1,508	99,998	100,000	0,228
	$(skewness(s_{dc}))^{-1}$	100,000	2,084	0,658	2,081	99,999	100,002	0,437
	$(kurtosis(k_{dc}))^{-1}$	100,000	1,607	0,534	1,690	99,999	100,001	0,288
	$meff/Var(y)$	99,999	0,859	0,479	1,513	99,998	100,000	0,230
	$(meff)^{-1}$	99,999	0,866	0,476	1,507	99,998	100,000	0,228
10	1/n	99,999	0,210	0,346	1,093	99,998	99,999	0,119
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,533	0,344	1,088	99,998	100,000	0,119
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,402	0,344	1,089	99,998	99,999	0,119
	$(skewness(s_{dc}))^{-1}$	100,000	1,764	0,544	1,719	99,999	100,001	0,299
	$(kurtosis(k_{dc}))^{-1}$	99,999	1,024	0,376	1,191	99,999	100,000	0,143
	$meff/Var(y)$	99,999	0,336	0,345	1,090	99,998	99,999	0,119
	$(meff)^{-1}$	99,999	0,407	0,345	1,091	99,998	99,999	0,119
15	1/n	99,998	-0,010	0,282	0,892	99,998	99,999	0,079
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,273	0,281	0,887	99,998	99,999	0,079
	$(CV(\hat{\theta}_a))^{-1}$	99,998	0,161	0,281	0,888	99,998	99,999	0,079
	$(skewness(s_{dc}))^{-1}$	100,000	1,297	0,520	1,643	99,999	100,001	0,272
	$(kurtosis(k_{dc}))^{-1}$	99,999	0,654	0,311	0,984	99,998	100,000	0,097
	$meff/Var(y)$	99,998	0,130	0,281	0,888	99,998	99,999	0,079
	$(meff)^{-1}$	99,999	0,190	0,281	0,888	99,998	99,999	0,079

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy

Table 4.26 continued: Scenario 3: Comparing weighting strategies for a finite population generated using the t distribution with three degrees of freedom, $t(3)(100,25)$, $N=1000\ 000$ and $n=10000$ $\theta=99.998$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	$1/n$	99,998	0,006	0,240	0,759	99,998	99,999	0,058
	$(Var(\hat{\theta}_d))^{-1}$	99,999	0,275	0,238	0,753	99,998	99,999	0,057
	$(CV(\hat{\theta}_d))^{-1}$	99,998	0,171	0,239	0,755	99,998	99,999	0,057
	$(skewness(s_{dc}))^{-1}$	99,999	0,703	0,506	1,599	99,998	100,000	0,256
	$(kurtosis(k_{dc}))^{-1}$	99,999	0,538	0,266	0,841	99,998	99,999	0,071
	$mef f / Var(y)$	99,998	0,131	0,239	0,757	99,998	99,999	0,057
	$(mef f)^{-1}$	99,999	0,183	0,238	0,754	99,998	99,999	0,057

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

All the weighting strategies resulted in similar MSE values except the inverse of skewness, just like in scenarios one and two.

Summary

The weighting strategies resulted in similar MSE values. It appears most of the weighting strategies studied could be used with a t distribution with 3 degrees of freedom superpopulation model. The inverse of skewness did not perform well, resulting in distinctly higher MSE values. The inverse of kurtosis did not perform as well as expected, resulting in MSE values comparable to the other weighting strategies even though the superpopulation distribution had a higher amount of kurtosis.

4.3.4 Simulation results for a finite population generated using skewed t- distribution with three degrees of freedom, under stratified sampling

Weighting strategies were compared where a finite population was generated using a skewed t-distribution with three degrees of freedom. A location parameter of -0.7891, a scale parameter of 0.746 and an asymmetry parameter of 3 were used to generate the skewed distribution. Appendix D1 presents further results of this comparison, with results from design consistent weighting strategies presented in tables 4.27 to 4.29.

Table 4. 27: Scenario 1: Comparing weighting strategies for a finite population generated using a skewed t distribution with three degrees of freedom, at(3)(100,25), N=1000 000 and n=1000 $\theta=99.970$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^2$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,974	0,396	3,509	1,110	99,967	99,981	1,233
2	1/n	99,975	0,490	2,469	0,781	99,970	99,980	0,612
	$(Var(\hat{\theta}_a))^{-1}$	99,965	-0,563	2,474	0,783	99,960	99,969	0,615
	$(CV(\hat{\theta}_a))^{-1}$	99,969	-0,098	2,459	0,778	99,964	99,974	0,605
	$(skewness(s_{dc}))^{-1}$	99,960	-1,016	2,626	0,831	99,955	99,965	0,700
	$(kurtosis(k_{dc}))^{-1}$	99,960	-1,004	2,681	0,848	99,955	99,965	0,729
	$meff/Var(y)$	99,970	-0,041	2,455	0,776	99,965	99,975	0,603
	$(meff)^{-1}$	99,969	-0,157	2,473	0,782	99,964	99,974	0,612
3	1/n	99,974	0,373	1,980	0,626	99,970	99,978	0,394
	$(Var(\hat{\theta}_a))^{-1}$	99,961	-0,912	1,980	0,626	99,957	99,965	0,400
	$(CV(\hat{\theta}_a))^{-1}$	99,967	-0,331	1,964	0,621	99,963	99,971	0,387
	$(skewness(s_{dc}))^{-1}$	99,955	-1,514	2,134	0,675	99,951	99,959	0,478
	$(kurtosis(k_{dc}))^{-1}$	99,956	-1,428	2,182	0,690	99,952	99,960	0,497
	$meff/Var(y)$	99,968	-0,253	1,959	0,620	99,964	99,972	0,384
	$(meff)^{-1}$	99,966	-0,411	1,981	0,627	99,962	99,970	0,394
5	1/n	99,971	0,033	1,543	0,488	99,968	99,974	0,238
	$(Var(\hat{\theta}_a))^{-1}$	99,956	-1,448	1,528	0,484	99,953	99,959	0,255
	$(CV(\hat{\theta}_a))^{-1}$	99,963	-0,745	1,517	0,480	99,960	99,966	0,236
	$(skewness(s_{dc}))^{-1}$	99,948	-2,242	1,657	0,524	99,945	99,951	0,325
	$(kurtosis(k_{dc}))^{-1}$	99,949	-2,082	1,665	0,527	99,946	99,953	0,320
	$meff/Var(y)$	99,964	-0,639	1,508	0,477	99,961	99,967	0,231
	$(meff)^{-1}$	99,962	-0,848	1,536	0,486	99,959	99,965	0,243
10	1/n	99,971	0,075	1,077	0,341	99,969	99,973	0,116
	$(Var(\hat{\theta}_a))^{-1}$	99,955	-1,514	1,072	0,339	99,953	99,957	0,138
	$(CV(\hat{\theta}_a))^{-1}$	99,963	-0,753	1,062	0,336	99,961	99,965	0,118
	$(skewness(s_{dc}))^{-1}$	99,946	-2,438	1,190	0,377	99,944	99,948	0,201
	$(kurtosis(k_{dc}))^{-1}$	99,948	-2,221	1,189	0,376	99,946	99,950	0,191
	$meff/Var(y)$	99,964	-0,643	1,062	0,336	99,962	99,966	0,117
	$(meff)^{-1}$	99,962	-0,860	1,071	0,339	99,960	99,964	0,122
15	1/n	99,971	0,109	0,879	0,278	99,970	99,973	0,077
	$(Var(\hat{\theta}_a))^{-1}$	99,955	-1,522	0,876	0,277	99,953	99,957	0,100
	$(CV(\hat{\theta}_a))^{-1}$	99,963	-0,740	0,868	0,274	99,961	99,965	0,081
	$(skewness(s_{dc}))^{-1}$	99,946	-2,455	0,980	0,310	99,944	99,948	0,156
	$(kurtosis(k_{dc}))^{-1}$	99,948	-2,210	0,971	0,307	99,946	99,950	0,143
	$meff/Var(y)$	99,964	-0,610	0,868	0,275	99,962	99,966	0,079
	$(meff)^{-1}$	99,962	-0,858	0,876	0,277	99,960	99,963	0,084

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.27 continued: Scenario 1: Comparing weighting strategies for a finite population generated using a skewed t distribution with three degrees of freedom, at(3)(100,25), N=1000 000 and n=1000 $\theta=99.970$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	1/n	99,970	0,025	0,773	0,245	99,969	99,972	0,060
	$(Var(\hat{\theta}_d))^{-1}$	99,954	-1,637	0,767	0,243	99,952	99,955	0,086
	$(CV(\hat{\theta}_d))^{-1}$	99,962	-0,840	0,761	0,241	99,960	99,963	0,065
	$(skewness(s_{dc}))^{-1}$	99,944	-2,607	0,864	0,273	99,942	99,946	0,142
	$(kurtosis(k_{dc}))^{-1}$	99,947	-2,346	0,854	0,270	99,945	99,948	0,128
	$meff/Var(y)$	99,963	-0,702	0,763	0,241	99,962	99,965	0,063
	$(meff)^{-1}$	99,961	-0,964	0,764	0,242	99,959	99,962	0,068

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Unexpectedly, the inverses of skewness and kurtosis appeared to have slightly higher MSE values. In general, combining samples resulted in improved estimates which kept improving as the number of samples combined was increased.

Table 4. 28: Scenario 2: Comparing weighting strategies for a finite population generated using a skewed t distribution with three degrees of freedom, at(3)(100,25), N=1000 000 and n=5000 $\theta=99.970$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		99,971	0,607	1,533	4,848	99,968	99,974	2,350
2	1/n	99,972	1,321	1,079	3,413	99,969	99,974	1,166
	$(\text{Var}(\hat{\theta}_a))^{-1}$	99,969	-0,763	1,082	3,423	99,967	99,972	1,171
	$(\text{CV}(\hat{\theta}_a))^{-1}$	99,971	0,295	1,082	3,422	99,968	99,973	1,170
	$(\text{skewness}(s_{dc}))^{-1}$	99,967	-2,793	1,130	3,574	99,965	99,970	1,285
	$(\text{kurtosis}(k_{dc}))^{-1}$	99,967	-3,599	1,186	3,751	99,964	99,969	1,419
	$\text{meff}/\text{Var}(y)$	99,971	0,429	1,081	3,418	99,969	99,973	1,168
	$(\text{meff})^{-1}$	99,970	0,143	1,084	3,430	99,968	99,973	1,176
3	1/n	99,971	0,276	0,859	2,716	99,969	99,972	0,738
	$(\text{Var}(\hat{\theta}_a))^{-1}$	99,968	-2,510	0,864	2,733	99,966	99,969	0,753
	$(\text{CV}(\hat{\theta}_a))^{-1}$	99,969	-1,090	0,862	2,727	99,967	99,971	0,744
	$(\text{skewness}(s_{dc}))^{-1}$	99,965	-5,156	0,917	2,902	99,963	99,967	0,868
	$(\text{kurtosis}(k_{dc}))^{-1}$	99,964	-6,340	0,976	3,088	99,962	99,966	0,993
	$\text{meff}/\text{Var}(y)$	99,969	-0,846	0,860	2,722	99,968	99,971	0,741
	$(\text{meff})^{-1}$	99,969	-1,348	0,865	2,736	99,967	99,971	0,750
5	1/n	99,971	0,298	0,676	2,137	99,969	99,972	0,456
	$(\text{Var}(\hat{\theta}_a))^{-1}$	99,967	-3,077	0,674	2,133	99,966	99,968	0,464
	$(\text{CV}(\hat{\theta}_a))^{-1}$	99,969	-1,404	0,673	2,128	99,968	99,970	0,454
	$(\text{skewness}(s_{dc}))^{-1}$	99,964	-6,190	0,716	2,266	99,963	99,965	0,552
	$(\text{kurtosis}(k_{dc}))^{-1}$	99,963	-7,573	0,763	2,413	99,961	99,964	0,639
	$\text{meff}/\text{Var}(y)$	99,969	-1,040	0,674	2,133	99,968	99,971	0,456
	$(\text{meff})^{-1}$	99,969	-1,685	0,677	2,140	99,967	99,970	0,461
10	1/n	99,971	0,582	0,488	1,542	99,970	99,972	0,238
	$(\text{Var}(\hat{\theta}_a))^{-1}$	99,967	-3,225	0,480	1,520	99,966	99,968	0,241
	$(\text{CV}(\hat{\theta}_a))^{-1}$	99,969	-1,342	0,482	1,525	99,968	99,970	0,234
	$(\text{skewness}(s_{dc}))^{-1}$	99,963	-6,963	0,511	1,615	99,962	99,964	0,309
	$(\text{kurtosis}(k_{dc}))^{-1}$	99,962	-8,589	0,544	1,722	99,961	99,963	0,370
	$\text{meff}/\text{Var}(y)$	99,969	-1,004	0,483	1,528	99,968	99,970	0,234
	$(\text{meff})^{-1}$	99,969	-1,696	0,483	1,527	99,968	99,969	0,236
15	1/n	99,971	0,396	0,406	1,284	99,970	99,971	0,165
	$(\text{Var}(\hat{\theta}_a))^{-1}$	99,967	-3,550	0,402	1,271	99,966	99,967	0,174
	$(\text{CV}(\hat{\theta}_a))^{-1}$	99,969	-1,602	0,402	1,273	99,968	99,969	0,164
	$(\text{skewness}(s_{dc}))^{-1}$	99,963	-7,491	0,427	1,352	99,962	99,964	0,239
	$(\text{kurtosis}(k_{dc}))^{-1}$	99,961	-9,198	0,456	1,443	99,960	99,962	0,293
	$\text{meff}/\text{Var}(y)$	99,969	-1,316	0,403	1,276	99,968	99,970	0,164
	$(\text{meff})^{-1}$	99,968	-1,970	0,403	1,275	99,967	99,969	0,166

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.28 continued: Scenario 2: Comparing weighting strategies for a finite population generated using a skewed t distribution with three degrees of freedom, at(3)(100,25), N=1000 000 and n=5000 $\theta=99.970$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	$1/n$	99,970	0,236	0,342	1,083	99,970	99,971	0,117
	$(Var(\hat{\theta}_d))^{-1}$	99,966	-3,774	0,338	1,070	99,966	99,967	0,129
	$(CV(\hat{\theta}_d))^{-1}$	99,968	-1,794	0,339	1,072	99,968	99,969	0,118
	$(skewness(s_{dc}))^{-1}$	99,963	-7,716	0,361	1,141	99,962	99,963	0,190
	$(kurtosis(k_{dc}))^{-1}$	99,961	-9,421	0,387	1,223	99,960	99,962	0,238
	$meff/Var(y)$	99,969	-1,411	0,340	1,076	99,968	99,970	0,118
	$(meff)^{-1}$	99,968	-2,152	0,340	1,074	99,967	99,969	0,120

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

The inverses of skewness and kurtosis resulted in higher MSE values compared to all the other weighting strategies for all the numbers of samples combined. The other weighting strategies resulted in similar MSE values.

Table 4. 29: Scenario 3: Comparing weighting strategies for a finite population generated using a skewed t distribution with three degrees of freedom, at(3)(100,25), N=1000 000 and n=10000 $\theta=99.970$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	$SE(\hat{\theta}_c)$ $\times 10^{-3}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-2}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-3}$
						LB	UB	
1		99,971	1,173	1,083	3,425	99,969	99,974	1,174
2	1/n	99,971	0,885	0,757	2,395	99,970	99,973	0,574
	$(Var(\hat{\theta}_a))^{-1}$	99,970	-0,312	0,763	2,413	99,968	99,971	0,582
	$(CV(\hat{\theta}_a))^{-1}$	99,970	0,229	0,761	2,409	99,969	99,972	0,580
	$(skewness(s_{dc}))^{-1}$	99,969	-1,508	0,796	2,516	99,967	99,970	0,635
	$(kurtosis(k_{dc}))^{-1}$	99,968	-2,202	0,835	2,643	99,966	99,970	0,703
	$meff/Var(y)$	99,971	0,320	0,762	2,411	99,969	99,972	0,581
	$(meff)^{-1}$	99,970	0,129	0,761	2,408	99,969	99,972	0,579
3	1/n	99,971	0,761	0,623	1,969	99,970	99,972	0,388
	$(Var(\hat{\theta}_a))^{-1}$	99,969	-0,833	0,625	1,978	99,968	99,971	0,392
	$(CV(\hat{\theta}_a))^{-1}$	99,970	-0,092	0,625	1,977	99,969	99,971	0,391
	$(skewness(s_{dc}))^{-1}$	99,968	-2,501	0,667	2,109	99,966	99,969	0,451
	$(kurtosis(k_{dc}))^{-1}$	99,967	-3,406	0,722	2,284	99,965	99,968	0,533
	$meff/Var(y)$	99,970	0,011	0,627	1,982	99,969	99,971	0,393
	$(meff)^{-1}$	99,970	-0,215	0,624	1,974	99,969	99,971	0,390
5	1/n	99,970	-0,027	0,494	1,564	99,969	99,971	0,245
	$(Var(\hat{\theta}_a))^{-1}$	99,969	-1,744	0,493	1,561	99,968	99,969	0,247
	$(CV(\hat{\theta}_a))^{-1}$	99,969	-0,885	0,493	1,560	99,968	99,970	0,244
	$(skewness(s_{dc}))^{-1}$	99,967	-3,717	0,521	1,648	99,966	99,968	0,285
	$(kurtosis(k_{dc}))^{-1}$	99,965	-4,806	0,565	1,788	99,964	99,967	0,343
	$meff/Var(y)$	99,969	-0,775	0,492	1,555	99,969	99,970	0,242
	$(meff)^{-1}$	99,969	-1,034	0,494	1,562	99,968	99,970	0,245
10	1/n	99,971	0,439	0,353	1,116	99,970	99,971	0,125
	$(Var(\hat{\theta}_a))^{-1}$	99,969	-1,495	0,351	1,111	99,968	99,969	0,126
	$(CV(\hat{\theta}_a))^{-1}$	99,970	-0,530	0,351	1,111	99,969	99,970	0,124
	$(skewness(s_{dc}))^{-1}$	99,967	-3,705	0,371	1,175	99,966	99,967	0,152
	$(kurtosis(k_{dc}))^{-1}$	99,965	-5,003	0,404	1,278	99,964	99,966	0,188
	$meff/Var(y)$	99,970	-0,354	0,352	1,113	99,969	99,971	0,124
	$(meff)^{-1}$	99,970	-0,724	0,351	1,111	99,969	99,970	0,124
15	1/n	99,971	0,412	0,277	0,876	99,970	99,971	0,077
	$(Var(\hat{\theta}_a))^{-1}$	99,969	-1,603	0,275	0,868	99,968	99,969	0,078
	$(CV(\hat{\theta}_a))^{-1}$	99,970	-0,599	0,275	0,870	99,969	99,970	0,076
	$(skewness(s_{dc}))^{-1}$	99,966	-3,858	0,293	0,927	99,966	99,967	0,101
	$(kurtosis(k_{dc}))^{-1}$	99,965	-5,191	0,323	1,021	99,964	99,966	0,131
	$meff/Var(y)$	99,970	-0,385	0,278	0,878	99,969	99,970	0,077
	$(meff)^{-1}$	99,969	-0,810	0,275	0,869	99,969	99,970	0,076

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, $SE(\hat{\theta}_c)$ =standard error of the combined estimate, $CV(\hat{\theta}_c)$ (%) = coefficient of variation, $CI_{95\%}(\hat{\theta}_c)$ = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.29 continued: Scenario 3: Comparing weighting strategies for a finite population generated using a skewed t distribution with three degrees of freedom, at(3)(100,25), N=1000 000 and n=10000 $\theta=99.970$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	$1/n$	99,971	0,467	0,242	0,767	99,970	99,971	0,059
	$(Var(\hat{\theta}_d))^{-1}$	99,969	-1,598	0,241	0,761	99,968	99,969	0,060
	$(CV(\hat{\theta}_d))^{-1}$	99,970	-0,569	0,241	0,762	99,969	99,970	0,058
	$(skewness(s_{dc}))^{-1}$	99,966	-3,948	0,257	0,814	99,966	99,967	0,082
	$(kurtosis(k_{dc}))^{-1}$	99,965	-5,336	0,283	0,894	99,964	99,965	0,108
	$mef f / Var(y)$	99,970	-0,390	0,242	0,766	99,969	99,970	0,059
	$(mef f)^{-1}$	99,969	-0,772	0,241	0,762	99,969	99,970	0,059

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Weighting samples by the inverses of skewness and kurtosis resulted in higher MSE values compared to the other weighting strategies. The other weighting strategies performed comparably well.

Summary

All the weighting strategies appeared to work well for the finite population generated using a skewed t distribution with three degrees of freedom except the inverses of skewness and kurtosis.

4.3.7 Summary of simulation results under stratified random sampling

Under stratified random sampling, the superpopulation appeared to influence the choice of the weighting strategy. The MSE values tended to decrease with increasing sample size and number of samples combined. Combining samples resulted in better estimates compared to using only one sample. All the weighting strategies except the inverse of skewness worked well with the normal distribution superpopulation. The weighting strategies performed similarly with the skewed normal distribution. All the weighting strategies, except the inverse of skewness worked well with the symmetric t-distribution with three degrees of freedom. This was expected of symmetric superpopulations. The inverses of skewness and kurtosis appeared not to work well with the skewed t-distributed superpopulation with three degrees of freedom. This was not expected and warrants further investigation with larger numbers of samples combined. The use of naïve estimates which do not consider sampling design also influenced the performance of weighting strategies resulting in higher or lower MSE values depending on the weighting strategy and superpopulation distribution.

4.4 Simulation results under two-stage simple cluster sampling

Enumeration areas are defined as the smallest geographical unit (piece of land) into which the country is divided for census or survey enumeration, of a size able to be enumerated by one fieldworker in the allocated period. In South Africa, enumeration areas (EAs) typically contain between 100 and 250 households. A clustered population of size 1 000 000 and clusters of size 250 (to resemble a typical size of EAs) was simulated according to the super-population model

$$Y_{ij} = \mu + u_i + v_{ij},$$

where μ is the overall mean of the observed characteristic Y_{ij} , u_i are the cluster random effects (characteristic common to all members of the same cluster) and v_{ij} are the individual random effects (individual specific characteristic). The cluster random effects were simulated to have a mean of zero and a variance of five whereas the individual random effects had a mean of zero and a variance of twenty for all the distributions considered.

For the sample size of 1000, fifty clusters of size two hundred and fifty were selected by simple random sampling without replacement, and twenty units were then selected also by simple random sampling without replacement from each of the selected clusters. Similarly, twenty units were selected from 250 clusters for the sample size 5000, and twenty units were selected from 500 clusters for the sample size 10 000.

Misspecification effects were calculated for each sample replicate. The misspecification effects for each distribution and sample size are presented in table 4.4.1.

Table 4. 30: Misspecification effects for all the distributions and sample sizes considered under cluster sampling

Distribution	Sample size	Misspecification effect
Normal	1000	2,540
	5000	4,976
	10000	4,893
Skewed normal	1000	4.898
	5000	5.011
	10000	4,846
t, 3 d.f	1000	4.564
	5000	4.683
	10000	4.589
skewed t, 3 d.f	1000	3,679
	5000	4.054
	10000	5.227

The misspecification effects were larger than one; hence ignoring clustering in the analysis would result in underestimation of variance, $Var(\hat{\theta})$. They were mostly close to five.

Similarly, to simple random sampling and stratified sampling, weighting by a weight of one and the inverse of sample size resulted in identical MSE values. We present only the results from weighting by the inverse of sample size.

4.4.1 Simulation results for the normal distribution, under cluster sampling

Multiple samples were selected from a finite population generated under the normal distribution. The results of comparing the design consistent weighting strategies for combining the multiple surveys are presented in the tables below. Appendix A3 presents the more comprehensive results, which include the naïve weighting strategies.

Table 4. 31: Scenario 1: Combining D samples of size n=1000, from a population of size N=1000 000, distributed as the normal distribution N(100,25), $\theta=100.031$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,975	-5,665	2,565	2,566	99,924	100,025	6,902
2	1/n	99,996	-3,477	1,371	1,371	99,970	100,023	2,002
	$(Var(\hat{\theta}_a))^{-1}$	99,985	-4,601	1,672	1,672	99,952	100,018	3,006
	$(CV(\hat{\theta}_a))^{-1}$	99,986	-4,528	1,662	1,662	99,953	100,019	2,966
	$(skewness(s_{dc}))^{-1}$	100,006	-2,533	1,794	1,794	99,971	100,041	3,282
	$(kurtosis(k_{dc}))^{-1}$	99,986	-4,519	1,662	1,662	99,953	100,019	2,967
	$meff/Var(y)$	99,986	-4,516	1,655	1,655	99,954	100,019	2,943
	$(meff)^{-1}$	99,985	-4,590	1,673	1,673	99,953	100,018	3,010
3	1/n	99,990	-4,168	1,175	1,175	99,967	100,013	1,553
	$(Var(\hat{\theta}_a))^{-1}$	99,987	-4,397	1,340	1,341	99,961	100,014	1,990
	$(CV(\hat{\theta}_a))^{-1}$	99,987	-4,377	1,321	1,321	99,962	100,013	1,937
	$(skewness(s_{dc}))^{-1}$	100,008	-2,360	1,681	1,681	99,975	100,041	2,882
	$(kurtosis(k_{dc}))^{-1}$	99,986	-4,486	1,318	1,318	99,961	100,012	1,938
	$meff/Var(y)$	99,989	-4,238	1,338	1,339	99,963	100,015	1,971
	$(meff)^{-1}$	99,987	-4,449	1,333	1,333	99,961	100,013	1,974
5	1/n	99,995	-3,579	1,103	1,103	99,974	100,017	1,345
	$(Var(\hat{\theta}_a))^{-1}$	99,996	-3,510	1,108	1,108	99,974	100,018	1,352
	$(CV(\hat{\theta}_a))^{-1}$	99,996	-3,501	1,102	1,103	99,975	100,018	1,338
	$(skewness(s_{dc}))^{-1}$	99,999	-3,191	1,527	1,527	99,969	100,029	2,434
	$(kurtosis(k_{dc}))^{-1}$	99,995	-3,610	1,108	1,109	99,973	100,017	1,359
	$meff/Var(y)$	99,999	-3,207	1,213	1,213	99,975	100,023	1,575
	$(meff)^{-1}$	99,997	-3,428	1,161	1,161	99,974	100,020	1,466
10	1/n	99,996	-3,501	0,728	0,728	99,982	100,010	0,653
	$(Var(\hat{\theta}_a))^{-1}$	99,997	-3,438	0,727	0,727	99,983	100,011	0,647
	$(CV(\hat{\theta}_a))^{-1}$	99,997	-3,419	0,724	0,724	99,983	100,011	0,641
	$(skewness(s_{dc}))^{-1}$	100,002	-2,926	1,370	1,370	99,975	100,029	1,961
	$(kurtosis(k_{dc}))^{-1}$	99,996	-3,530	0,727	0,727	99,982	100,010	0,653
	$meff/Var(y)$	99,995	-3,644	0,759	0,759	99,980	100,010	0,708
	$(meff)^{-1}$	99,996	-3,564	0,740	0,740	99,981	100,010	0,674
15	1/n	99,999	-3,242	0,611	0,611	99,987	100,011	0,478
	$(Var(\hat{\theta}_a))^{-1}$	99,998	-3,286	0,594	0,594	99,987	100,010	0,460
	$(CV(\hat{\theta}_a))^{-1}$	99,999	-3,217	0,600	0,600	99,987	100,011	0,463
	$(skewness(s_{dc}))^{-1}$	100,001	-3,036	1,355	1,355	99,974	100,027	1,929
	$(kurtosis(k_{dc}))^{-1}$	99,999	-3,272	0,613	0,613	99,986	100,011	0,483
	$meff/Var(y)$	99,997	-3,429	0,646	0,646	99,984	100,010	0,535
	$(meff)^{-1}$	99,997	-3,419	0,610	0,610	99,985	100,009	0,489

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.31 continued: Scenario 1: Combining D samples of size n=1000, from a population of size N=1000 000, distributed as the normal distribution N(100,25), $\theta=100.031$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	1/n	100,000	-3,098	0,537	0,537	99,990	100,011	0,384
	$(Var(\hat{\theta}_d))^{-1}$	100,000	-3,129	0,528	0,528	99,990	100,010	0,377
	$(CV(\hat{\theta}_d))^{-1}$	100,001	-3,061	0,530	0,530	99,990	100,011	0,375
	$(skewness(s_{dc}))^{-1}$	100,006	-2,536	1,304	1,303	99,980	100,031	1,764
	$(kurtosis(k_{dc}))^{-1}$	100,000	-3,131	0,540	0,540	99,989	100,010	0,389
	$meff/Var(y)$	100,001	-3,055	0,539	0,539	99,990	100,011	0,384
	$(meff)^{-1}$	100,000	-3,125	0,529	0,529	99,990	100,010	0,377

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

In this scenario, the weighting strategies performed similarly, resulting in MSE values that were close in magnitude except the inverse of skewness. The relative difference in magnitude between the MSE values resulting from the use of the inverse of skewness and the other weighting strategies became more pronounced as the number of samples combined increased. Combining the samples using any one of the weighting strategies resulted in better estimates compared to a single sample.

Table 4. 32: Scenario 2: Combining D samples of size n=5000, from a population of size N=1000 000, distributed as the normal distribution N(100,25), $\theta=100.031$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		100,035	4,140	1,404	1,403	100,008	100,063	1,972
2	1/n	100,024	-7,345	0,908	0,908	100,006	100,042	0,830
	$(Var(\hat{\theta}_a))^{-1}$	100,024	-6,854	1,041	1,041	100,004	100,045	1,089
	$(CV(\hat{\theta}_a))^{-1}$	100,024	-6,775	1,043	1,042	100,004	100,045	1,092
	$(skewness(s_{dc}))^{-1}$	100,032	0,862	1,314	1,314	100,006	100,058	1,727
	$(kurtosis(k_{dc}))^{-1}$	100,024	-7,210	1,048	1,048	100,003	100,045	1,104
	$meff/Var(y)$	100,024	-6,793	1,043	1,043	100,004	100,045	1,093
	$(meff)^{-1}$	100,024	-7,018	1,043	1,043	100,004	100,045	1,093
3	1/n	100,028	-3,697	0,711	0,710	100,014	100,041	0,506
	$(Var(\hat{\theta}_a))^{-1}$	100,029	-2,419	0,816	0,816	100,013	100,045	0,667
	$(CV(\hat{\theta}_a))^{-1}$	100,029	-2,445	0,820	0,820	100,013	100,045	0,673
	$(skewness(s_{dc}))^{-1}$	100,032	0,807	1,158	1,158	100,009	100,055	1,342
	$(kurtosis(k_{dc}))^{-1}$	100,028	-3,239	0,826	0,826	100,012	100,044	0,684
	$meff/Var(y)$	100,029	-2,539	0,822	0,822	100,013	100,045	0,676
	$(meff)^{-1}$	100,029	-2,520	0,819	0,819	100,013	100,045	0,671
5	1/n	100,031	-0,506	0,614	0,613	100,019	100,043	0,377
	$(Var(\hat{\theta}_a))^{-1}$	100,031	0,247	0,610	0,610	100,020	100,043	0,372
	$(CV(\hat{\theta}_a))^{-1}$	100,031	0,047	0,611	0,611	100,019	100,043	0,373
	$(skewness(s_{dc}))^{-1}$	100,038	6,994	1,076	1,075	100,017	100,059	1,162
	$(kurtosis(k_{dc}))^{-1}$	100,030	-0,854	0,613	0,613	100,018	100,042	0,376
	$meff/Var(y)$	100,030	-0,980	0,621	0,620	100,018	100,042	0,385
	$(meff)^{-1}$	100,031	0,199	0,611	0,610	100,019	100,043	0,373
10	1/n	100,027	-4,396	0,460	0,460	100,018	100,036	0,214
	$(Var(\hat{\theta}_a))^{-1}$	100,027	-4,376	0,460	0,460	100,018	100,036	0,213
	$(CV(\hat{\theta}_a))^{-1}$	100,027	-4,188	0,459	0,459	100,018	100,036	0,213
	$(skewness(s_{dc}))^{-1}$	100,037	5,317	0,982	0,982	100,017	100,056	0,968
	$(kurtosis(k_{dc}))^{-1}$	100,027	-4,529	0,461	0,461	100,018	100,036	0,215
	$meff/Var(y)$	100,027	-4,077	0,462	0,462	100,018	100,036	0,215
	$(meff)^{-1}$	100,027	-4,361	0,460	0,460	100,018	100,036	0,214
15	1/n	100,028	-2,724	0,389	0,389	100,021	100,036	0,152
	$(Var(\hat{\theta}_a))^{-1}$	100,028	-2,792	0,387	0,387	100,021	100,036	0,151
	$(CV(\hat{\theta}_a))^{-1}$	100,029	-2,553	0,388	0,388	100,021	100,036	0,151
	$(skewness(s_{dc}))^{-1}$	100,036	4,938	0,848	0,848	100,020	100,053	0,722
	$(kurtosis(k_{dc}))^{-1}$	100,028	-2,840	0,391	0,391	100,021	100,036	0,153
	$meff/Var(y)$	100,026	-5,200	0,468	0,468	100,017	100,035	0,222
	$(meff)^{-1}$	100,027	-4,270	0,422	0,422	100,019	100,035	0,180

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.32 continued: Scenario 2: Combining D samples of size n=5000, from a population of size N=1000 000, distributed as the normal distribution N(100,25), $\theta=100.031$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	1/n	100,026	-4,955	0,327	0,327	100,020	100,033	0,110
	$(Var(\hat{\theta}_d))^{-1}$	100,026	-4,810	0,326	0,325	100,020	100,033	0,108
	$(CV(\hat{\theta}_d))^{-1}$	100,027	-4,663	0,326	0,326	100,020	100,033	0,109
	$(skewness(s_{dc}))^{-1}$	100,036	4,864	0,753	0,753	100,021	100,051	0,570
	$(kurtosis(k_{dc}))^{-1}$	100,026	-4,954	0,328	0,328	100,020	100,033	0,110
	$meff/Var(y)$	100,026	-5,582	0,339	0,339	100,019	100,032	0,118
	$(meff)^{-1}$	100,027	-4,374	0,324	0,324	100,020	100,033	0,107

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Like the first scenario, the weighting strategies resulted in similar MSE values except the inverse of skewness which resulted in remarkably higher MSE values. The MSE values decreased as the number of samples combined was increased. Combining samples resulted in smaller MSE values compare to using estimates from one sample.

Table 4. 33: Scenario 3: Combining D samples of size n=10000, from a population of size N=1000 000, distributed as the normal distribution N(100,25), $\theta=100.031$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		100,020	-1,093	9,764	9,762	100,001	100,039	9,653
2	1/n	100,016	-1,476	6,104	6,103	100,004	100,028	3,944
	$(Var(\hat{\theta}_a))^{-1}$	100,028	-0,364	6,167	6,165	100,015	100,040	3,816
	$(CV(\hat{\theta}_a))^{-1}$	100,027	-0,403	6,170	6,168	100,015	100,039	3,823
	$(skewness(s_{dc}))^{-1}$	100,027	-0,414	6,607	6,605	100,014	100,040	4,383
	$(kurtosis(k_{dc}))^{-1}$	100,027	-0,453	6,175	6,173	100,015	100,039	3,834
	$meff/Var(y)$	100,027	-0,435	6,180	6,178	100,015	100,039	3,838
	$(meff)^{-1}$	100,027	-0,381	6,162	6,161	100,015	100,039	3,812
3	1/n	100,017	-1,465	5,223	5,223	100,006	100,027	2,943
	$(Var(\hat{\theta}_a))^{-1}$	100,028	-0,311	5,021	5,019	100,018	100,038	2,530
	$(CV(\hat{\theta}_a))^{-1}$	100,028	-0,350	5,020	5,019	100,018	100,038	2,532
	$(skewness(s_{dc}))^{-1}$	100,019	-1,202	6,199	6,197	100,007	100,031	3,987
	$(kurtosis(k_{dc}))^{-1}$	100,027	-0,410	5,015	5,014	100,017	100,037	2,532
	$meff/Var(y)$	100,028	-0,344	5,016	5,015	100,018	100,038	2,528
	$(meff)^{-1}$	100,028	-0,337	5,018	5,016	100,018	100,038	2,529
5	1/n	100,030	-0,078	4,067	4,066	100,022	100,038	1,655
	$(Var(\hat{\theta}_a))^{-1}$	100,031	-0,049	4,087	4,086	100,023	100,039	1,671
	$(CV(\hat{\theta}_a))^{-1}$	100,031	-0,055	4,076	4,075	100,023	100,039	1,662
	$(skewness(s_{dc}))^{-1}$	100,029	-0,211	5,433	5,431	100,018	100,040	2,956
	$(kurtosis(k_{dc}))^{-1}$	100,030	-0,086	4,063	4,062	100,022	100,038	1,652
	$meff/Var(y)$	100,030	-0,103	4,011	4,010	100,022	100,038	1,610
	$(meff)^{-1}$	100,031	-0,071	4,052	4,050	100,023	100,038	1,642
10	1/n	100,028	-0,340	2,808	2,807	100,022	100,033	0,800
	$(Var(\hat{\theta}_a))^{-1}$	100,028	-0,320	2,822	2,821	100,022	100,034	0,807
	$(CV(\hat{\theta}_a))^{-1}$	100,028	-0,320	2,814	2,813	100,023	100,034	0,802
	$(skewness(s_{dc}))^{-1}$	100,032	0,107	4,409	4,408	100,024	100,041	1,945
	$(kurtosis(k_{dc}))^{-1}$	100,028	-0,349	2,810	2,809	100,022	100,033	0,802
	$meff/Var(y)$	100,028	-0,358	2,816	2,815	100,022	100,033	0,806
	$(meff)^{-1}$	100,028	-0,290	2,829	2,829	100,023	100,034	0,809
15	1/n	100,031	0,026	2,468	2,467	100,027	100,036	0,609
	$(Var(\hat{\theta}_a))^{-1}$	100,032	0,059	2,469	2,468	100,027	100,037	0,610
	$(CV(\hat{\theta}_a))^{-1}$	100,032	0,052	2,467	2,466	100,027	100,037	0,609
	$(skewness(s_{dc}))^{-1}$	100,036	0,527	4,337	4,336	100,028	100,045	1,909
	$(kurtosis(k_{dc}))^{-1}$	100,031	0,020	2,471	2,470	100,027	100,036	0,610
	$meff/Var(y)$	100,030	-0,136	2,745	2,745	100,024	100,035	0,756
	$(meff)^{-1}$	100,031	0,017	2,437	2,436	100,027	100,036	0,594

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.33 continued: Scenario 3: Combining D samples of size n=10000, from a population of size N=1000 000, distributed as the normal distribution N(100,25), $\theta=100.031$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	1/n	100,033	0,137	1,982	1,981	100,029	100,036	0,395
	$(Var(\hat{\theta}_d))^{-1}$	100,033	0,167	1,978	1,977	100,029	100,037	0,394
	$(CV(\hat{\theta}_d))^{-1}$	100,033	0,162	1,978	1,977	100,029	100,037	0,394
	$(skewness(s_{dc}))^{-1}$	100,034	0,285	4,785	4,783	100,025	100,043	2,298
	$(kurtosis(k_{dc}))^{-1}$	100,033	0,133	1,984	1,984	100,029	100,036	0,395
	$meff/Var(y)$	100,032	0,093	2,049	2,048	100,028	100,036	0,421
	$(meff)^{-1}$	100,033	0,139	2,000	2,000	100,029	100,037	0,402

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Again, all the weighting strategies resulted in similar MSE values apart from the inverse of skewness which resulted in higher MSE values. The difference in magnitude of the MSE values became more marked as more samples were combined.

Summary

All the weighting strategies performed comparably well with a normal distribution superpopulation model at all the sample sizes studied. The inverse of skewness did not work well as a weighting strategy, resulting in markedly higher MSE values. Combining samples resulted in better estimates compared estimates calculated from one sample.

4.4.2 Simulation results for a finite population generated using the skewed normal distribution, under cluster sampling

The finite population was generated by the skewed normal distribution. A location parameter of -1.15, scale parameter of 1.541 and an asymmetry parameter 3 were used to generate strongly asymmetric data with an overall mean of 100 and variance 25. The results of comparing the weighting strategies are presented below, for the design consistent weighting strategies and in Appendix B7 for all the weighting strategies.

Table 4. 34: Scenario 1: Combining D samples of size n=1000, from a population of size N=1000 000, distributed as the skewed normal distribution SN(100,25), $\theta=99.914$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-1}$
						LB	UB	
1		99,897	-1,688	3,184	3,187	99,834	99,959	1,017
2	1/n	99,905	-0,834	2,179	2,181	99,863	99,948	0,475
	$(\text{Var}(\hat{\theta}_a))^{-1}$	99,922	0,846	2,443	2,445	99,874	99,970	0,597
	$(\text{CV}(\hat{\theta}_a))^{-1}$	99,928	1,410	2,421	2,422	99,880	99,975	0,588
	$(\text{skewness}(s_{dc}))^{-1}$	99,935	2,113	2,436	2,437	99,887	99,982	0,598
	$(\text{kurtosis}(k_{dc}))^{-1}$	99,934	2,025	2,417	2,419	99,886	99,981	0,588
	$m_{eff}/\text{Var}(y)$	99,928	1,471	2,406	2,407	99,881	99,975	0,581
	$(m_{eff})^{-1}$	99,926	1,237	2,438	2,440	99,878	99,974	0,596
3	1/n	99,898	-1,526	1,959	1,961	99,860	99,937	0,386
	$(\text{Var}(\hat{\theta}_a))^{-1}$	99,919	0,541	1,964	1,965	99,881	99,958	0,386
	$(\text{CV}(\hat{\theta}_a))^{-1}$	99,928	1,435	1,964	1,965	99,889	99,966	0,388
	$(\text{skewness}(s_{dc}))^{-1}$	99,937	2,361	1,997	1,998	99,898	99,976	0,404
	$(\text{kurtosis}(k_{dc}))^{-1}$	99,937	2,305	1,989	1,990	99,898	99,976	0,401
	$m_{eff}/\text{Var}(y)$	99,929	1,537	1,955	1,957	99,891	99,967	0,385
	$(m_{eff})^{-1}$	99,925	1,156	1,974	1,975	99,886	99,964	0,391
5	1/n	99,925	1,138	1,418	1,419	99,897	99,953	0,202
	$(\text{Var}(\hat{\theta}_a))^{-1}$	99,906	-0,766	1,404	1,405	99,878	99,933	0,198
	$(\text{CV}(\hat{\theta}_a))^{-1}$	99,916	0,258	1,404	1,405	99,889	99,944	0,197
	$(\text{skewness}(s_{dc}))^{-1}$	99,927	1,342	1,393	1,394	99,900	99,954	0,196
	$(\text{kurtosis}(k_{dc}))^{-1}$	99,926	1,219	1,413	1,414	99,898	99,953	0,201
	$m_{eff}/\text{Var}(y)$	99,918	0,431	1,407	1,408	99,890	99,945	0,198
	$(m_{eff})^{-1}$	99,914	0,054	1,441	1,442	99,886	99,942	0,208
10	1/n	99,941	2,698	1,031	1,032	99,920	99,961	0,114
	$(\text{Var}(\hat{\theta}_a))^{-1}$	99,916	0,193	1,073	1,074	99,895	99,937	0,115
	$(\text{CV}(\hat{\theta}_a))^{-1}$	99,929	1,525	1,046	1,047	99,908	99,949	0,112
	$(\text{skewness}(s_{dc}))^{-1}$	99,943	2,943	1,034	1,035	99,923	99,963	0,116
	$(\text{kurtosis}(k_{dc}))^{-1}$	99,942	2,822	1,033	1,033	99,922	99,962	0,115
	$m_{eff}/\text{Var}(y)$	99,935	2,180	1,109	1,109	99,914	99,957	0,128
	$(m_{eff})^{-1}$	99,925	1,189	1,068	1,069	99,905	99,946	0,115
15	1/n	99,932	1,806	0,861	0,861	99,915	99,949	0,077
	$(\text{Var}(\hat{\theta}_a))^{-1}$	99,906	-0,715	0,883	0,883	99,889	99,924	0,078
	$(\text{CV}(\hat{\theta}_a))^{-1}$	99,920	0,635	0,866	0,867	99,903	99,937	0,075
	$(\text{skewness}(s_{dc}))^{-1}$	99,937	2,333	0,857	0,857	99,920	99,954	0,079
	$(\text{kurtosis}(k_{dc}))^{-1}$	99,934	2,018	0,854	0,854	99,917	99,951	0,077
	$m_{eff}/\text{Var}(y)$	99,923	0,964	0,871	0,872	99,906	99,940	0,077
	$(m_{eff})^{-1}$	99,915	0,132	0,873	0,874	99,898	99,932	0,076

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.34 continued: Scenario 1: Combining D samples of size n=1000, from a population of size N=1000 000, distributed as the skewed normal distribution SN(100,25), $\theta=99.914$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	1/n	99,931	1,702	0,720	0,720	99,917	99,945	0,055
	$(Var(\hat{\theta}_d))^{-1}$	99,906	-0,796	0,739	0,740	99,891	99,920	0,055
	$(CV(\hat{\theta}_d))^{-1}$	99,919	0,541	0,724	0,724	99,905	99,933	0,053
	$(skewness(s_{dc}))^{-1}$	99,937	2,349	0,723	0,724	99,923	99,951	0,058
	$(kurtosis(k_{dc}))^{-1}$	99,933	1,963	0,712	0,713	99,919	99,947	0,055
	$mef f / Var(y)$	99,924	1,014	0,729	0,729	99,909	99,938	0,054
	$(mef f)^{-1}$	99,917	0,297	0,766	0,766	99,902	99,932	0,059

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Combining samples yielded MSE values which were smaller than those resulting from using one sample. The MSE values decreased as the number of samples combined was increased. The weighting strategies resulted in similar MSE values.

Table 4. 35: Scenario 2: Combining D samples of size n=5000, from a population of size N=1000 000, distributed as the skewed normal distribution SN(100,25), $\theta=99.914$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,909	-0,491	1,493	1,495	99,879	99,938	2,232
2	1/n	99,889	-2,481	0,946	0,947	99,870	99,907	0,957
	$(Var(\hat{\theta}_a))^{-1}$	99,894	-1,946	1,024	1,025	99,874	99,914	1,086
	$(CV(\hat{\theta}_a))^{-1}$	99,895	-1,825	1,021	1,022	99,875	99,915	1,075
	$(skewness(s_{dc}))^{-1}$	99,898	-1,610	1,018	1,019	99,878	99,917	1,063
	$(kurtosis(k_{dc}))^{-1}$	99,897	-1,693	1,018	1,019	99,877	99,917	1,065
	$meff/Var(y)$	99,895	-1,818	1,020	1,021	99,875	99,915	1,074
	$(meff)^{-1}$	99,895	-1,858	1,022	1,023	99,875	99,915	1,078
3	1/n	99,891	-2,287	0,770	0,771	99,876	99,906	0,646
	$(Var(\hat{\theta}_a))^{-1}$	99,898	-1,599	0,860	0,861	99,881	99,914	0,765
	$(CV(\hat{\theta}_a))^{-1}$	99,899	-1,498	0,859	0,860	99,882	99,915	0,760
	$(skewness(s_{dc}))^{-1}$	99,900	-1,334	0,857	0,858	99,883	99,917	0,752
	$(kurtosis(k_{dc}))^{-1}$	99,900	-1,403	0,857	0,858	99,883	99,916	0,754
	$meff/Var(y)$	99,899	-1,482	0,857	0,858	99,882	99,916	0,756
	$(meff)^{-1}$	99,898	-1,513	0,860	0,861	99,882	99,915	0,762
5	1/n	99,907	-0,649	0,612	0,613	99,895	99,919	0,379
	$(Var(\hat{\theta}_a))^{-1}$	99,904	-0,985	0,619	0,620	99,892	99,916	0,393
	$(CV(\hat{\theta}_a))^{-1}$	99,906	-0,800	0,615	0,616	99,894	99,918	0,385
	$(skewness(s_{dc}))^{-1}$	99,907	-0,612	0,612	0,612	99,896	99,919	0,378
	$(kurtosis(k_{dc}))^{-1}$	99,907	-0,639	0,611	0,611	99,895	99,919	0,377
	$meff/Var(y)$	99,905	-0,856	0,623	0,623	99,893	99,917	0,395
	$(meff)^{-1}$	99,905	-0,894	0,621	0,621	99,893	99,917	0,393
10	1/n	99,907	-0,663	0,420	0,420	99,899	99,915	0,180
	$(Var(\hat{\theta}_a))^{-1}$	99,903	-1,069	0,420	0,420	99,895	99,911	0,187
	$(CV(\hat{\theta}_a))^{-1}$	99,905	-0,846	0,419	0,420	99,897	99,913	0,183
	$(skewness(s_{dc}))^{-1}$	99,908	-0,587	0,419	0,419	99,900	99,916	0,179
	$(kurtosis(k_{dc}))^{-1}$	99,907	-0,648	0,418	0,419	99,899	99,915	0,179
	$meff/Var(y)$	99,905	-0,878	0,426	0,427	99,896	99,913	0,189
	$(meff)^{-1}$	99,903	-1,039	0,432	0,432	99,895	99,912	0,197
15	1/n	99,908	-0,543	0,355	0,355	99,901	99,915	0,129
	$(Var(\hat{\theta}_a))^{-1}$	99,904	-1,000	0,354	0,355	99,897	99,911	0,135
	$(CV(\hat{\theta}_a))^{-1}$	99,906	-0,751	0,354	0,355	99,899	99,913	0,131
	$(skewness(s_{dc}))^{-1}$	99,909	-0,455	0,353	0,354	99,902	99,916	0,127
	$(kurtosis(k_{dc}))^{-1}$	99,908	-0,517	0,353	0,354	99,902	99,915	0,127
	$meff/Var(y)$	99,908	-0,598	0,372	0,372	99,900	99,915	0,142
	$(meff)^{-1}$	99,906	-0,803	0,361	0,361	99,899	99,913	0,137

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.35 continued: Scenario 2: Combining D samples of size n=5000, from a population of size N=1000 000, distributed as the skewed normal distribution SN(100,25), $\theta=99.914$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	1/n	99,909	-0,479	0,300	0,300	99,903	99,915	0,092
	$(Var(\hat{\theta}_d))^{-1}$	99,904	-0,945	0,300	0,300	99,898	99,910	0,099
	$(CV(\hat{\theta}_d))^{-1}$	99,907	-0,691	0,300	0,300	99,901	99,913	0,094
	$(skewness(s_{dc}))^{-1}$	99,910	-0,360	0,297	0,298	99,904	99,916	0,090
	$(kurtosis(k_{dc}))^{-1}$	99,909	-0,437	0,299	0,299	99,903	99,915	0,091
	$meff/Var(y)$	99,908	-0,592	0,302	0,302	99,902	99,914	0,095
	$(meff)^{-1}$	99,906	-0,762	0,302	0,302	99,900	99,912	0,097

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

The weighting strategies performed comparably, resulting in similar MSE values for the same numbers of samples combined. Combined estimates resulted in smaller MSE values compared to the estimate calculated from one sample. Increasing the number of samples combined tended to decrease the MSE values.

Table 4. 36: Scenario 3: Combining D samples of size n=10000, from a population of size N=1000 000, distributed as the skewed normal distribution SN(100,25), $\theta=99.914$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,912	-1,717	1,011	1,012	99,892	99,932	1,022
2	1/n	99,910	-3,600	0,668	0,668	99,897	99,923	0,447
	$(Var(\hat{\theta}_a))^{-1}$	99,906	-8,014	0,722	0,723	99,891	99,920	0,528
	$(CV(\hat{\theta}_a))^{-1}$	99,906	-7,243	0,722	0,722	99,892	99,921	0,526
	$(skewness(s_{dc}))^{-1}$	99,908	-5,840	0,727	0,728	99,894	99,922	0,532
	$(kurtosis(k_{dc}))^{-1}$	99,907	-6,441	0,723	0,723	99,893	99,921	0,526
	$meff/Var(y)$	99,907	-6,931	0,722	0,722	99,893	99,921	0,526
	$(meff)^{-1}$	99,906	-7,661	0,722	0,722	99,892	99,920	0,526
3	1/n	99,914	0,257	0,552	0,552	99,903	99,925	0,305
	$(Var(\hat{\theta}_a))^{-1}$	99,909	-4,652	0,584	0,584	99,898	99,920	0,343
	$(CV(\hat{\theta}_a))^{-1}$	99,910	-3,422	0,583	0,583	99,899	99,922	0,341
	$(skewness(s_{dc}))^{-1}$	99,912	-1,579	0,584	0,584	99,901	99,923	0,341
	$(kurtosis(k_{dc}))^{-1}$	99,912	-2,115	0,583	0,583	99,900	99,923	0,340
	$meff/Var(y)$	99,910	-3,194	0,582	0,583	99,899	99,922	0,340
	$(meff)^{-1}$	99,910	-3,931	0,584	0,584	99,898	99,921	0,342
5	1/n	99,914	0,630	0,446	0,446	99,906	99,923	0,199
	$(Var(\hat{\theta}_a))^{-1}$	99,912	-1,857	0,444	0,444	99,903	99,920	0,197
	$(CV(\hat{\theta}_a))^{-1}$	99,913	-0,521	0,445	0,445	99,904	99,922	0,198
	$(skewness(s_{dc}))^{-1}$	99,915	1,352	0,451	0,451	99,906	99,924	0,203
	$(kurtosis(k_{dc}))^{-1}$	99,914	0,879	0,450	0,450	99,906	99,923	0,202
	$meff/Var(y)$	99,914	0,326	0,446	0,446	99,905	99,923	0,199
	$(meff)^{-1}$	99,912	-1,367	0,446	0,447	99,904	99,921	0,199
10	1/n	99,911	-2,448	0,306	0,307	99,905	99,917	0,095
	$(Var(\hat{\theta}_a))^{-1}$	99,909	-4,869	0,311	0,311	99,903	99,915	0,099
	$(CV(\hat{\theta}_a))^{-1}$	99,910	-3,555	0,308	0,309	99,904	99,916	0,096
	$(skewness(s_{dc}))^{-1}$	99,912	-1,623	0,307	0,307	99,906	99,918	0,094
	$(kurtosis(k_{dc}))^{-1}$	99,911	-2,180	0,307	0,307	99,905	99,917	0,095
	$meff/Var(y)$	99,910	-3,640	0,304	0,304	99,904	99,916	0,094
	$(meff)^{-1}$	99,909	-4,582	0,307	0,307	99,903	99,915	0,096
15	1/n	99,911	-2,283	0,236	0,237	99,907	99,916	0,056
	$(Var(\hat{\theta}_a))^{-1}$	99,909	-4,670	0,239	0,239	99,904	99,914	0,059
	$(CV(\hat{\theta}_a))^{-1}$	99,910	-3,374	0,237	0,238	99,906	99,915	0,058
	$(skewness(s_{dc}))^{-1}$	99,912	-1,615	0,238	0,238	99,907	99,917	0,057
	$(kurtosis(k_{dc}))^{-1}$	99,912	-2,065	0,237	0,237	99,907	99,916	0,057
	$meff/Var(y)$	99,910	-3,742	0,246	0,246	99,905	99,915	0,062
	$(meff)^{-1}$	99,910	-3,880	0,238	0,238	99,905	99,914	0,058

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.36 continued: Scenario 3: Combining D samples of size n=10000, from a population of size N=1000 000, distributed as the skewed normal distribution SN(100,25), $\theta=99.914$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	1/n	99,911	-2,840	0,217	0,217	99,907	99,915	0,048
	$(Var(\hat{\theta}_d))^{-1}$	99,908	-5,184	0,219	0,219	99,904	99,913	0,051
	$(CV(\hat{\theta}_d))^{-1}$	99,910	-3,910	0,218	0,218	99,905	99,914	0,049
	$(skewness(s_{dc}))^{-1}$	99,911	-2,205	0,218	0,218	99,907	99,916	0,048
	$(kurtosis(k_{dc}))^{-1}$	99,911	-2,643	0,217	0,218	99,907	99,915	0,048
	$meff/Var(y)$	99,910	-3,478	0,216	0,216	99,906	99,914	0,048
	$(meff)^{-1}$	99,909	-4,760	0,226	0,226	99,904	99,913	0,053

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

The weighting strategies resulted in comparable MSE values which decreased as the number of samples combined was increased. The MSE value resulting from the one sample estimate was markedly higher than the MSE values resulting from the combined estimates.

Summary

Combining samples resulted in better estimates compared to using estimates derived from one sample. The weighting strategies performed similarly, with comparable MSE values. At sample sizes 5000 and 10000, the weighting strategies mostly underestimated the population parameter value as shown by the negative relative bias and 95% confidence intervals.

4.4.3 Simulation results for the t- distribution with three degrees of freedom, under cluster sampling

The finite population, from which multiple surveys were sampled, was generated by a t- distribution with three degrees of freedom in each cluster for this simulation. The finite population was simulated in such a way that the overall mean would be 100 and the variance 25. The results of comparing the design consistent weighting strategies for combining the surveys under this distribution are presented below and in Appendix C2 for all the weighting strategies studied.

Table 4. 37: Scenario 1: Combining D samples of size n=1000, from a population of size N=1000 000, distributed as the t-distribution with three degrees of freedom t(3)(100,25), $\theta=100.010$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,981	-2,931	3,051	3,051	99,921	100,041	9,393
2	1/n	100,015	0,461	1,881	1,880	99,978	100,052	3,539
	$(Var(\hat{\theta}_a))^{-1}$	100,006	-0,389	2,199	2,199	99,963	100,049	4,839
	$(CV(\hat{\theta}_a))^{-1}$	100,005	-0,480	2,176	2,176	99,963	100,048	4,739
	$(skewness(s_{dc}))^{-1}$	100,002	-0,791	2,538	2,538	99,952	100,052	6,448
	$(kurtosis(k_{dc}))^{-1}$	99,999	-1,145	2,367	2,367	99,952	100,045	5,614
	$meff/Var(y)$	100,003	-0,727	2,175	2,175	99,960	100,046	4,735
	$(meff)^{-1}$	100,007	-0,358	2,200	2,199	99,963	100,050	4,839
3	1/n	100,019	0,908	1,731	1,731	99,985	100,053	3,005
	$(Var(\hat{\theta}_a))^{-1}$	100,009	-0,069	1,968	1,968	99,971	100,048	3,875
	$(CV(\hat{\theta}_a))^{-1}$	100,009	-0,081	1,958	1,958	99,971	100,048	3,834
	$(skewness(s_{dc}))^{-1}$	100,010	-0,058	2,394	2,394	99,963	100,056	5,731
	$(kurtosis(k_{dc}))^{-1}$	100,001	-0,923	2,201	2,201	99,958	100,044	4,852
	$meff/Var(y)$	100,005	-0,531	2,005	2,005	99,966	100,044	4,023
	$(meff)^{-1}$	100,009	-0,153	1,995	1,995	99,970	100,048	3,981
5	1/n	100,007	-0,338	1,469	1,469	99,978	100,036	2,159
	$(Var(\hat{\theta}_a))^{-1}$	100,005	-0,480	1,477	1,477	99,976	100,034	2,185
	$(CV(\hat{\theta}_a))^{-1}$	100,006	-0,376	1,459	1,458	99,978	100,035	2,129
	$(skewness(s_{dc}))^{-1}$	99,987	-2,302	2,000	2,000	99,948	100,026	4,052
	$(kurtosis(k_{dc}))^{-1}$	100,000	-0,968	1,576	1,576	99,970	100,031	2,492
	$meff/Var(y)$	100,005	-0,528	1,443	1,443	99,977	100,033	2,085
	$(meff)^{-1}$	100,005	-0,541	1,516	1,516	99,975	100,034	2,301
10	1/n	100,003	-0,750	0,903	0,903	99,985	100,020	0,821
	$(Var(\hat{\theta}_a))^{-1}$	99,994	-1,590	0,930	0,930	99,976	100,012	0,891
	$(CV(\hat{\theta}_a))^{-1}$	99,999	-1,143	0,901	0,901	99,981	100,016	0,825
	$(skewness(s_{dc}))^{-1}$	100,004	-0,619	1,764	1,764	99,969	100,039	3,116
	$(kurtosis(k_{dc}))^{-1}$	100,003	-0,683	1,033	1,033	99,983	100,024	1,073
	$meff/Var(y)$	100,002	-0,813	0,917	0,917	99,984	100,020	0,848
	$(meff)^{-1}$	99,995	-1,548	0,976	0,976	99,976	100,014	0,977
15	1/n	100,010	0,002	0,749	0,749	99,995	100,025	0,561
	$(Var(\hat{\theta}_a))^{-1}$	99,996	-1,426	0,729	0,729	99,982	100,010	0,552
	$(CV(\hat{\theta}_a))^{-1}$	100,003	-0,707	0,728	0,728	99,989	100,017	0,535
	$(skewness(s_{dc}))^{-1}$	100,005	-0,466	1,786	1,786	99,970	100,040	3,191
	$(kurtosis(k_{dc}))^{-1}$	100,012	0,137	0,884	0,884	99,994	100,029	0,782
	$meff/Var(y)$	100,004	-0,593	0,758	0,758	99,989	100,019	0,578
	$(meff)^{-1}$	99,996	-1,430	0,773	0,774	99,981	100,011	0,619

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.37 continued: Scenario 1: Combining D samples of size n=1000, from a population of size N=1000 000, distributed as the t-distribution with three degrees of freedom t(3)(100,25), $\theta=100.010$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	1/n	100,014	0,362	0,746	0,746	99,999	100,028	0,557
	$(Var(\hat{\theta}_d))^{-1}$	100,000	-1,027	0,733	0,733	99,986	100,014	0,548
	$(CV(\hat{\theta}_d))^{-1}$	100,007	-0,319	0,732	0,731	99,993	100,021	0,536
	$(skewness(s_{dc}))^{-1}$	100,002	-0,787	1,695	1,695	99,969	100,036	2,878
	$(kurtosis(k_{dc}))^{-1}$	100,013	0,321	0,843	0,843	99,997	100,030	0,711
	$meff/Var(y)$	100,007	-0,351	0,789	0,789	99,991	100,022	0,624
	$(meff)^{-1}$	99,998	-1,213	0,864	0,864	99,981	100,015	0,762

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Combining samples resulted in lower MSE values compared to estimates obtained from one sample. Weighting by the inverse of the sample size resulted in the lowest MSE value for 2, 3 or 10 samples were combined. The inverse of skewness resulted in markedly higher MSE values compared to the other weighting strategies.

Table 4. 38: Scenario 2: Combining D samples of size n=5000, from a population of size N=1000 000, distributed as the t-distribution with three degrees of freedom t(3)(100,25), $\theta=100.010$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,994	-1,627	1,527	1,527	99,964	100,024	2,357
2	1/n	100,025	1,532	0,945	0,944	100,007	100,044	0,916
	$(\text{Var}(\hat{\theta}_a))^{-1}$	100,010	0,010	1,062	1,062	99,989	100,031	1,128
	$(\text{CV}(\hat{\theta}_a))^{-1}$	100,012	0,161	1,060	1,059	99,991	100,033	1,123
	$(\text{skewness}(s_{dc}))^{-1}$	100,007	-0,346	1,354	1,354	99,980	100,033	1,834
	$(\text{kurtosis}(k_{dc}))^{-1}$	100,014	0,421	1,250	1,250	99,990	100,039	1,565
	$m_{eff}/\text{Var}(y)$	100,012	0,195	1,069	1,069	99,991	100,033	1,143
	$(m_{eff})^{-1}$	100,011	0,110	1,057	1,056	99,991	100,032	1,117
3	1/n	100,021	1,090	0,715	0,714	100,007	100,035	0,522
	$(\text{Var}(\hat{\theta}_a))^{-1}$	100,008	-0,265	0,801	0,801	99,992	100,023	0,642
	$(\text{CV}(\hat{\theta}_a))^{-1}$	100,009	-0,069	0,797	0,797	99,994	100,025	0,636
	$(\text{skewness}(s_{dc}))^{-1}$	100,002	-0,854	1,129	1,129	99,979	100,024	1,281
	$(\text{kurtosis}(k_{dc}))^{-1}$	100,009	-0,138	0,977	0,977	99,990	100,028	0,955
	$m_{eff}/\text{Var}(y)$	100,011	0,042	0,814	0,814	99,995	100,027	0,663
	$(m_{eff})^{-1}$	100,009	-0,086	0,791	0,791	99,994	100,025	0,625
5	1/n	100,007	-0,364	0,670	0,670	99,993	100,020	0,450
	$(\text{Var}(\hat{\theta}_a))^{-1}$	100,004	-0,644	0,669	0,669	99,991	100,017	0,452
	$(\text{CV}(\hat{\theta}_a))^{-1}$	100,005	-0,489	0,668	0,668	99,992	100,018	0,449
	$(\text{skewness}(s_{dc}))^{-1}$	99,998	-1,254	0,990	0,990	99,978	100,017	0,997
	$(\text{kurtosis}(k_{dc}))^{-1}$	100,005	-0,554	0,770	0,770	99,990	100,020	0,596
	$m_{eff}/\text{Var}(y)$	100,005	-0,525	0,680	0,680	99,992	100,018	0,465
	$(m_{eff})^{-1}$	100,005	-0,559	0,669	0,669	99,991	100,018	0,451
10	1/n	100,006	-0,398	0,457	0,457	99,997	100,015	0,211
	$(\text{Var}(\hat{\theta}_a))^{-1}$	100,003	-0,743	0,436	0,436	99,994	100,011	0,196
	$(\text{CV}(\hat{\theta}_a))^{-1}$	100,005	-0,552	0,445	0,445	99,996	100,013	0,201
	$(\text{skewness}(s_{dc}))^{-1}$	99,997	-1,280	0,907	0,907	99,980	100,015	0,840
	$(\text{kurtosis}(k_{dc}))^{-1}$	100,008	-0,241	0,536	0,536	99,997	100,018	0,287
	$m_{eff}/\text{Var}(y)$	100,005	-0,482	0,459	0,459	99,996	100,014	0,213
	$(m_{eff})^{-1}$	100,003	-0,742	0,452	0,452	99,994	100,012	0,210
15	1/n	100,006	-0,371	0,370	0,370	99,999	100,014	0,139
	$(\text{Var}(\hat{\theta}_a))^{-1}$	100,002	-0,790	0,353	0,353	99,995	100,009	0,131
	$(\text{CV}(\hat{\theta}_a))^{-1}$	100,004	-0,566	0,360	0,360	99,997	100,012	0,133
	$(\text{skewness}(s_{dc}))^{-1}$	99,999	-1,080	0,760	0,760	99,984	100,014	0,590
	$(\text{kurtosis}(k_{dc}))^{-1}$	100,007	-0,338	0,404	0,404	99,999	100,015	0,165
	$m_{eff}/\text{Var}(y)$	100,006	-0,461	0,365	0,365	99,998	100,013	0,135
	$(m_{eff})^{-1}$	100,003	-0,723	0,362	0,362	99,996	100,010	0,136

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.38 continued: Scenario 2: Combining D samples of size n=5000, from a population of size N=1000 000, distributed as the t- distribution with three degrees of freedom t(3)(100,25), $\theta=100.010$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	1/n	100,008	-0,199	0,327	0,327	100,002	100,015	0,107
	$(Var(\hat{\theta}_d))^{-1}$	100,004	-0,635	0,319	0,319	99,998	100,010	0,106
	$(CV(\hat{\theta}_d))^{-1}$	100,006	-0,404	0,322	0,322	100,000	100,012	0,105
	$(skewness(s_{dc}))^{-1}$	100,004	-0,574	0,752	0,752	99,990	100,019	0,568
	$(kurtosis(k_{dc}))^{-1}$	100,007	-0,276	0,368	0,367	100,000	100,015	0,136
	$mef f / Var(y)$	100,007	-0,343	0,332	0,332	100,000	100,013	0,111
	$(mef f)^{-1}$	100,003	-0,672	0,340	0,340	99,997	100,010	0,120

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Similarly, to the scenario where a sample size of 1000 was used, combining samples resulted in lower MSE values compared to using estimates obtained from one sample. Weighting the samples by the inverses of skewness and kurtosis resulted in remarkably higher MSE values compared to the other weighting strategies. MSE values resulting from using the inverse of kurtosis as a weighting strategy became closer to the other weighting strategies as the number of samples combined was increased.

Table 4. 39: Scenario 3: Combining D samples of size n=10000, from a population of size N=1000 000, distributed as the t-distribution with three degrees of freedom t(3)(100,25), $\theta=100.010$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		99,994	-1,574	9,449	9,449	99,976	100,013	9,176
2	1/n	99,999	-1,082	6,461	6,461	99,987	100,012	4,291
	$(Var(\hat{\theta}_a))^{-1}$	100,003	-0,690	6,410	6,410	99,991	100,016	4,156
	$(CV(\hat{\theta}_a))^{-1}$	100,004	-0,617	6,410	6,410	99,991	100,017	4,147
	$(skewness(s_{dc}))^{-1}$	100,008	-0,208	8,589	8,588	99,991	100,025	7,382
	$(kurtosis(k_{dc}))^{-1}$	100,006	-0,438	7,682	7,682	99,991	100,021	5,921
	$meff/Var(y)$	100,005	-0,543	6,491	6,491	99,992	100,017	4,243
	$(meff)^{-1}$	100,003	-0,704	6,379	6,379	99,991	100,016	4,119
3	1/n	99,999	-1,090	5,228	5,228	99,989	100,010	2,852
	$(Var(\hat{\theta}_a))^{-1}$	100,003	-0,694	5,267	5,266	99,993	100,014	2,822
	$(CV(\hat{\theta}_a))^{-1}$	100,004	-0,628	5,255	5,255	99,994	100,014	2,801
	$(skewness(s_{dc}))^{-1}$	100,008	-0,252	7,140	7,139	99,994	100,022	5,104
	$(kurtosis(k_{dc}))^{-1}$	100,007	-0,323	6,070	6,070	99,995	100,019	3,695
	$meff/Var(y)$	100,005	-0,561	5,263	5,263	99,994	100,015	2,802
	$(meff)^{-1}$	100,003	-0,734	5,274	5,274	99,992	100,013	2,836
5	1/n	100,002	-0,850	4,491	4,491	99,993	100,010	2,089
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-0,964	4,489	4,489	99,992	100,009	2,108
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-0,899	4,487	4,487	99,992	100,010	2,094
	$(skewness(s_{dc}))^{-1}$	100,010	0,021	6,777	6,777	99,997	100,024	4,593
	$(kurtosis(k_{dc}))^{-1}$	100,004	-0,626	5,214	5,214	99,994	100,014	2,758
	$meff/Var(y)$	100,002	-0,840	4,545	4,545	99,993	100,011	2,137
	$(meff)^{-1}$	100,001	-0,922	4,512	4,512	99,992	100,010	2,121
10	1/n	100,004	-0,642	3,162	3,162	99,998	100,010	1,041
	$(Var(\hat{\theta}_a))^{-1}$	100,003	-0,756	3,205	3,205	99,996	100,009	1,084
	$(CV(\hat{\theta}_a))^{-1}$	100,003	-0,691	3,179	3,179	99,997	100,009	1,058
	$(skewness(s_{dc}))^{-1}$	100,006	-0,394	5,886	5,886	99,995	100,018	3,480
	$(kurtosis(k_{dc}))^{-1}$	100,006	-0,406	3,931	3,931	99,998	100,014	1,562
	$meff/Var(y)$	100,004	-0,604	3,219	3,219	99,998	100,010	1,073
	$(meff)^{-1}$	100,003	-0,735	3,140	3,140	99,997	100,009	1,040
15	1/n	100,007	-0,279	2,501	2,501	100,002	100,012	0,633
	$(Var(\hat{\theta}_a))^{-1}$	100,006	-0,420	2,471	2,471	100,001	100,011	0,628
	$(CV(\hat{\theta}_a))^{-1}$	100,007	-0,279	2,501	2,501	100,002	100,012	0,633
	$(skewness(s_{dc}))^{-1}$	100,007	-0,341	2,481	2,481	100,002	100,012	0,627
	$(kurtosis(k_{dc}))^{-1}$	100,005	-0,534	5,437	5,436	99,994	100,015	2,984
	$meff/Var(y)$	100,009	-0,108	3,058	3,058	100,003	100,015	0,936
	$(meff)^{-1}$	100,007	-0,324	2,552	2,552	100,002	100,012	0,662

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.39 continued: Scenario 3: Combining D samples of size n=10000, from a population of size N=1000 000, distributed as the t- distribution with three degrees of freedom t(3)(100,25), $\theta=100.010$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	1/n	100,005	-0,480	2,566	2,566	100,000	100,010	0,681
	$(Var(\hat{\theta}_d))^{-1}$	100,007	-0,313	2,291	2,291	100,003	100,012	0,535
	$(CV(\hat{\theta}_d))^{-1}$	100,005	-0,470	2,259	2,259	100,001	100,010	0,532
	$(skewness(s_{dc}))^{-1}$	100,006	-0,382	2,271	2,270	100,002	100,011	0,530
	$(kurtosis(k_{dc}))^{-1}$	100,003	-0,711	4,919	4,919	99,993	100,013	2,470
	$mef f / Var(y)$	100,009	-0,094	2,594	2,594	100,004	100,014	0,674
	$(mef f)^{-1}$	100,007	-0,327	2,284	2,284	100,002	100,011	0,532

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

The weighting methods more or less performed similarly except the inverses of skewness and kurtosis which resulted in higher MSE values. The MSE values resulting from combining samples were much lower compared to MSE values resulting from using estimates calculated from one sample, regardless of number of samples combined.

Summary

The weighting strategies generally performed similarly with the exception of the inverses of skewness and kurtosis. Combining estimates from different samples yielded better estimates compared to using estimates from one sample.

4.4.4 Simulation results for a finite population generated using the skewed t- distribution with three degrees of freedom, under cluster sampling

Weighting strategies were compared where a finite population was generated using a skewed t-distribution with three degrees of freedom in each cluster. A location parameter of -0.7891, a scale parameter of 0.746 and an asymmetry parameter of 3 were used to generate the skewed distribution. Appendix D2 presents comprehensive results of this comparison and the tables below present the results for design consistent weighting strategies.

Table 4. 40: Scenario 1: Combining D samples of size n=1000, from a population of size N=1000 000, distributed as the skewed t-distribution with three degrees of freedom at(3)(100,25), $\theta=99.940$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,936	-0,595	2,697	2,698	99,883	99,989	7,276
2	1/n	99,918	-2,374	1,835	1,836	99,882	99,954	3,423
	$(Var(\hat{\theta}_a))^{-1}$	99,899	-4,257	1,978	1,980	99,861	99,938	4,095
	$(CV(\hat{\theta}_a))^{-1}$	99,921	-2,088	1,947	1,948	99,883	99,959	3,834
	$(skewness(s_{dc}))^{-1}$	99,924	-1,815	2,010	2,012	99,884	99,963	4,073
	$(kurtosis(k_{dc}))^{-1}$	99,919	-2,329	2,080	2,082	99,878	99,959	4,382
	$meff/Var(y)$	99,916	-2,627	1,973	1,974	99,877	99,954	3,961
	$(meff)^{-1}$	99,927	-1,510	1,990	1,992	99,888	99,966	3,983
3	1/n	99,927	-1,529	1,523	1,524	99,897	99,956	2,343
	$(Var(\hat{\theta}_a))^{-1}$	99,888	-5,384	1,567	1,568	99,857	99,919	2,743
	$(CV(\hat{\theta}_a))^{-1}$	99,918	-2,419	1,571	1,572	99,887	99,948	2,525
	$(skewness(s_{dc}))^{-1}$	99,921	-2,106	1,692	1,694	99,888	99,954	2,909
	$(kurtosis(k_{dc}))^{-1}$	99,913	-2,904	1,809	1,811	99,877	99,948	3,357
	$meff/Var(y)$	99,914	-2,743	1,595	1,596	99,883	99,946	2,618
	$(meff)^{-1}$	99,924	-1,804	1,631	1,632	99,892	99,956	2,693
5	1/n	99,935	-0,688	1,316	1,317	99,909	99,961	1,736
	$(Var(\hat{\theta}_a))^{-1}$	99,869	-7,298	1,295	1,297	99,844	99,894	2,208
	$(CV(\hat{\theta}_a))^{-1}$	99,901	-4,059	1,264	1,265	99,877	99,926	1,762
	$(skewness(s_{dc}))^{-1}$	99,908	-3,339	1,339	1,340	99,882	99,935	1,903
	$(kurtosis(k_{dc}))^{-1}$	99,904	-3,781	1,400	1,402	99,877	99,932	2,104
	$meff/Var(y)$	99,901	-4,038	1,313	1,314	99,876	99,927	1,886
	$(meff)^{-1}$	99,908	-3,404	1,329	1,330	99,882	99,934	1,881
10	1/n	99,927	-1,439	0,922	0,922	99,909	99,946	0,870
	$(Var(\hat{\theta}_a))^{-1}$	99,862	-7,945	0,989	0,991	99,843	99,882	1,608
	$(CV(\hat{\theta}_a))^{-1}$	99,895	-4,727	0,920	0,921	99,877	99,913	1,070
	$(skewness(s_{dc}))^{-1}$	99,908	-3,339	0,982	0,983	99,889	99,928	1,076
	$(kurtosis(k_{dc}))^{-1}$	99,908	-3,395	1,062	1,063	99,887	99,929	1,244
	$meff/Var(y)$	99,894	-4,816	0,925	0,926	99,876	99,912	1,086
	$(meff)^{-1}$	99,896	-4,559	0,965	0,966	99,877	99,915	1,138
15	1/n	99,932	-1,023	0,781	0,782	99,916	99,947	0,621
	$(Var(\hat{\theta}_a))^{-1}$	99,865	-7,727	0,828	0,829	99,848	99,881	1,281
	$(CV(\hat{\theta}_a))^{-1}$	99,898	-4,416	0,777	0,778	99,882	99,913	0,799
	$(skewness(s_{dc}))^{-1}$	99,910	-3,208	0,787	0,788	99,894	99,925	0,722
	$(kurtosis(k_{dc}))^{-1}$	99,910	-3,229	0,842	0,843	99,893	99,926	0,813
	$meff/Var(y)$	99,899	-4,305	0,801	0,802	99,883	99,915	0,827
	$(meff)^{-1}$	99,901	-4,098	0,817	0,818	99,885	99,917	0,835

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.40 continued: Scenario 1: Combining D samples of size n=1000, from a population of size N=1000 000, distributed as the skewed t-distribution with three degrees of freedom at(3)(100,25), $\theta=99.940$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	1/n	99,933	-0,910	0,681	0,682	99,919	99,946	0,473
	$(Var(\hat{\theta}_d))^{-1}$	99,867	-7,459	0,684	0,685	99,854	99,881	1,023
	$(CV(\hat{\theta}_d))^{-1}$	99,899	-4,249	0,664	0,665	99,886	99,912	0,621
	$(skewness(s_{dc}))^{-1}$	99,912	-3,019	0,694	0,695	99,898	99,925	0,573
	$(kurtosis(k_{dc}))^{-1}$	99,912	-2,974	0,748	0,748	99,897	99,927	0,648
	$meff/Var(y)$	99,899	-4,329	0,666	0,666	99,886	99,912	0,630
	$(meff)^{-1}$	99,905	-3,710	0,762	0,762	99,890	99,920	0,717

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Estimates from one sample resulted in higher MSE values compared to estimates obtained by combining estimates from different samples. Weighting strategies generally performed similarly. However, the inverse of sample size outperformed the other weighting strategies as the number of samples combined was increased.

Table 4. 41: Scenario 2: Combining D samples of size n=5000, from a population of size N=1000 000, distributed as the skewed t-distribution with three degrees of freedom at(3) (100,25), $\theta=99.940$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,922	-1,777	1,494	1,495	99,893	99,951	2,262
2	1/n	99,931	-0,853	0,992	0,992	99,912	99,951	0,991
	$(Var(\hat{\theta}_a))^{-1}$	99,918	-2,164	1,041	1,042	99,898	99,938	1,131
	$(CV(\hat{\theta}_a))^{-1}$	99,927	-1,261	1,027	1,028	99,907	99,947	1,071
	$(skewness(s_{dc}))^{-1}$	99,926	-1,349	1,069	1,070	99,905	99,947	1,162
	$(kurtosis(k_{dc}))^{-1}$	99,923	-1,673	1,138	1,139	99,901	99,945	1,323
	$meff/Var(y)$	99,927	-1,235	1,032	1,033	99,907	99,948	1,081
	$(meff)^{-1}$	99,927	-1,248	1,028	1,028	99,907	99,947	1,072
3	1/n	99,930	-0,926	0,824	0,825	99,914	99,947	0,688
	$(Var(\hat{\theta}_a))^{-1}$	99,924	-1,540	0,824	0,824	99,908	99,940	0,702
	$(CV(\hat{\theta}_a))^{-1}$	99,932	-0,755	0,827	0,827	99,916	99,948	0,689
	$(skewness(s_{dc}))^{-1}$	99,933	-0,686	0,886	0,886	99,915	99,950	0,789
	$(kurtosis(k_{dc}))^{-1}$	99,931	-0,877	0,984	0,985	99,912	99,950	0,977
	$meff/Var(y)$	99,932	-0,756	0,837	0,838	99,916	99,949	0,707
	$(meff)^{-1}$	99,933	-0,701	0,831	0,831	99,916	99,949	0,695
5	1/n	99,931	-0,892	0,716	0,717	99,917	99,945	0,521
	$(Var(\hat{\theta}_a))^{-1}$	99,910	-2,972	0,648	0,649	99,897	99,923	0,508
	$(CV(\hat{\theta}_a))^{-1}$	99,920	-2,011	0,671	0,672	99,906	99,933	0,491
	$(skewness(s_{dc}))^{-1}$	99,921	-1,830	0,754	0,755	99,907	99,936	0,603
	$(kurtosis(k_{dc}))^{-1}$	99,918	-2,130	0,859	0,860	99,902	99,935	0,783
	$meff/Var(y)$	99,920	-1,957	0,700	0,700	99,906	99,934	0,528
	$(meff)^{-1}$	99,918	-2,121	0,658	0,658	99,906	99,931	0,477
10	1/n	99,936	-0,382	0,474	0,474	99,927	99,945	0,225
	$(Var(\hat{\theta}_a))^{-1}$	99,910	-2,971	0,448	0,448	99,901	99,919	0,289
	$(CV(\hat{\theta}_a))^{-1}$	99,921	-1,819	0,451	0,451	99,913	99,930	0,237
	$(skewness(s_{dc}))^{-1}$	99,927	-1,272	0,484	0,485	99,917	99,936	0,251
	$(kurtosis(k_{dc}))^{-1}$	99,925	-1,453	0,542	0,542	99,915	99,936	0,314
	$meff/Var(y)$	99,923	-1,694	0,444	0,444	99,914	99,931	0,225
	$(meff)^{-1}$	99,918	-2,156	0,439	0,440	99,910	99,927	0,239
15	1/n	99,937	-0,274	0,373	0,374	99,930	99,944	0,140
	$(Var(\hat{\theta}_a))^{-1}$	99,911	-2,843	0,348	0,348	99,904	99,918	0,201
	$(CV(\hat{\theta}_a))^{-1}$	99,923	-1,699	0,350	0,351	99,916	99,930	0,151
	$(skewness(s_{dc}))^{-1}$	99,928	-1,175	0,394	0,394	99,920	99,936	0,169
	$(kurtosis(k_{dc}))^{-1}$	99,926	-1,321	0,452	0,453	99,918	99,935	0,222
	$meff/Var(y)$	99,926	-1,348	0,392	0,393	99,919	99,934	0,172
	$(meff)^{-1}$	99,922	-1,793	0,355	0,355	99,915	99,929	0,158

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.41 continued: Scenario 2: Combining D samples of size n=5000, from a population of size N=1000 000, distributed as the skewed t-distribution with three degrees of freedom at(3) (100,25), $\theta=99.940$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	1/n	99,936	-0,355	0,303	0,304	99,930	99,942	0,093
	$(Var(\hat{\theta}_d))^{-1}$	99,909	-3,031	0,305	0,306	99,903	99,915	0,185
	$(CV(\hat{\theta}_d))^{-1}$	99,921	-1,855	0,299	0,299	99,915	99,927	0,124
	$(skewness(s_{dc}))^{-1}$	99,926	-1,395	0,345	0,345	99,919	99,932	0,138
	$(kurtosis(k_{dc}))^{-1}$	99,923	-1,627	0,399	0,400	99,916	99,931	0,186
	$meff/Var(y)$	99,923	-1,643	0,307	0,308	99,917	99,929	0,121
	$(meff)^{-1}$	99,920	-1,980	0,297	0,298	99,914	99,926	0,128

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

In general, the weighting strategies resulted in similar MSE values. However, the inverse of sample size resulted in the least MSE values except when five samples were combined. The MSE values decreased as the number of samples combined increased.

Table 4. 42: Scenario 3: Combining D samples of size n=10000, from a population of size N=1000 000, distributed as the skewed t-distribution with three degrees of freedom at(3) (100,25), $\theta=99.940$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,920	-1,985	1,228	1,229	99,896	99,944	1,547
2	1/n	99,934	-0,542	0,728	0,729	99,920	99,949	0,533
	$(Var(\hat{\theta}_a))^{-1}$	99,921	-1,823	0,755	0,756	99,907	99,936	0,604
	$(CV(\hat{\theta}_a))^{-1}$	99,926	-1,343	0,756	0,756	99,911	99,941	0,589
	$(skewness(s_{dc}))^{-1}$	99,930	-1,009	0,770	0,771	99,914	99,945	0,603
	$(kurtosis(k_{dc}))^{-1}$	99,931	-0,908	0,814	0,815	99,915	99,947	0,671
	$meff/Var(y)$	99,928	-1,200	0,755	0,755	99,913	99,942	0,584
	$(meff)^{-1}$	99,925	-1,488	0,761	0,761	99,910	99,940	0,601
3	1/n	99,937	-0,282	0,601	0,602	99,925	99,949	0,362
	$(Var(\hat{\theta}_a))^{-1}$	99,925	-1,490	0,602	0,603	99,913	99,937	0,385
	$(CV(\hat{\theta}_a))^{-1}$	99,931	-0,837	0,607	0,607	99,919	99,943	0,375
	$(skewness(s_{dc}))^{-1}$	99,934	-0,573	0,650	0,650	99,921	99,947	0,426
	$(kurtosis(k_{dc}))^{-1}$	99,934	-0,610	0,725	0,726	99,919	99,948	0,530
	$meff/Var(y)$	99,933	-0,693	0,610	0,611	99,921	99,945	0,377
	$(meff)^{-1}$	99,930	-0,984	0,611	0,611	99,918	99,942	0,383
5	1/n	99,940	0,035	0,496	0,497	99,930	99,950	0,246
	$(Var(\hat{\theta}_a))^{-1}$	99,925	-1,440	0,477	0,477	99,916	99,935	0,248
	$(CV(\hat{\theta}_a))^{-1}$	99,932	-0,734	0,481	0,481	99,923	99,942	0,236
	$(skewness(s_{dc}))^{-1}$	99,935	-0,477	0,510	0,510	99,925	99,945	0,263
	$(kurtosis(k_{dc}))^{-1}$	99,933	-0,661	0,566	0,566	99,922	99,944	0,325
	$meff/Var(y)$	99,934	-0,598	0,487	0,487	99,924	99,943	0,241
	$(meff)^{-1}$	99,930	-0,949	0,485	0,485	99,921	99,940	0,244
10	1/n	99,942	0,195	0,335	0,335	99,935	99,948	0,113
	$(Var(\hat{\theta}_a))^{-1}$	99,926	-1,325	0,323	0,323	99,920	99,933	0,122
	$(CV(\hat{\theta}_a))^{-1}$	99,934	-0,606	0,325	0,325	99,927	99,940	0,109
	$(skewness(s_{dc}))^{-1}$	99,937	-0,259	0,351	0,351	99,930	99,944	0,124
	$(kurtosis(k_{dc}))^{-1}$	99,936	-0,351	0,396	0,396	99,928	99,944	0,158
	$meff/Var(y)$	99,935	-0,426	0,329	0,330	99,929	99,942	0,110
	$(meff)^{-1}$	99,932	-0,807	0,325	0,325	99,925	99,938	0,112
15	1/n	99,940	0,032	0,260	0,260	99,935	99,945	0,068
	$(Var(\hat{\theta}_a))^{-1}$	99,924	-1,567	0,253	0,253	99,919	99,929	0,089
	$(CV(\hat{\theta}_a))^{-1}$	99,931	-0,819	0,253	0,253	99,927	99,936	0,071
	$(skewness(s_{dc}))^{-1}$	99,934	-0,523	0,268	0,268	99,929	99,940	0,074
	$(kurtosis(k_{dc}))^{-1}$	99,933	-0,689	0,308	0,308	99,927	99,939	0,100
	$meff/Var(y)$	99,935	-0,468	0,288	0,288	99,929	99,941	0,085
	$(meff)^{-1}$	99,931	-0,913	0,264	0,264	99,925	99,936	0,078

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Table 4.42 continued: Scenario 3: Combining D samples of size n=10000, from a population of size N=1000 000, distributed as the skewed t-distribution with three degrees of freedom at(3) (100,25), $\theta=99.940$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
20	1/n	99,940	0,007	0,238	0,238	99,935	99,944	0,057
	$(Var(\hat{\theta}_d))^{-1}$	99,923	-1,634	0,229	0,229	99,919	99,928	0,079
	$(CV(\hat{\theta}_d))^{-1}$	99,931	-0,866	0,230	0,230	99,927	99,936	0,060
	$(skewness(s_{dc}))^{-1}$	99,934	-0,576	0,249	0,249	99,929	99,939	0,065
	$(kurtosis(k_{dc}))^{-1}$	99,932	-0,794	0,281	0,282	99,926	99,937	0,086
	$meff/Var(y)$	99,932	-0,746	0,252	0,252	99,927	99,937	0,069
	$(meff)^{-1}$	99,929	-1,085	0,231	0,231	99,924	99,933	0,065

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Weighting samples by the inverse of sample size resulted in the least MSE values except when ten samples were combined. Combining samples appeared to improve estimates.

Summary

For the superpopulation skewed t-distribution with three degrees of freedom, the inverse of sample size mostly resulted in the least MSE values. Combining estimates resulted in lower MSE values compared to using estimates from one sample for all the weighting strategies considered.

4.4.7 Summary of simulation results under two stage simple cluster sampling

Under cluster random sampling, the weighting methods generally performed similarly. As expected, the inverse of skewness did not work well with the symmetric superpopulation distributions. The inverse of sample size performed best for the t distribution superpopulations. Combining samples appeared to improve estimates, with decreasing MSE values as the number of samples combined increased. Design consistent weighting strategies largely resulted in different MSE values compared to the naïve weighting strategies.

Chapter 5

Methods for combining surveys: Variance estimators under the separate approach

The estimation of variance is an important aspect of statistical inference because it gives a measure of accuracy for the estimate and allows, for example, the construction of confidence intervals. Variance estimators derived under simple random sampling generally underestimate the variance when a complex survey design is used, especially as a result of clustering and unequal probabilities of selection. Design effects are often reported as a measure of the effect of the survey design on the variance. Variance estimation becomes more complex in the context of combined surveys. The estimation of variances needs to account for the sampling designs of the individual surveys together with the method used to combine the surveys.^[32]

5.1 Traditional methods

In this section we discuss some variance estimators which have been used in classical statistics and meta-analysis. We consider the advantages and disadvantages of using these estimators in the context of combining simple and complex sampling surveys, accounting for the survey sample designs.

In classical statistics, the variance of a sum of random variables, say $Y_1 + Y_2$ is calculated as

$$\text{Var}(Y_1 + Y_2) = \text{Var}(Y_1) + \text{Var}(Y_2) + 2\text{Cov}(Y_1, Y_2).$$

If we consider the simplest composite estimator, $\hat{\theta}_c = \alpha\hat{\theta}_1 + (1 - \alpha)\hat{\theta}_2$, the variance may be expressed as,

$$\text{Var}(\hat{\theta}_c) = \alpha^2 \text{var}(\hat{\theta}_1) + (1 - \alpha)^2 \text{var}(\hat{\theta}_2),$$

with $\alpha = w_1$, the weight allocated to survey 1. The covariance term is equal to zero since the surveys are assumed to be independent.^[59] In general, the classic variance of a sum of random variables may be expressed as

$$\text{var}(\hat{\theta}_c) = \sum_{d=1}^D w_d^2 * \text{var}(\hat{\theta}_d),$$

if the weights sum up to one and the surveys are independent.^[62] The composite estimator allows for different weights to be allocated to the surveys. However, this variance estimator is not be studied further in this thesis because the weighting strategies under consideration do not necessarily have to sum up to one.

When surveys are combined and parameter estimates are calculated as a straight forward average of the individual survey estimates, the variance may be expressed as^[74]

$$var(\hat{\theta}_c) = \frac{1}{D^2} \sum_{d=1}^D var(\hat{\theta}_d), \quad (5.1)$$

where $var(\hat{\theta}_d)$ is the estimator of the design consistent variance of the individual survey estimators $\hat{\theta}_d$. This variance estimator, henceforth referred to as the average variance estimators, was studied and simulation results are presented in chapter 6. The average variance estimator only accounts for the complex sampling design within individual surveys. It does not acknowledge the differences in sample size, precision or sampling design between the different surveys being combined. Equal weights are given to the surveys. Variance estimates calculated using the average variance estimator were compared to the empirical variance for the combined estimates obtained using the different weighting strategies considered in this thesis.

$$var_{empirical}(\hat{\theta}_c) = \frac{1}{\sum_{d=1}^D w_d} \sum_{d=1}^D \frac{w_d (\hat{\theta}_d - \hat{\theta}_c)^2}{D},$$

where $\hat{\theta}_c$ is the combined estimate, w_d is the weight allocated to survey $d = 1, \dots, D$ and $\hat{\theta}_d$ is the point estimate calculated from survey d .

In section 3.1 of this thesis, we discussed the use of the inverse of variance as a weighting strategy and assessed its performance through simulation in chapter 4. It is usually used in meta-analysis and has been used in combining surveys.^{[75][32]} Meta-analysis is usually used in clinical studies which aim to measure the effect size of an intervention. Combining clinical studies may result in a more accurate estimation of the effect size of the intervention. Under the fixed effects and random effects models, the combined effect is estimated as

$$\hat{\theta}_c = \frac{\sum_d^k w_d \hat{\theta}_d}{\sum_d^k w_d}$$

with $w_d = \frac{1}{\sigma_d^2}$ under the fixed effects model and $w_d = \frac{1}{\sigma_d^2 + \tau^2}$ under the random effects model. The variance of the estimate for the effect size for study d , is denoted σ_d^2 and τ^2 is the between study variance. Because the actual variances are usually unknown, their estimates may be used, leading to estimated weights $\hat{w}_d = \frac{1}{\hat{\sigma}_d^2}$; or $\hat{w}_d = \frac{1}{\hat{\sigma}_d^2 + \hat{\tau}^2}$. For simplicity, we used the fixed effects model in our simulations. The variance of the combined effect is usually estimated by^{[76][31]}

$$var(\hat{\theta}_c) = \frac{1}{\sum_d \hat{w}_d}, \quad (5.2)$$

where $\hat{w}_d = \frac{1}{\hat{\sigma}_d^2}$ under the fixed effects model.

If the estimates of the weights differ from the actual weights, then the variance is over-estimated.^[29,77] A weighted sample variance estimator and a robust variance estimator, which is robust to errors in the estimated variances from the individual studies, has been proposed for meta-analysis^[77]. The weighted sample variance estimator is

$$var_w(\hat{\theta}_c) = \frac{1}{D-1} \frac{\sum_{d=1}^D \hat{w}_d (\hat{\theta}_d - \hat{\theta}_c)^2}{\sum_{d=1}^D \hat{w}_d}, \quad (5.3)$$

The robust variance estimator is calculated as

$$var_R(\hat{\theta}_c) = \frac{1}{(\sum_{d=1}^D \hat{w}_d)^2} \sum_{d=1}^D \hat{w}_d^2 (\hat{\theta}_d - \hat{\theta}_c)^2. \quad (5.4)$$

The weighted meta-analysis variance estimator (equation 5.2), the weighted meta-analysis variance estimator (equation 5.2) and the robust meta-analysis variance estimator (equation 5.4) all account for the weights allocated to the different surveys being combined. These variance estimators were therefore evaluated through simulation in chapter 6, for use with the combined estimates investigated in chapter 4.

5.2 Replication methods

Replication methods can be used when the micro data of the individual surveys to be combined is available. These include, among others, the bootstrap and Jackknife methods of estimation.

The bootstrap method was proposed by Efron in 1979.^[78] It is a simulation method where new data sets are generated from the original or observed data. No prior assumptions about the distribution of the estimators and the observations are required. Techniques for estimating variance under complex survey designs are often cumbersome making resampling techniques more attractive.

The jackknife method was first introduced by Quenouille as a means of reducing bias in parameter estimates.^{[79][21]} Tukey suggested that the jackknife method could also be used to produce variance estimates for many estimators, including parameter estimates, given finite population samples.^{[21][80]} For variance estimation, the sample is split into equal sized, disjoint and exhaustive subsamples. Each of these subsamples is dropped out in turn; and the parameter of interest is estimated from the remaining sample (replicate sample). The variability among these estimates is then used to estimate the variance of the original sample estimator.^[80]

In the delete-1 jackknife method, the sample is split into samples of size 1. The replicate samples used for estimation are obtained by deleting one observation from the sample. For a sample of size n , there are n replicate samples. Let the estimate of the parameter of interest, estimated from the sample where observation i has been deleted be $\hat{\theta}_{(i)}$, then the estimate of the variance is given by

$$Var_{jack}(\hat{\theta}) = \frac{n-1}{n} \sum_{i=1}^n (\hat{\theta}_{(i)} - \hat{\theta}_{(\cdot)})^2$$

where $\hat{\theta}_{(\cdot)}$ is the average of the jackknife estimates. The method can be generalized to delete-k jackknife method.^[81]

Miller suggested research in the use of the jackknife method in the multi-sample situation. Considering the two-sample problem, one could delete observations from the first sample in turn leaving the second sample intact; then delete observations from the second sample leaving the first sample intact. Alternatively, one could jackknife the samples separately then combine the results.^{[82][80]}

Although the jackknife makes no assumptions about the shape of the underlying probability distribution, it requires that the observations are independent of each other. When the independence assumption is violated, the jackknife method underestimates the variance in the dataset which makes the data look more reliable than they are.

The properties of the traditional variance estimators will be studied in the context of combining surveys. Replication methods will not be studied further in this thesis. We leave them for future studies. We decided to focus on meta-analysis linked variance estimators as it is closely related to combining surveys.

Chapter 6

Simulation study 2 (Variance estimators in chapter 5)

The variance gives an indication of the precision of point estimates. It is therefore important to report point estimates together with the estimates of their variances. In this chapter, results are presented for repeated simulations that were conducted where sample point estimates were combined using the following different weighting strategies: the inverse of the sample size, the inverse of the variance, the inverse of the coefficient of variation, the inverse of skewness, the inverse of kurtosis and the inverse of the D'Agostino-Pearson statistic for samples collected using simple random sampling. Additionally, for survey data collected using stratified random sampling, a weight of 1, the inverse of m_{eff} and $m_{eff}/Var(y)$ were also evaluated as weighting strategies.

The simulation process was repeated to allow for the estimation of variance of the combined estimate, for each weighting strategy, using selected variance estimators discussed in chapter 5. The resulting variance estimates were then compared to the empirical variance estimate discussed in section 5.1.

The simulation study presented in the current chapter aimed to identify the variance estimator which works best for each weighting strategy, for different scenarios. The variance estimators evaluated in this study are the average variance estimator (equation 5.1) the meta-analysis variance estimator (equation 5.2), weighted meta-analysis variance estimator (equation 5.3) and the robust meta-analysis variance estimator (equation 5.4), discussed in chapter 5.

Moreover, the current simulation study aimed to evaluate the performance of the different variance estimators under different sampling strategies, that is: simple random sampling, stratified random sampling and cluster random sampling. However, simulation results for cluster sampling are not presented in this thesis because they were not according to expectation and have been left for further study. Most of the relative differences observed between the variance obtained by simulation and the variances calculated using the studied variance estimators were unacceptably high and will not be presented in this thesis.

6.1 Simulation set up

In a similar way to the simulation described in chapter 4, finite populations of size 100 000 were simulated from finite populations with different distributions: normal distribution, skewed normal distribution, t distribution with three degrees of freedom, t distribution with four degrees of freedom, skewed t distribution with three degrees of freedom and skewed t distribution with four degrees of freedom. For each finite population, $D= 2, 3, 5, 10, 15$ or 20 simple random samples of size $n = 1200$ were drawn and used to estimate the population parameter θ . The estimates were then combined, using the different weighting strategies described in chapter 4 to yield combined estimate $\hat{\theta}_c$. The combined point estimates were repeatedly estimated using the different weighting strategies as described in chapter 4. This process was repeated $K = 1\ 000$ times. Variance estimates were calculated using the average variance estimator (equation 5.1), the meta-analysis variance estimator (equation 5.2), the weighted meta-analysis variance estimator (equation 5.3) and the robust meta-analysis variance estimator (equation 5.4). The corresponding standard errors were then compared to the standard error obtained empirically in chapter 4, using the relative difference. In short, a finite population was generated from which $D=2, 3, 5, 10, 15$ or 20 samples were selected and an estimate of the mean $\hat{\theta}_d$ was calculated from each of the samples before being combined. The standard error based on the empirical variance discussed in section 5.1 was calculated as

$$s.e_{empirical} = \sqrt{\frac{1}{D \sum_{d=1}^D w_d} \sum_{d=1}^D w_d (\hat{\theta}_d - \hat{\theta}_c)^2},$$

where $\hat{\theta}_c$ is the combined estimate for each of weighting strategies considered. The relative differences were calculated for each weighting strategy and variance estimator. The relative difference was calculated as

$$Rel. diff = \frac{s.e_{estimator} - s.e_{empirical}}{s.e_{empirical}},$$

where

$$s.e_{empirical}(\hat{\theta}_c) = \sqrt{var_{empirical}(\hat{\theta}_c)}$$

The simulation set up is summarized in figure 9.

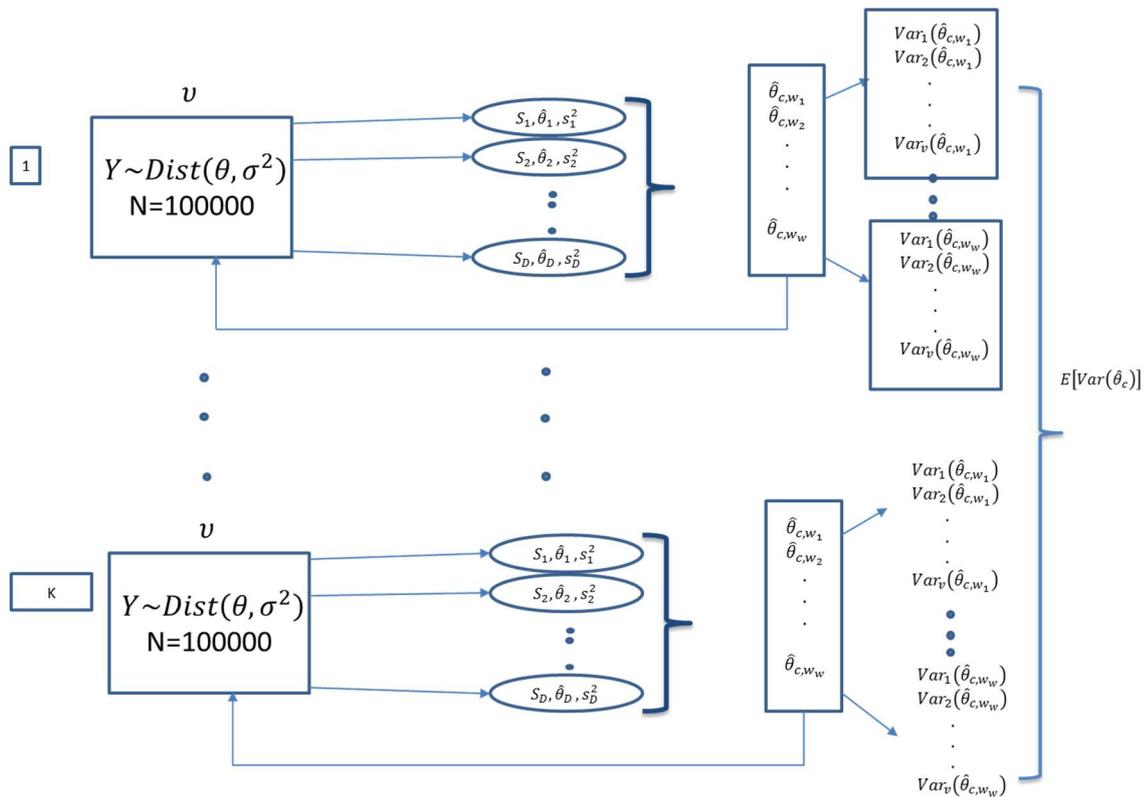


Figure 9: Evaluating variance estimators

In this simulation, the finite population is simulated K times. From each finite population, samples, denoted by S_d , $d = 1, \dots, D$, were collected. The estimate of the mean $\hat{\theta}_d$, and the estimate of the variance s_d^2 were then calculated using the sample data S_d . In the diagram, $\hat{\theta}_{c,w_w}$ is the combined estimate, calculated using weighting strategy $w = 1, 2, \dots, W$ and $\text{Var}_v(\hat{\theta}_{c,w_w})$ is the variance of the combined estimate calculate using weighting strategy w and variance estimator $v = 1, \dots, V$. For example, $\text{Var}_1(\hat{\theta}_{c,w_1})$ would be the variance obtained using the average variance estimator for the combined estimate of θ , calculated using the inverse of the sample size as a weighting strategy.

Parallel computing in a computer server was used to reduce the amount of time required for a simulation to run because of the large number of sampling replicates required. The R library doParallel was used.^[83] When we ask a computer to do a computation, for example calculating a variance, the computation is broken down into a series of instructions which are then executed one after the other on a single computer processor. The computer processor executes the program instructions in a step by step fashion, sequentially as illustrated below.^[84]

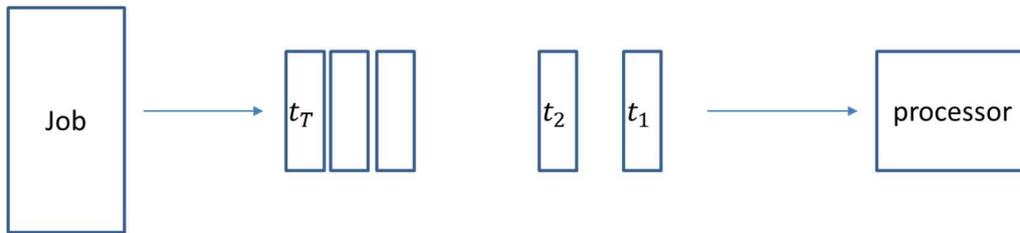


Figure 10: Serial computing

In figure 10, $t_t, t = 1, \dots, T$ represents time. However, the job can be done faster if some instructions can be executed simultaneously. In parallel computing, the job is partitioned into different parts which can be executed concurrently. Each part is then broken down into a series of instructions which will then be processed simultaneously by different processors as illustrated in figure 11.^[84]

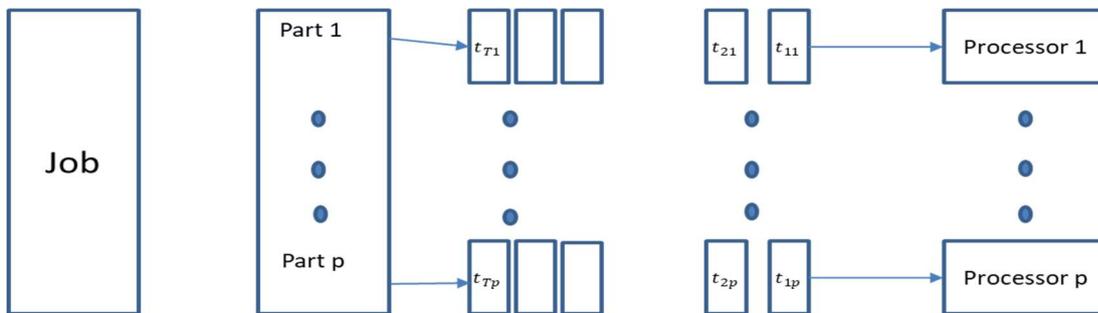


Figure 11: Parallel computing

6.2 Simulation results for variance estimators under simple random sampling

Variance estimators were compared for each weighting strategy considered in this thesis. The normal distribution superpopulation was used to generate a finite population for the simulation study. Other superpopulation distributions have been left for future studies. Relative differences were used to compare the variance estimators.

Unlike the situation with the point estimators, using the inverse of the sample size did not yield the same variance estimation results as using a weight of one when combining the surveys. Hence the relative differences for both weighting strategies have been presented. Relative differences where the variance estimator performed relatively well with the weighting strategy have been highlighted in bold.

The meta-analysis-based variance estimators considered in thesis appeared not to work well with non-normal superpopulation distributions. The empirical variance estimates were then used to assess precision for point estimates obtained from these superpopulation distributions in the supplementary analysis, section 6.3.

Section 6.2.1 below presents simulation results comparing the variance estimators under the simple random sampling with a normal distribution superpopulation.

6.2.1 Simulation results for the normal distribution under simple random sampling

The relative differences between the variance estimate obtained through simulation and the variance estimates calculated using the four variance estimators considered in this study are presented in table 6.1, for each weighting strategy assessed through simulation in chapter 4. Variance estimators 1 to 4 were presented in chapter 5 as equations 5.1 to 5.4 respectively.

Table 6. 1: Comparing Relative differences across weighting strategies and number of samples combined, for each variance estimator for the normal distribution under simple random sampling, n=1000

D	Variance estimator	Relative difference x 100% for Weighting strategy						
		n	$(Var(\hat{\theta}_d))^{-1}$	1	$(CV(\hat{\theta}_d))^{-1}$	$(skewness(s_d))^{-1}$	$(kurtosis(K_d))^{-1}$	$(K^2)^{-1}$
1		-0,0072						
2	1	2,0523	2,0485	2,0523	2,0512	1,7360	2,0420	1,6683
	2	-0,8018	-0,0117	5,2676	-0,7511	0,2066	9,7967	5,2530
	3	-0,3444	-0,3451	-0,3444	-0,3445	-0,4966	-0,3462	-0,5203
	4	-0,5364	-0,5375	-0,5364	-0,5367	-0,6736	-0,5385	-0,6870
3	1	1,4111	1,4079	1,4111	1,4103	0,9589	1,4079	0,9580
	2	-0,8030	-0,0176	5,2299	-0,7525	0,0174	9,7536	4,2881
	3	-0,2375	-0,2391	-0,2375	-0,2380	-0,4678	-0,2388	-0,4828
	4	-0,3774	-0,3784	-0,3774	-0,3776	-0,5931	-0,3792	-0,5922
5	1	-0,7532	-0,7510	-0,7532	-0,7527	-0,2784	-0,7499	-0,2950
	2	0,8038	0,0218	-5,2032	0,7536	0,1743	-9,7024	-3,3858
	3	0,1799	0,1814	0,1799	0,1803	0,5029	0,1818	0,4901
	4	0,2665	0,2677	0,2665	0,2668	0,5501	0,2680	0,5247
10	1	0,0347	0,0348	0,0347	0,0350	-0,3960	0,0348	-0,3579
	2	-0,8025	-0,0143	5,2439	-0,7518	-0,4084	9,7880	2,3052
	3	-0,1219	-0,1221	-0,1219	-0,1218	-0,5702	-0,1222	-0,5333
	4	-0,1670	-0,1667	-0,1670	-0,1667	-0,5243	-0,1669	-0,4675
15	1	-0,3573	-0,3577	-0,3573	-0,3574	-0,6870	-0,3587	-0,6541
	2	-0,7966	0,0144	5,4307	-0,7445	-0,5200	10,0855	1,7707
	3	-0,0656	-0,0664	-0,0656	-0,0659	-0,6181	-0,0677	-0,5646
	4	-0,0973	-0,0977	-0,0973	-0,0975	-0,5286	-0,0984	-0,4381
20	1	-0,7716	-0,7719	-0,7716	-0,7718	-0,8963	-0,7718	-0,8849
	2	-0,7959	0,0175	5,4551	-0,7436	-0,5677	10,1446	1,5580
	3	-0,0483	-0,0502	-0,0483	-0,0491	-0,6383	-0,0490	-0,5790
	4	-0,0724	-0,0740	-0,0724	-0,0732	-0,5125	-0,0722	-0,4182

D=number of samples combined, $\hat{\theta}_d$ = estimate calculated from each individual survey, K^2 = D'Agostino Pearson statistic, CV=coefficient of variation, 1=average variance estimator, 2=meta-analysis variance estimator, 3= weighted meta-analysis variance estimator, 4=robust meta-analysis variance estimator.

When fifteen or more samples were combined, weighted meta-analysis and the robust meta-analysis variance estimators yielded relative differences less than 10% when the sample estimates were combined without weighting or the inverses of sample size, variance and kurtosis were used as weighting strategies. The meta-analysis variance estimator resulted in the lowest relative difference when the

inverse of variance was used as the weighting strategy. This is consistent with meta-analysis theory. None of the variance estimators worked well when the inverse of skewness or the inverse of the D'Agostino-Pearson statistic were used as the weighting strategies for combining the samples. This was expected as the finite population was generated from a symmetric superpopulation distribution.

Table 6. 2: Comparing Relative differences across weighting strategies and number of samples combined, for each variance estimator for the normal distribution under simple random sampling, n=5000

D	Variance estimator	Relative difference x 100% for Weighting strategy					
		n	$(Var(\hat{\theta}_d))^{-1}$	1	$(CV(\hat{\theta}_d))^{-1}$	$(skewness(s_d))^{-1}$	$(kurtosis(K_d))^{-1}$
1		-0,0312					
2	1	1,9778	1,9788	1,9778	1,9780	1,6335	1,97650
	2	-0,8063	-0,0335	12,698	-0,6361	0,7107	22,7035
	3	-0,3072	-0,3074	-0,3072	-0,3073	-0,4991	-0,3076
	4	-0,5101	-0,5102	-0,5101	-0,5101	-0,6764	-0,5105
3	1	1,3919	1,3922	1,3919	1,3922	0,9479	1,3889
	2	-0,8042	-0,0232	12,8459	-0,6322	0,4850	22,9395
	3	-0,2539	-0,2540	-0,2539	-0,2539	-0,4925	-0,2546
	4	-0,3908	-0,3908	-0,3908	-0,3907	-0,6048	-0,3914
5	1	0,8226	0,82307	0,8226	0,8229	0,2807	0,8206
	2	-0,7957	0,01926	13,4463	-0,6163	0,2170	23,9806
	3	-0,1321	-0,1320	-0,1321	-0,1320	-0,4893	-0,1329
	4	-0,2238	-0,2235	-0,2238	-0,2236	-0,5274	-0,2243
10	1	0,0775	0,07716	0,0775	0,0774	-0,3675	0,0763
	2	-0,7940	0,02720	13,5691	-0,6131	-0,0689	24,1925
	3	-0,0811	-0,0814	-0,0811	-0,0812	-0,5424	-0,0821
	4	-0,1283	-0,1284	-0,1283	-0,1283	-0,4893	-0,1290
15	1	-0,3593	-0,3591	-0,3593	-0,3591	-0,6626	-0,3601
	2	-0,7968	0,0135	13,3676	-0,6184	-0,2193	23,8436
	3	-0,0659	-0,0657	-0,0659	-0,0657	-0,5813	-0,0669
	4	-0,0976	-0,0972	-0,0976	-0,0973	-0,4808	-0,0983
20	1	-0,7697	-0,7695	-0,7697	-0,7696	-0,8935	-0,7700
	2	-0,7935	0,0304	13,6021	-0,6121	-0,3308	24,2485
	3	-0,0379	-0,0374	-0,0379	-0,0376	-0,6193	-0,0390
	4	-0,0623	-0,0617	-0,0623	-0,0619	-0,4820	-0,0631

D=number of samples combined, $\hat{\theta}_d$ = estimate calculated from each individual survey, K^2 = D'Agostino Pearson statistic, CV=coefficient of variation, 1=average variance estimator, 2=meta-analysis variance estimator, 3= weighted meta-analysis variance estimator, 4= robust meta-analysis variance estimator

The meta-analysis variance estimator worked well when the samples were combined using the inverse of variance as the weighting strategy, resulting in relative

differences less than 10% for all the numbers of samples combined. The weighted meta-analysis variance estimator performed well when ten or more samples were combined. It resulted in relative differences less than 10 % for all the weighting strategies except the inverse of skewness. The robust variance estimator worked well with all weighting strategies, again with the exception of the inverse of skewness when fifteen or twenty samples were combined. The finite population was generated from a symmetric superpopulation distribution therefore it was not expected that the inverse of skewness weighting strategy would work well with any variance estimator.

Table 6. 3: Comparing Relative differences across weighting strategies and number of samples combined, for each variance estimator for the normal distribution under simple random sampling, n=10000

D	Variance estimator	Relative difference x 100% for Weighting strategy					
		n	$(Var(\hat{\theta}_d))^{-1}$	1	$(CV(\hat{\theta}_d))^{-1}$	$(skewness(s_d))^{-1}$	$(kurtosis(K_d))^{-1}$
1		-0,0010					
2	1	2,0042	2,0034	2,0042	2,0038	1,7478	2,0030
	2	-0,8041	-0,0258	18,5924	-0,5631	1,2295	32,9078
	3	-0,3032	-0,3036	-0,3032	-0,3034	-0,4543	-0,3035
	4	-0,5073	-0,5075	-0,5073	-0,5074	-0,6410	-0,5076
3	1	1,4241	1,4242	1,4241	1,4242	1,0451	1,4240
	2	-0,8011	-0,0104	18,8931	-0,5563	0,9071	33,4424
	3	-0,2384	-0,2383	-0,2384	-0,2383	-0,4713	-0,2385
	4	-0,3781	-0,3781	-0,3781	-0,3781	-0,5864	-0,3783
5	1	0,7489	0,7498	0,7489	0,7494	0,2900	0,7495
	2	-0,8035	-0,0219	18,6526	-0,5616	0,5060	33,0373
	3	-0,1797	-0,1792	-0,1797	-0,1795	-0,5025	-0,1796
	4	-0,2663	-0,2659	-0,2663	-0,2661	-0,5462	-0,2662
10	1	0,0609	0,0611	0,0609	0,0610	-0,3734	0,0610
	2	-0,7967	0,0115	19,3325	-0,5465	0,1044	34,2091
	3	-0,0945	-0,0943	-0,0945	-0,0944	-0,5468	-0,0943
	4	-0,1410	-0,1407	-0,1410	-0,1408	-0,4858	-0,1405
15	1	-0,3630	-0,3628	-0,3630	-0,3629	-0,68106	-0,3631
	2	-0,7976	0,0074	19,2449	-0,5484	-0,1295	34,0463
	3	-0,0717	-0,0713	-0,0717	-0,0715	-0,6059	-0,0718
	4	-0,1031	-0,1028	-0,1031	-0,1029	-0,5032	-0,1031
20	1	-0,7774	-0,7773	-0,7774	-0,7774	-0,8985	-0,7775
	2	-0,7999	-0,0043	19,0078	-0,5536	-0,2382	33,6243
	3	-0,0718	-0,0714	-0,0718	-0,0715	-0,6390	-0,0724
	4	-0,0953	-0,0948	-0,0953	-0,0950	-0,5109	-0,0958

D=number of samples combined, $\hat{\theta}_d$ = estimate calculated from each individual survey, K^2 = D'Agostino Pearson statistic, CV=coefficient of variation, 1=average variance estimator, 2=meta-analysis variance estimator, 3= weighted meta-analysis variance estimator, 4= robust meta-analysis variance estimator

Similar to sample sizes 1000 and 5000, the meta-analysis variance estimator resulted in the least relative differences when the inverse of variance was used as the weighting strategy when combining the samples. When twenty samples were combined, the weighted meta-analysis and robust meta-analysis variance estimators performed well with all the weighting strategies apart from the inverse of skewness.

The weighted meta-analysis variance estimator appears to work well when ten or more samples are combined.

6.1.4 Summary of variance estimator simulation results under simple random sampling

Comparisons between the variance estimators under consideration were made, assuming a superpopulation with a normal distribution. The meta-analysis variance estimator consistently resulted in low relative differences, less than 10% when the inverse of variance was used as the weighting strategy when combining samples. This was observed for all sample sizes and all numbers of samples combined. The weighted meta-analysis and the robust meta-analysis variance estimators worked well with all the weighting strategies except the inverse of skewness when ten or more samples were combined. This was expected because the finite population used in the simulation was generated from a symmetric superpopulation distribution. The average variance estimator did not work well with any of the weighting strategies.

The variance estimators did not work well with all the other superpopulation distributions considered in this thesis and results will not be presented here. The meta-analysis variance estimators assume normal distribution^[77].

6.3 Supplementary Analysis

Simulation indicated that the variance estimators based on meta-analysis, studied in this thesis were not suitable for evaluating precision of estimates when the superpopulation was not distributed as the normal distribution. In addition, the proposed variance estimators did not work well with stratified or cluster random sampling. Precision of the point estimators was only evaluated by the empirical variance. Standard errors were compared across the sampling distributions for each weighting strategy and number of samples combined. Tables of comparison for the finite population size $N=1\ 000\ 000$ and sample size $n=10\ 000$ are presented below.

Table 6. 4: Comparing standard errors across sampling strategies, weighting strategies and number of samples combined, for a finite population distributed the normal distribution $N(100,25)$, $n=10000$.

D	Weighting strategy	Simple random sampling $\times 10^{-2}$	Stratified random sampling $\times 10^{-3}$	Cluster random sampling $\times 10^{-3}$
1		4,978	1,086	9,764
2	$1/n$	3,609	0,778	6,104
	$(Var(\hat{\theta}_a))^{-1}$	3,610	0,782	6,167
	$(CV(\hat{\theta}_a))^{-1}$	3,610	0,782	6,170
	$(skewness(s_{dc}))^{-1}$	3,946	0,881	6,607
	$(kurtosis(k_{dc}))^{-1}$	3,611	0,782	6,175
	$mef f / Var(y)$		0,782	6,180
	$(mef f)^{-1}$		0,782	6,162
3	$1/n$	2,902	0,646	5,223
	$(Var(\hat{\theta}_a))^{-1}$	2,902	0,648	5,021
	$(CV(\hat{\theta}_a))^{-1}$	2,902	0,648	5,020
	$(skewness(s_{dc}))^{-1}$	3,440	0,778	6,199
	$(kurtosis(k_{dc}))^{-1}$	2,902	0,647	5,015
	$mef f / Var(y)$		0,647	5,016
	$(mef f)^{-1}$		0,648	5,018
5	$1/n$	2,276	0,489	4,067
	$(Var(\hat{\theta}_a))^{-1}$	2,274	0,489	4,087
	$(CV(\hat{\theta}_a))^{-1}$	2,275	0,489	4,076
	$(skewness(s_{dc}))^{-1}$	3,085	0,673	5,433
	$(kurtosis(k_{dc}))^{-1}$	2,275	0,489	4,063
	$mef f / Var(y)$		0,489	4,011
	$(mef f)^{-1}$		0,490	4,052
10	$1/n$	1,555	0,348	2,808
	$(Var(\hat{\theta}_a))^{-1}$	1,555	0,348	2,822
	$(CV(\hat{\theta}_a))^{-1}$	1,555	0,348	2,814
	$(skewness(s_{dc}))^{-1}$	2,633	0,595	4,409
	$(kurtosis(k_{dc}))^{-1}$	1,555	0,348	2,810
	$mef f / Var(y)$		0,350	2,816
	$(mef f)^{-1}$		0,349	2,829

D	Weighting strategy	Simple random sampling $\times 10^{-2}$	Stratified random sampling $\times 10^{-3}$	Cluster random sampling $\times 10^{-3}$
15	$1/n$	1,275	0,282	2,468
	$(\text{Var}(\hat{\theta}_a))^{-1}$	1,275	0,283	2,469
	$(\text{CV}(\hat{\theta}_a))^{-1}$	1,275	0,283	2,467
	$(\text{skewness}(s_{dc}))^{-1}$	2,546	0,550	4,337
	$(\text{kurtosis}(k_{dc}))^{-1}$	1,276	0,283	2,471
	$mef f / \text{Var}(y)$		0,283	2,745
	$(mef f)^{-1}$		0,283	2,437
20	$1/n$	1,118	0,243	1,982
	$(\text{Var}(\hat{\theta}_a))^{-1}$	1,117	0,243	1,978
	$(\text{CV}(\hat{\theta}_a))^{-1}$	1,117	0,243	1,978
	$(\text{skewness}(s_{dc}))^{-1}$	2,451	0,525	4,785
	$(\text{kurtosis}(k_{dc}))^{-1}$	1,118	0,243	1,984
	$mef f / \text{Var}(y)$		0,246	2,049
	$(mef f)^{-1}$		0,245	2,000

The lowest standard errors obtained errors were obtained when stratified random sampling was used for all the weighting methods. This was consistent for all numbers of samples combined.

The standard errors resulting from using a superpopulation distributed as the skewed normal distribution are presented in table 6.5.

Table 6. 5: Comparing standard errors across sampling strategies, weighting strategies and number of samples combined, for the finite population distributed as the skewed normal distribution, SN(100,25), n=10000

D	Weighting strategy	Simple random sampling $\times 10^{-3}$	Stratified random sampling $\times 10^{-3}$	Cluster random sampling $\times 10^{-2}$
1		1,595	1,086	1,011
2	$1/n$	1,143	0,777	0,668
	$(Var(\hat{\theta}_a))^{-1}$	1,143	0,783	0,722
	$(CV(\hat{\theta}_a))^{-1}$	1,143	0,783	0,722
	$(skewness(s_{dc}))^{-1}$	1,144	0,783	0,727
	$(kurtosis(k_{dc}))^{-1}$	1,143	0,782	0,723
	$mef f / Var(y)$		0,782	0,722
	$(mef f)^{-1}$		0,783	0,722
3	$1/n$	0,953	0,635	0,552
	$(Var(\hat{\theta}_a))^{-1}$	0,952	0,638	0,584
	$(CV(\hat{\theta}_a))^{-1}$	0,952	0,638	0,583
	$(skewness(s_{dc}))^{-1}$	0,953	0,640	0,584
	$(kurtosis(k_{dc}))^{-1}$	0,953	0,638	0,583
	$mef f / Var(y)$		0,638	0,582
	$(mef f)^{-1}$		0,638	0,584
5	$1/n$	0,726	0,475	0,446
	$(Var(\hat{\theta}_a))^{-1}$	0,726	0,475	0,444
	$(CV(\hat{\theta}_a))^{-1}$	0,726	0,475	0,445
	$(skewness(s_{dc}))^{-1}$	0,727	0,476	0,451
	$(kurtosis(k_{dc}))^{-1}$	0,727	0,475	0,450
	$mef f / Var(y)$		0,475	0,446
	$(mef f)^{-1}$		0,476	0,446
10	$1/n$	0,498	0,342	0,306
	$(Var(\hat{\theta}_a))^{-1}$	0,498	0,342	0,311
	$(CV(\hat{\theta}_a))^{-1}$	0,498	0,342	0,308
	$(skewness(s_{dc}))^{-1}$	0,499	0,342	0,307
	$(kurtosis(k_{dc}))^{-1}$	0,498	0,342	0,307
	$mef f / Var(y)$		0,342	0,304
	$(mef f)^{-1}$		0,342	0,307

D	Weighting strategy	Simple random sampling $\times 10^{-2}$	Stratified random sampling $\times 10^{-3}$	Cluster random sampling $\times 10^{-3}$
15	$1/n$	0,405	0,280	0,236
	$(\text{Var}(\hat{\theta}_a))^{-1}$	0,405	0,280	0,239
	$(\text{CV}(\hat{\theta}_a))^{-1}$	0,405	0,280	0,237
	$(\text{skewness}(s_{dc}))^{-1}$	0,406	0,281	0,238
	$(\text{kurtosis}(k_{dc}))^{-1}$	0,405	0,280	0,237
	$mef f / \text{Var}(y)$		0,281	0,246
	$(mef f)^{-1}$		0,281	0,238
20	$1/n$	0,355	0,244	0,217
	$(\text{Var}(\hat{\theta}_a))^{-1}$	0,355	0,244	0,219
	$(\text{CV}(\hat{\theta}_a))^{-1}$	0,355	0,244	0,218
	$(\text{skewness}(s_{dc}))^{-1}$	0,355	0,245	0,218
	$(\text{kurtosis}(k_{dc}))^{-1}$	0,355	0,244	0,217
	$mef f / \text{Var}(y)$		0,245	0,216
	$(mef f)^{-1}$		0,244	0,226

In this simulation, stratified random sampling appeared to increase the efficiency. The empirical standard errors obtained through simple random sampling were larger than those obtained when stratified random sampling was used, for all the numbers of samples combined. Cluster sampling resulted in larger empirical standard errors. This is consistent with known statistical sampling theory.

Standard errors obtained when the superpopulation was distributed as the t-distribution with three degrees of freedom were compared in table 6.6.

Table 6. 6: Comparing standard errors across sampling strategies, weighting strategies and number of samples combined, for the finite population distributed as the t distribution,3d.f, $t(100,25)$, $n=10000$

D	Weighting strategy	Simple random sampling $\times 10^{-3}$	Stratified random sampling $\times 10^{-3}$	Cluster random sampling $\times 10^{-3}$
1		1,546	1,093	9,449
2	$1/n$	1,093	0,783	6,461
	$(Var(\hat{\theta}_a))^{-1}$	1,087	0,789	6,410
	$(CV(\hat{\theta}_a))^{-1}$	1,089	0,788	6,410
	$(skewness(s_{dc}))^{-1}$	1,213	0,887	8,589
	$(kurtosis(k_{dc}))^{-1}$	1,169	0,860	7,682
	$mef f / Var(y)$		0,787	6,491
	$(mef f)^{-1}$		0,788	6,379
3	$1/n$	0,887	0,635	5,228
	$(Var(\hat{\theta}_a))^{-1}$	0,881	0,635	5,267
	$(CV(\hat{\theta}_a))^{-1}$	0,882	0,635	5,255
	$(skewness(s_{dc}))^{-1}$	1,086	0,776	7,140
	$(kurtosis(k_{dc}))^{-1}$	0,979	0,703	6,070
	$mef f / Var(y)$		0,635	5,263
	$(mef f)^{-1}$		0,635	5,274
5	$1/n$	0,665	0,479	4,491
	$(Var(\hat{\theta}_a))^{-1}$	0,660	0,476	4,489
	$(CV(\hat{\theta}_a))^{-1}$	0,661	0,477	4,487
	$(skewness(s_{dc}))^{-1}$	0,917	0,658	6,777
	$(kurtosis(k_{dc}))^{-1}$	0,743	0,534	5,214
	$mef f / Var(y)$		0,479	4,545
	$(mef f)^{-1}$		0,476	4,512
10	$1/n$	0,482	0,346	3,162
	$(Var(\hat{\theta}_a))^{-1}$	0,480	0,344	3,205
	$(CV(\hat{\theta}_a))^{-1}$	0,480	0,344	3,179
	$(skewness(s_{dc}))^{-1}$	0,785	0,544	5,886
	$(kurtosis(k_{dc}))^{-1}$	0,543	0,376	3,931
	$mef f / Var(y)$		0,345	3,219
	$(mef f)^{-1}$		0,345	3,140

D	Weighting strategy	Simple random sampling $\times 10^{-2}$	Stratified random sampling $\times 10^{-3}$	Cluster random sampling $\times 10^{-3}$
15	$1/n$	0,401	0,282	2,501
	$(\text{Var}(\hat{\theta}_a))^{-1}$	0,399	0,281	2,471
	$(\text{CV}(\hat{\theta}_a))^{-1}$	0,399	0,281	2,501
	$(\text{skewness}(s_{dc}))^{-1}$	0,714	0,520	2,481
	$(\text{kurtosis}(k_{dc}))^{-1}$	0,443	0,311	5,437
	$m_{eff}/\text{Var}(y)$		0,281	3,058
	$(m_{eff})^{-1}$		0,281	2,552
20	$1/n$	0,347	0,240	2,566
	$(\text{Var}(\hat{\theta}_a))^{-1}$	0,345	0,238	2,291
	$(\text{CV}(\hat{\theta}_a))^{-1}$	0,346	0,239	2,259
	$(\text{skewness}(s_{dc}))^{-1}$	0,673	0,506	2,271
	$(\text{kurtosis}(k_{dc}))^{-1}$	0,377	0,266	4,919
	$m_{eff}/\text{Var}(y)$		0,239	2,594
	$(m_{eff})^{-1}$		0,238	2,284

Stratified random sampling resulted in the smallest empirical standard errors, followed by simple random sampling, with cluster random sampling resulting in the largest empirical standard errors.

The same comparisons were done for a finite population resulting from a superpopulation distributed as the skewed t-distribution with three degrees of freedom. The results are presented in table 6.7 below.

Table 6. 7: Comparing standard errors across sampling strategies, weighting strategies and number of samples combined, for the finite population distributed as the skewed t distribution,3d.f, at(3)(100,25), n=10000

D	Weighting strategy	Simple random sampling $\times 10^{-3}$	Stratified random sampling $\times 10^{-3}$	Cluster random sampling $\times 10^{-2}$
1		1,573	1,083	1,228
2	$1/n$	1,095	0,757	0,728
	$(Var(\hat{\theta}_a))^{-1}$	1,091	0,763	0,755
	$(CV(\hat{\theta}_a))^{-1}$	1,092	0,761	0,756
	$(skewness(s_{dc}))^{-1}$	1,114	0,796	0,770
	$(kurtosis(k_{dc}))^{-1}$	1,184	0,835	0,814
	$mef f/Var(y)$		0,762	0,755
	$(mef f)^{-1}$		0,761	0,761
3	$1/n$	0,905	0,623	0,601
	$(Var(\hat{\theta}_a))^{-1}$	0,903	0,625	0,602
	$(CV(\hat{\theta}_a))^{-1}$	0,903	0,625	0,607
	$(skewness(s_{dc}))^{-1}$	0,937	0,667	0,650
	$(kurtosis(k_{dc}))^{-1}$	1,026	0,722	0,725
	$mef f/Var(y)$		0,627	0,610
	$(mef f)^{-1}$		0,624	0,611
5	$1/n$	0,716	0,494	0,496
	$(Var(\hat{\theta}_a))^{-1}$	0,717	0,493	0,477
	$(CV(\hat{\theta}_a))^{-1}$	0,715	0,493	0,481
	$(skewness(s_{dc}))^{-1}$	0,745	0,521	0,510
	$(kurtosis(k_{dc}))^{-1}$	0,821	0,565	0,566
	$mef f/Var(y)$		0,492	0,487
	$(mef f)^{-1}$		0,494	0,485
10	$1/n$	0,510	0,353	0,335
	$(Var(\hat{\theta}_a))^{-1}$	0,512	0,351	0,323
	$(CV(\hat{\theta}_a))^{-1}$	0,510	0,351	0,325
	$(skewness(s_{dc}))^{-1}$	0,535	0,371	0,351
	$(kurtosis(k_{dc}))^{-1}$	0,593	0,404	0,396
	$mef f/Var(y)$		0,352	0,329
	$(mef f)^{-1}$		0,351	0,325

D	Weighting strategy	Simple random sampling $\times 10^{-2}$	Stratified random sampling $\times 10^{-3}$	Cluster random sampling $\times 10^{-3}$
15	$1/n$	0,411	0,277	0,260
	$(\text{Var}(\hat{\theta}_a))^{-1}$	0,411	0,275	0,253
	$(\text{CV}(\hat{\theta}_a))^{-1}$	0,410	0,275	0,253
	$(\text{skewness}(s_{dc}))^{-1}$	0,426	0,293	0,268
	$(\text{kurtosis}(k_{dc}))^{-1}$	0,469	0,323	0,308
	$mef f / \text{Var}(y)$		0,278	0,288
	$(mef f)^{-1}$		0,275	0,264
20	$1/n$	0,364	0,242	0,238
	$(\text{Var}(\hat{\theta}_a))^{-1}$	0,362	0,241	0,229
	$(\text{CV}(\hat{\theta}_a))^{-1}$	0,362	0,241	0,230
	$(\text{skewness}(s_{dc}))^{-1}$	0,375	0,257	0,249
	$(\text{kurtosis}(k_{dc}))^{-1}$	0,413	0,283	0,281
	$mef f / \text{Var}(y)$		0,242	0,252
	$(mef f)^{-1}$		0,241	0,231

Consistent with classic statistical theory, stratification appeared to reduce the standard errors whereas clustering increased the standard errors for all the superpopulation distributions considered.

Chapter 7 (Application)

This application aims to find a more accurate estimate of the average age of the household heads by combining estimates calculated from two surveys, and therefore to apply methods discussed in this Thesis. In South Africa, there appears to be an increase in the number of child-headed households because of the HIV/AIDS epidemic. These children might require special social support systems since they are unlikely to have developed enough skills to support other minors left in their care. A study in Iran found a significant association between the use of health services and the age of the head of the household.^[56] Another study in Kenya showed that there was a significant association between the use of improved drinking water sources and the age of the household head.^[57] It is therefore important to know the average age of household heads in the population.

Two surveys were analysed and combined using the selected methods studied in this thesis, that is, the Community Survey of 2016 and the General Household survey of 2016. The Community Survey was used to illustrate the use of the studied weighting methods under simple random sampling. The weighting methods under complex sampling were illustrated using the General Household survey. The two surveys are large-scale national surveys. A brief introduction of each survey is given below.

7.1 Community Survey 2016

The aim of the community survey is to provide reliable demographic and socio-economic data at local municipality level in South Africa between censuses. It aims to estimate the population count by local municipality, the household count by local municipality and measure demographic factors such as fertility, mortality and migration.^[85] The survey also measures socio-economic factors, e.g. unemployment and poverty, and access to facilities and services, e.g. piped water, sanitation and electricity for lighting is also assessed. Reliable information is required to promote optimal allocation and use of resources in all spheres of government to reduce poverty and vulnerability among all South Africans. The target population in this survey is the non-institutional population residing in private dwellings, in South

Africa. It excludes homeless people and people in hospitals, prisons, military barracks, non-residential hotels, frail care centres, orphanages, boarding school hostels, initiation schools, religious houses, ships in harbour and refugee camps. The sampling frame used in this survey was a geo-referenced dwelling frame updated from the 2011 census and very small EAs were excluded. A stratified single-stage sample design with systematic random sampling was used, selecting eight percent of dwelling units in an enumeration area with adjustments to very small and very large EAs.^{[85][86]} However, for operational feasibility, a lower limit of five and an upper limit of sixty six dwelling units were placed on the sample size for each enumeration area. Computer assisted personal interviewing was used to enhance data quality. A total of 1 370 809 households were sampled, with a 90.52 % response rate. After editing and quality checks, a sample of size 984 627 remained.

7.2 General Household Survey 2016

The General Household Survey was established to determine speed of progress in South Africa.^[87] It measures quality of service delivery and performance of programmes in the country. Its target population is people residing in private households in South Africa, including people in workers hostels. The survey excludes homeless people and people in hospitals, prisons, military barracks, non-residential hotels, frail care centres, orphanage, boarding school hostels, initiation schools, religious houses, ships in harbour and refugee camps. After editing and quality checks, a sample of size 21 218 households remained. The sampling frame was based on information from the 2011 census. A stratified two-stage sampling design was used with stratification based on geographic type. The EAs were sampled with probability proportional to size (PPS) and dwelling units were then sampled using systematic sampling

7.3 Simple random sampling application

To illustrate the weighting strategies studied through simulation, two simple random samples were selected from the Community Survey data, considered here as the target population. The mean of the age of the head of the household was estimated from each of the sample surveys selected. These were then combined using each of the studied weighting methods in this thesis, to obtain a combined estimate of the mean age of the household age in the Community Survey. That is, the inverse of the sample size, the inverse of variance, the inverse of the coefficient of variation (equation 3.1), the inverse of skewness (equation 3.2), the inverse of kurtosis (equation 3.3) and the inverse of the D'Agostino-Pearson statistic (equation 3.4) were used to combine the two surveys. Confidence intervals for the combined mean estimate were constructed using each of the variance estimators presented in Chapter 5. That is, variance estimator 1 (the average variance estimator, equation 5.1), variance estimator 2 (the meta-analysis variance estimator, equation 5.2), variance estimator 4 (the weighted meta-analysis variance estimator, equation 5.3) and variance estimator 4 (the robust meta-analysis variance estimator, equation 5.4). The results are presented in table 7.1 below.

Table 7. 1: Combined estimates of mean age of the household head and their confidence intervals under simple random sampling

Sample size	Weighting Method	Population Mean (θ)	Combined Mean	Sampling Error	LCL1	UCL1	LCL2	UCL2	LCL3	UCL3	LCL4	UCL4
500	$\frac{1}{n}$	47.4847744	47.7460000	-0.261225563	47.023935139	48.46806486	47.684019358	47.80798064	46.33594858	49.15605142	46.74894308	48.74305692
	$(Var(\hat{\theta}_d))^{-1}$	47.4847744	47.7054939	-0.220719507	46.983429083	48.42755881	46.709630819	48.70135707	46.29713081	49.11385708	46.70963082	48.70135707
	1	47.4847744	47.7460000	-0.261225563	47.023935139	48.46806486	46.360070709	49.13192929	46.33594858	49.15605142	46.74894308	48.74305692
	$(CV(\hat{\theta}_d))^{-1}$	47.4847744	47.7400913	-0.255316878	47.018026454	48.46215618	47.570009035	47.91017360	46.33028604	49.14989659	46.74335713	48.73682550
	$(skewness(sk_{dc}))^{-1}$	47.4847744	47.9811627	-0.496388234	47.259097810	48.70322753	47.130030014	48.83229533	46.56134941	49.40097593	46.93143108	49.03089427
	$(kurtosis(K_{dc}))^{-1}$	47.4847744	47.7881353	-0.303360894	47.066070470	48.51020019	45.559241920	50.01702874	46.37632987	49.19994079	46.78731240	48.78895826
	$(K^2)^{-1}$	47.4847744	48.0157848	-0.531010410	47.293719986	48.73784971	42.603077159	53.42849254	46.59454005	49.43702965	46.95213377	49.07943593
900	$\frac{1}{n}$	47.4847744	47.4822222	0.002552215	46.956386204	48.00805824	47.436024579	47.52841987	46.43878711	48.52565734	46.74440218	48.220042267
	$(Var(\hat{\theta}_d))^{-1}$	47.4847744	47.4743146	0.010459805	46.948478614	48.00015065	46.736587276	48.21204199	46.43101060	48.51761866	46.73658728	48.212041988
	1	47.4847744	47.4822222	0.002552215	46.956386204	48.00805824	46.096292931	48.86815151	46.43878711	48.52565734	46.74440218	48.220042267
	$(CV(\hat{\theta}_d))^{-1}$	47.4847744	47.4835100	0.001264397	46.957674022	48.00934606	47.336760122	47.63025996	46.44005358	48.52696650	46.74565735	48.221362731
	$(skewness(sk_{dc}))^{-1}$	47.4847744	47.4457991	0.038975367	46.919963052	47.97163509	46.504573646	48.38702449	46.40296787	48.48863027	46.70686730	48.184730843
	$(kurtosis(K_{dc}))^{-1}$	47.4847744	47.4549631	0.029811348	46.929127071	47.98079911	45.187643293	49.72228289	46.41197991	48.49794627	46.71668081	48.193245371
	$(K^2)^{-1}$	47.4847744	47.4585710	0.026200463	46.932737956	47.98440999	39.463434784	55.45371316	46.41553092	48.50161703	46.72047942	48.196668532
1200	$\frac{1}{n}$	47.4847744	47.3654167	0.119357770	46.908524097	47.82230924	47.325408335	47.40542410	46.45460576	48.27622757	46.72137610	48.00945724
	$(Var(\hat{\theta}_d))^{-1}$	47.4847744	47.3651741	0.119600355	46.908281512	47.82206665	46.721147294	48.00920087	46.45438266	48.27596550	46.72114729	48.00920087
	1	47.4847744	47.3654167	0.119357770	46.908524097	47.82230924	45.979487376	48.75134596	46.45460576	48.27622757	46.72137610	48.00945724
	$(CV(\hat{\theta}_d))^{-1}$	47.4847744	47.3653244	0.119450031	46.908431836	47.82221698	47.228048746	47.50260007	46.45452091	48.27612790	46.72129233	48.00935649
	$(skewness(sk_{dc}))^{-1}$	47.4847744	47.3689815	0.115792913	46.912088954	47.82587409	46.483377897	48.25458515	46.45788428	48.28007877	46.72156893	48.01639411
	$(kurtosis(K_{dc}))^{-1}$	47.4847744	47.3658079	0.118966494	46.908915373	47.82270051	45.168035525	49.56358036	46.45496560	48.27665028	46.72168708	48.00992881
	$(K^2)^{-1}$	47.4847744	47.3699787	0.114795746	46.913086121	47.82687126	38.638963563	56.10099382	46.45880137	48.28115602	46.72056871	48.01938867

$\hat{\theta}_d$ =point estimate calculated from survey d, LC1= lower confidence limit calculated using variance estimator 1, UC1=upper confidence limit calculated using variance estimator 1, LC2= lower confidence limit calculated using variance estimator 2, UC2=upper confidence limit calculated using variance estimator 2, LC3= lower confidence limit calculated using variance estimator 3, UC3=upper confidence limit calculated using variance estimator 3, LC4= lower confidence limit calculated using variance estimator 4, UC4=upper confidence limit calculated using variance estimator 4

Table 7.1: Combined estimates of mean age of the household head and their confidence intervals under simple random sampling (cont'd)

Sample size	Weighting Method	Population Mean (θ)	Combined Mean	Sampling Error	LCL1	UCL1	LCL2	UCL2	LCL3	UCL3	LCL4	UCL4
2000	$\frac{1}{n}$	47.4847744	47.4557500	0.029024437	47.098052659	47.81344734	47.424759679	47.48674032	46.74249231	48.16900769	46.95140065	47.96009935
	$(\text{Var}(\hat{\theta}_d))^{-1}$	47.4847744	47.4537231	0.031051336	47.096025760	47.81142044	46.949382834	47.95806337	46.74047826	48.16696795	46.94938284	47.95806337
	1	47.4847744	47.4557500	0.029024437	47.098052659	47.81344734	46.069820709	48.84167929	46.74249231	48.16900769	46.95140065	47.96009935
	$(\text{CV}(\hat{\theta}_d))^{-1}$	47.4847744	47.4571404	0.027634046	47.099443050	47.81483773	47.335775365	47.57850542	46.74387389	48.17040689	46.95277431	47.96150647
	$(\text{skewness}(sk_{dc}))^{-1}$	47.4847744	47.4801098	0.004664599	47.122412497	47.83780718	46.622862168	48.33735751	46.76669781	48.19352187	46.97423273	47.98598695
	$(\text{kurtosis}(K_{dc}))^{-1}$	47.4847744	47.4547329	0.030041526	47.097035570	47.81243025	45.265601375	49.64386445	46.74148167	48.167984155	46.95039039	47.95907543
	$(K^2)^{-1}$	47.4847744	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
5000	$\frac{1}{n}$	47.4847744	47.4440000	0.040774437	47.218621653	47.66937835	47.424400000	47.46360000	46.99353560	47.89446440	47.12547357	47.76252644
	$(\text{Var}(\hat{\theta}_d))^{-1}$	47.4847744	47.4441978	0.040576599	47.218819491	47.66957619	47.125671672	47.76272400	46.99373381	47.89466186	47.12567167	47.76272400
	1	47.4847744	47.4440000	0.040774437	47.218621653	47.66937835	46.058070709	48.82992929	46.99353560	47.89446440	47.12547357	47.76252644
	$(\text{CV}(\hat{\theta}_d))^{-1}$	47.4847744	47.4445885	0.040185978	47.219210112	47.66996681	47.348127397	47.54104953	46.99412518	47.89505173	47.12606125	47.76311567
	$(\text{skewness}(sk_{dc}))^{-1}$	47.4847744	47.4379072	0.046867216	47.212528874	47.66328557	46.585141984	48.29067245	46.98743113	47.88838331	47.11910982	47.75670462
	$(\text{kurtosis}(K_{dc}))^{-1}$	47.4847744	47.4440587	0.040715760	47.218680330	47.66943702	45.246781592	49.64133576	46.99359439	47.89452297	47.12553238	47.76258498
	$(K^2)^{-1}$	47.4847744	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

$\hat{\theta}_d$ =point estimate calculated from survey d, LC1= lower confidence limit calculated using variance estimator 1, UC1=upper confidence limit calculated using variance estimator 1, LC2= lower confidence limit calculated using variance estimator 2, UC2=upper confidence limit calculated using variance estimator 2, LC3= lower confidence limit calculated using variance estimator 3, UC3=upper confidence limit calculated using variance estimator 3, LC4= lower confidence limit calculated using variance estimator 4, UC4=upper confidence limit calculated using variance estimator 4,

The combined mean estimates for age of the head of the household, resulting from all the weighting methods under consideration were very close to the population mean. The D’Agostino-Pearson Statistic could not be calculated for sample sizes 2000 and 5000. This may be an artefact of the r-code used in the simulations. The combined estimate was not very close to the population mean for the smaller sample sizes considered. Sampling errors became smaller as the sample size was increased from 500 to 5000, which is expected. 95% confidence intervals for the combined estimate were calculated using variance estimates obtained using the variance estimators assessed in chapter 6. Most of the 95% confidence intervals included the population mean value. An exception

was observed for the sample size weighting method when used with the meta-analysis variance estimator. The population mean value was either under-estimated, or just included in the 95% confidence interval. Variance estimates resulting from combining the two simple random samples and using the variance estimators under consideration in this study were compared to the variance estimate obtained from one random sample. The results are presented in table 7.2 below.

Table 7. 2: Comparison of standard errors from one sample to standard errors resulting from combined samples

Sample size	Weighting method	Variance estimator				One sample
		1	2	3	4	
500	$\frac{1}{n}$	0.3684	0.031623	0.719414	0.508703	0.736801
	$(Var(\hat{\theta}_d))^{-1}$	0.3684	0.508093	0.718553	0.508093	0.736801
	1	0.3684	0.707107	0.719414	0.508703	0.736801
	$(CV(\hat{\theta}_d))^{-1}$	0.3684	0.086777	0.719288	0.508538	0.736801
	$(skewness(sk_{dc}))^{-1}$	0.3684	0.434251	0.724395	0.535577	0.736801
	$(kurtosis(K_{dc}))^{-1}$	0.3684	1.137191	0.720309	0.510624	0.736801
900	$\frac{1}{n}$	0.268284	0.02357	0.532365	0.376439	0.536567
	$(Var(\hat{\theta}_d))^{-1}$	0.268284	0.376392	0.532298	0.376392	0.536567
	1	0.268284	0.707107	0.532365	0.376439	0.536567
	$(CV(\hat{\theta}_d))^{-1}$	0.268284	0.074872	0.532376	0.376455	0.536567
	$(skewness(sk_{dc}))^{-1}$	0.268284	0.480217	0.532057	0.377006	0.536567
	$(kurtosis(K_{dc}))^{-1}$	0.268284	1.156796	0.532134	0.376675	0.536567
1200	$\frac{1}{n}$	0.233108	0.020412	0.464699	0.328592	0.466217
	$(Var(\hat{\theta}_d))^{-1}$	0.233108	0.328585	0.464689	0.328585	0.466217
	1	0.233108	0.707107	0.464699	0.328592	0.466217
	$(CV(\hat{\theta}_d))^{-1}$	0.233108	0.070039	0.464696	0.328588	0.466217
	$(skewness(sk_{dc}))^{-1}$	0.233108	0.451839	0.464846	0.330313	0.466217
	$(kurtosis(K_{dc}))^{-1}$	0.233108	1.121312	0.464715	0.328633	0.466217
2000	$\frac{1}{n}$	0.182499	0.015811	0.363907	0.257321	0.364997
	$(Var(\hat{\theta}_d))^{-1}$	0.182499	0.257316	0.3639	0.257316	0.364997
	1	0.182499	0.707107	0.363907	0.257321	0.364997
	$(CV(\hat{\theta}_d))^{-1}$	0.182499	0.061921	0.363911	0.25733	0.364997
	$(skewness(sk_{dc}))^{-1}$	0.182499	0.437371	0.363986	0.258101	0.364997
	$(kurtosis(K_{dc}))^{-1}$	0.182499	1.116904	0.363904	0.257318	0.364997
5000	$\frac{1}{n}$	0.114989	0.01	0.229829	0.162513	0.229978
	$(Var(\hat{\theta}_d))^{-1}$	0.114989	0.162513	0.229829	0.162513	0.229978
	1	0.114989	0.707107	0.229829	0.162513	0.229978
	$(CV(\hat{\theta}_d))^{-1}$	0.114989	0.049215	0.229828	0.162514	0.229978
	$(skewness(sk_{dc}))^{-1}$	0.114989	0.435084	0.229835	0.162652	0.229978
	$(kurtosis(K_{dc}))^{-1}$	0.114989	1.12106	0.229829	0.162513	0.229978
	$(K^2)^{-1}$	0.114989	NA	NA	NA	0.229978

$\hat{\theta}_d$ = estimate calculated from each individual survey, K^2 = D'Agostino Pearson statistic, 1=average variance estimator, 2=meta-analysis variance estimator, 3= weighted meta-analysis variance estimator, 4=robust meta-analysis variance estimator.

The results of the comparison show that combining surveys leads to smaller standard errors for most of the weighting methods and variance estimators compared to using only one survey, across all the sample sizes considered. In this example, only two samples were combined, but the improvement in the standard errors is quite substantial. Standard errors calculated using the meta-analysis variance estimator, when kurtosis and the D'Agostino-Pearson statistic were used to weight samples were generally larger than the standard errors obtained from one sample. When the sample size was greater than 1200, the standard errors calculated using the meta-analysis variance estimator were large when no weight was employed.

7.4 Stratified random sampling application

The General household survey was used to illustrate combining surveys where stratified random sampling has been used to collect the data. As with the simple random sampling case, two stratified samples were selected from the General Household survey data and the mean age of the household head was estimated from each sample. The two estimates were then combined using each of the weighting methods evaluated using simulation in chapter 4. That is, the inverse of the sample size, the inverse of variance, the inverse of the coefficient of variation (equation 3.1), the inverse of skewness (equation 3.2), the inverse of kurtosis (equation 3.3), the inverse of the D'Agostino-Pearson statistic (equation 3.4), the inverse of misspecification effect (equation 3.9) and the ratio of misspecification effect to the variance (equation 3.10), were used to combine the two surveys. Again, 95% confidence intervals were calculated using the variance estimates obtained from the variance estimators studied in this thesis, equations 5.1 to 5.4. The results are presented in table 7.3 below.

Table 7. 3: Combined estimates of mean age of household head and confidence intervals under stratified random sampling

Sample size	Weighting method	Population mean	Combined mean	Sampling error	LCL1	UCL1	LCL2	UCL2	LCL3	UCL3	LCL4	UCL4
500	$\frac{1}{n}$	47.822038	47.801339	0.020698	47.61712130479	47.985557687	47.4746728	48.128006163	47.536245944	48.066433048	47.613890048	47.98878894
	$(\text{Var}(\hat{\theta}_d))^{-1}$	47.822038	48.622170	-0.800132	48.43795201379	48.806388396	48.5540554	48.690285002	48.525841336	48.718499074	48.554055408	48.69028500
	1	47.822038	47.787486	0.0345522	47.60326753179	47.971703914	25.6295938	69.945377567	47.520457531	48.054513915	47.597287282	47.97768416
	$(\text{CV}(\hat{\theta}_d))^{-1}$	47.822038	47.801340	0.0206984	47.61712130479	47.985557687	46.4154102	49.187268787	47.536245944	48.066433048	47.613890048	47.98878894
	$(\text{skewness}(sk_{dc}))^{-1}$	47.822038	47.816288	0.0057503	47.63206941379	48.000505796	47.6370997	47.995475435	47.553297476	48.079277734	47.631797424	48.00077779
	$(\text{kurtosis}(K_{dc}))^{-1}$	47.822038	47.813902	0.0081352	47.62968445279	47.998120835	47.0097360	48.618069254	47.550575789	48.077229499	47.628940806	47.99886448
	$meff/\text{Var}(y)$	47.822038	47.827041	-0.0050032	47.64282295079	48.011259333	46.9585311	48.695551097	47.565574662	48.088507622	47.644675231	48.00940705
$(meff)^{-1}$	47.822038	47.844182	-0.0221443	47.65996396279	48.028400345	46.9111242	48.777240043	47.585162880	48.103201428	47.665193966	48.02317034	
900	$\frac{1}{n}$	47.822038	47.827582	-0.0055444	47.64336413279	48.011800515	45.5637309	50.091433694	47.566192757	48.088971891	47.645323214	48.00984143
	$(\text{Var}(\hat{\theta}_d))^{-1}$	47.822038	47.829038	-0.0070000	47.64481978479	48.013256167	45.5169991	50.141076830	47.567855402	48.090220550	47.647066086	48.01100987
	1	47.822038	48.620664	-0.7986261	48.43644580879	48.804882191	47.2495367	49.991791261	48.523755920	48.717572080	48.552548450	48.68877955
	$(\text{CV}(\hat{\theta}_d))^{-1}$	47.822038	46.981667	0.84037128	46.79744842379	47.165884806	46.5706071	47.392726091	46.619473443	47.343859787	46.626056868	47.33727636
	$(\text{skewness}(sk_{dc}))^{-1}$	47.822038	48.146955	-0.32491682	48.01510650344	48.278802912	47.8202880	48.473621375	47.958642665	48.335266751	48.013797985	48.28011143
	$(\text{kurtosis}(K_{dc}))^{-1}$	47.822038	48.248922	-0.42688421	48.11707390244	48.380770311	48.2120449	48.285799310	48.196769867	48.301074347	48.212044904	48.28579931
	$meff/\text{Var}(y)$	47.822038	48.143979	-0.32194147	48.01213116644	48.275827575	25.8928076	70.395151075	47.953147504	48.334811238	48.007231327	48.28072741
$(meff)^{-1}$	47.822038	48.146955	-0.32491681	48.01510650344	48.278802912	46.7610254	49.532883999	47.958642665	48.335266751	48.013797985	48.28011143	
1200	$\frac{1}{n}$	47.822038	48.146131	-0.32409330	48.01428299144	48.277979400	47.9944441	48.297818337	47.957118355	48.335144037	48.011981247	48.28028115
	$(\text{Var}(\hat{\theta}_d))^{-1}$	47.822038	48.145958	-0.32392023	48.01410991944	48.277806328	47.3432209	48.948695331	47.956798331	48.335117917	48.011599358	48.28031689
	1	47.822038	48.144441	-0.32240348	48.01259317244	48.276289581	47.2825140	49.006368762	47.953998598	48.334884156	48.008251490	48.28063126
	$(\text{CV}(\hat{\theta}_d))^{-1}$	47.822038	48.136978	-0.31494055	48.00513023344	48.268826642	47.2598759	49.014080959	47.940344794	48.333612082	47.991751780	48.28220510
	$(\text{skewness}(sk_{dc}))^{-1}$	47.822038	48.140121	-0.31808350	48.00827322844	48.271969637	45.9338494	50.34639342	47.946070977	48.334171889	47.998705781	48.28153708
	$(\text{kurtosis}(K_{dc}))^{-1}$	47.822038	48.137952	-0.31591420	48.00610389244	48.269800301	45.9254086	50.350495629	47.942115048	48.333789146	47.993906803	48.28199739
	$meff/\text{Var}(y)$	47.822038	48.248855	-0.42681764	48.11700732544	48.380703734	47.2181090	49.279602027	48.196498735	48.301212325	48.211978233	48.28573283
$(meff)^{-1}$	47.822038	48.044892	-0.22285435	47.91304404244	48.176740451	47.7477610	48.342023526	47.783676781	48.306107713	47.786233919	48.30355057	

$\hat{\theta}_d$ =point estimate from survey d, LC1= lower confidence limit calculated using variance estimator 1, UC1=upper confidence limit calculated using variance estimator 1, LC2= lower confidence limit calculated using variance estimator 2, UC2=upper confidence limit calculated using variance estimator 2, LC3= lower confidence limit calculated using variance estimator 3, UC3=upper confidence limit calculated using variance estimator 3, LC4= lower confidence limit calculated using variance estimator 4, UC4=upper confidence limit calculated using variance estimator 4

Table 7.3: Combined estimates of mean age of household head and confidence intervals: stratified random sampling (cont'd) ()

Sample size	Weighting Method	Population mean	Combined mean	Sampling error	LCL1	UCL1	LCL2	UCL2	LCL3	UCL3	LCL4	UCL4
2000	$\frac{1}{n}$	47.822038	47.844849	-0.02281117	47.73636344994	47.95333468	47.5181824	48.171515732	47.690229041	47.999469089	47.735516197	47.95418193
	$(Var(\hat{\theta}_d))^{-1}$	47.822038	47.806913	0.01512462	47.69842765594	47.915398886	47.7799542	47.833872360	47.768787362	47.845039180	47.779954182	47.83387236
	1	47.822038	47.845151	-0.02311289	47.73666516994	47.953636400	25.8493209	69.840980696	47.689954341	48.000347229	47.735001021	47.95530055
	$(CV(\hat{\theta}_d))^{-1}$	47.822038	47.844849	-0.02281117	47.73636344994	47.953334680	46.4589198	49.230778356	47.690229041	47.999469089	47.735516197	47.95418193
	$(skewness(sk_{dc}))^{-1}$	47.822038	47.845877	-0.02383865	47.73739092594	47.954362156	47.7067995	47.984953580	47.689302267	48.002450815	47.733760248	47.95799283
	$(kurtosis(K_{dc}))^{-1}$	47.822038	47.845140	-0.02310176	47.73665404094	47.953625271	47.0460040	48.644275362	47.689964436	48.000314876	47.735020031	47.95525928
	$meff/Var(y)$	47.822038	47.845171	-0.02313345	47.73668572794	47.953656958	46.9871964	48.703146288	47.689935701	48.000406985	47.734965903	47.95537678
	$(meff)^{-1}$	47.822038	47.848351	-0.02631357	47.73986584394	47.956837074	46.9693014	48.727401556	47.687167156	48.009535762	47.729513759	47.96718916
5000	$\frac{1}{n}$	47.822038	47.844774	-0.02273629	47.73628856294	47.953259793	45.5915146	50.098033735	47.690297554	47.999250802	47.735644003	47.95390435
	$(Var(\hat{\theta}_d))^{-1}$	47.822038	47.845675	-0.02363732	47.73718959794	47.954160828	45.5963331	50.095017294	47.689481937	48.001868490	47.734104661	47.95724576
	1	47.822038	47.806815	0.01522317	47.69832910794	47.915300338	46.9374502	48.676179287	47.769461606	47.844167840	47.779854229	47.83377522
	$(CV(\hat{\theta}_d))^{-1}$	47.822038	47.882989	-0.06095153	47.77450380794	47.991475038	47.6373443	48.128634589	47.667392119	48.098586727	47.668800697	48.09717815
	$(skewness(sk_{dc}))^{-1}$	47.822038	48.138949	-0.31691152	48.05002496049	48.227873872	47.8122827	48.465616083	48.012694338	48.265204494	48.049673594	48.22822524
	$(kurtosis(K_{dc}))^{-1}$	47.822038	47.917384	-0.09534606	47.82845949949	48.006308411	47.9016205	47.933147410	47.895091063	47.939676847	47.901620500	47.93314741
	$meff/Var(y)$	47.822038	48.137866	-0.31582814	48.04894158449	48.226790496	25.8842464	70.391485647	48.011910365	48.263821715	48.049013103	48.22671898
	$(meff)^{-1}$	47.822038	48.138949	-0.31691152	48.05002496049	48.227873872	46.7530201	49.524878707	48.012694338	48.265204493	48.049673594	48.22822523

LC1= lower confidence limit calculated using variance estimator 1, UC1=upper confidence limit calculated using variance estimator 1, LC2= lower confidence limit calculated using variance estimator 2, UC2=upper confidence limit calculated using variance estimator 2, LC3= lower confidence limit calculated using variance estimator 3, UC3=upper confidence limit calculated using variance estimator 3, LC4= lower confidence limit calculated using variance estimator 4, UC4=upper confidence limit calculated using variance estimator 4

At sample size n=500, all the weighting strategies apart from the inverse of variance resulted in point estimates close to the population parameter value, with sampling errors less than 4%. All the variance estimators appeared to work well with all the weighting strategies, except the inverse of variance which resulted in 95% confidence intervals which did not include the population

parameter value. Only the inverse of the sample size and the inverse of variance resulted in point estimates close to the population parameter value at sample size 900. All the variance estimators resulted in confidence intervals including the population mean for these weighting strategies. The meta-analysis variance estimator (variance estimator 2) resulted in a wider confidence interval compared to the other variance estimators. Additionally, the meta-analysis variance estimator appeared to work well with a weight of 1, the inverse of skewness, the inverse of misspecification error and $m_{eff}/Var(y)$. At sample size $n=1200$, all the variance estimators worked well with the inverse of misspecification error. The meta-analysis variance estimator worked with all weighting strategies except the inverse of sample size. All the other confidence intervals excluded the population mean. When the sample size was increased to 2000, all the point estimates were close to the population parameter value and all the confidence intervals included the population mean. All the variance estimators resulted in confidence intervals which included the population mean with the inverse of sample size, weight of 1, inverse of variance and inverse of coefficient of variation when the sample size was increased to 5000. The meta-analysis variance estimator worked well with all the weighting strategies except the inverse of kurtosis. The confidence intervals were very wide for the inverse of sample size, inverse of variance and $m_{eff}/Var(y)$ weighting strategies.

Variance estimates, resulting from combining the two stratified random samples and using the variance estimators under consideration in this study, were compared to the variance estimate obtained from one stratified random sample. The results are presented in table 7.4 below.

Table 7. 4: Comparison of standard errors from one sample to standard errors resulting from combined samples under stratified random sampling

Sample size	Weighting method	Variance Estimator				One sample
		1	2	3	4	
500	$\frac{1}{n}$	0.0939888730671814	0.1666666666666667	0.135251812152932	0.0956374735411073	0.187977746134363
	$(Var(\theta_a))^{-1}$	0.0939888730671814	0.0347524472161688	0.0491473821787611	0.0347524472161688	0.187977746134363
	1	0.0939888730671814	0.707106781186548	0.135251812152932	0.0956374735411073	0.187977746134363
	$(CV(\theta_a))^{-1}$	0.0939888730671814	0.0914223620067288	0.134178637310581	0.0941276430011106	0.187977746134363
	$(skewness(sk_{dc}))^{-1}$	0.0939888730671814	0.44311732400753	0.133401265537659	0.0930438321374779	0.187977746134363
	$(kurtosis(K_{dc}))^{-1}$	0.0939888730671814	1.15502620901315	0.133362023749108	0.0929893416524842	0.187977746134363
	$meff/Var(y)$	0.0939888730671814	0.699554725083475	0.0494428980218796	0.0347528318056822	0.187977746134363
	$(meff)^{-1}$	0.0939888730671814	0.209724222195848	0.184792434860796	0.181433544212019	0.187977746134363
900	$\frac{1}{n}$	0.0672694921203861	0.1666666666666667	0.0968592905254623	0.0684898611514723	0.134538984240772
	$(Var(\theta_a))^{-1}$	0.0672694921203861	0.0252843811210805	0.0357575146976423	0.0252843811210805	0.134538984240772
	1	0.0672694921203861	0.707106781186548	0.0968592905254623	0.0684898611514723	0.134538984240772
	$(CV(\theta_a))^{-1}$	0.0672694921203861	0.0773913986149171	0.097207945157778	0.0689840997452107	0.134538984240772
	$(skewness(sk_{dc}))^{-1}$	0.0672694921203861	0.439758869860557	0.0979194857898231	0.0699999709307912	0.134538984240772
	$(kurtosis(K_{dc}))^{-1}$	0.0672694921203861	1.12564897288303	0.0997154276212395	0.0726068587458853	0.134538984240772
	$meff/Var(y)$	0.0672694921203861	0.70662701982821	0.0358931376233206	0.0252844963445139	0.134538984240772
	$(meff)^{-1}$	0.0672694921203861	0.150401790749508	0.132282759157758	0.129897360410511	0.134538984240772
1200	$\frac{1}{n}$	0.055349803599578	0.1666666666666667	0.0797862947516343	0.0564174300646293	0.110699607199156
	$(Var(\theta_a))^{-1}$	0.055349803599578	0.0214338258132156	0.0303120071585921	0.0214338258132156	0.110699607199156
	1	0.055349803599578	0.707106781186548	0.0797862947516343	0.0564174300646293	0.110699607199156
	$(CV(\theta_a))^{-1}$	0.055349803599578	0.0709576725428504	0.0807491013572492	0.0577899983485213	0.110699607199156
	$(skewness(sk_{dc}))^{-1}$	0.055349803599578	0.4377423189399	0.0800895345777931	0.0568473647794136	0.110699607199156
	$(kurtosis(K_{dc}))^{-1}$	0.055349803599578	1.14962222278949	0.0797156667344986	0.0563176061016747	0.110699607199156
	$meff/Var(y)$	0.055349803599578	0.691820325439433	0.0297247517622134	0.0214364874064751	0.110699607199156
	$(meff)^{-1}$	0.055349803599578	0.124153278269877	0.109023930952586	0.10724099897998	0.110699607199156

Table 7.4: Comparison of standard errors from one sample to standard errors resulting from combined samples under stratified random sampling (cont'd)

Sample size	Weighting method	Variance Estimator				One sample
		1	2	3	4	
2000	$\frac{1}{n}$	0.0453696201580241	0.1666666666666667	0.0652485447321587	0.0461376884426632	0.0907392403160482
	$(\text{Var}(\hat{\theta}_d))^{-1}$	0.0453696201580241	0.0164879671774472	0.0233175067983083	0.0164879671774472	0.0907392403160482
	1	0.0453696201580241	0.707106781186548	0.0652485447321587	0.0461376884426632	0.0907392403160482
	$(\text{CV}(\hat{\theta}_d))^{-1}$	0.0453696201580241	0.0630118774413785	0.0651631421160472	0.0460170131212809	0.0907392403160482
	$(\text{skewness}(sk_{dc}))^{-1}$	0.0453696201580241	0.441668876386013	0.0644823293613776	0.0450623437382922	0.0907392403160482
	$(\text{kurtosis}(K_{dc}))^{-1}$	0.0453696201580241	1.12118910125204	0.0649454642551034	0.0457103562453551	0.0907392403160482
	$\text{meff}/\text{Var}(y)$	0.0453696201580241	0.695850828168822	0.0234287225650247	0.0164880775330626	0.0907392403160482
$(\text{meff})^{-1}$	0.0453696201580241	0.0997860000131511	0.0892075799109878	0.08759632085413	0.0907392403160482	
5000	$\frac{1}{n}$	0.0272931703527966	0.1666666666666667	0.0392311110391019	0.0277405846492314	0.0545863407055933
	$(\text{Var}(\hat{\theta}_d))^{-1}$	0.0272931703527966	0.00976422713438375	0.0138087024395369	0.00976422713438375	0.0545863407055933
	1	0.0272931703527966	0.707106781186548	0.0392311110391019	0.0277405846492314	0.0545863407055933
	$(\text{CV}(\hat{\theta}_d))^{-1}$	0.0272931703527966	0.048779987361175	0.0390065397545479	0.027424160458614	0.0545863407055933
	$(\text{skewness}(sk_{dc}))^{-1}$	0.0272931703527966	0.44773933854235	0.0391173179981151	0.0275799565803125	0.0545863407055933
	$(\text{kurtosis}(K_{dc}))^{-1}$	0.0272931703527966	1.13177993464756	0.0394872489766225	0.0281043331701814	0.0545863407055933
	$\text{meff}/\text{Var}(y)$	0.0272931703527966	0.693835834665335	0.0139157014492362	0.00976439422947314	0.0545863407055933
$(\text{meff})^{-1}$	0.0272931703527966	0.0596108579831122	0.0536526344185803	0.052672759088903	0.0545863407055933	

$\hat{\theta}_d$ = estimate calculated from each individual survey, CV=coefficient of variation, 1=average variance estimator, 2=meta-analysis variance estimator, 3= weighted meta-analysis variance estimator, 4=robust meta-analysis variance estimator.

When a sample size of 500 was used, combining sample surveys resulted in standard error estimates which were mostly smaller than the standard error estimate obtained using one sample. The weighted meta-analysis variance estimator resulted in larger standard error estimates when all the weighting strategies were used to weight the surveys except the inverse of sample size and the inverse of the variance. As the sample size increased, the weighted meta-analysis and robust meta-analysis variance estimators seemed to work better with standard error estimates mostly smaller than the standard errors obtained using one sample.

7.5 Conclusion

For this application, the inverses of sample size, kurtosis and the D'Agostino-Pearson statistic, would not be recommended as weighting strategies under simple random sampling. Large meta-analysis variance estimator standard errors which improved with increasing sample size were observed when the inverse of sample size was used as the weighting strategy. The inverses of kurtosis and the D'Agostino-Pearson statistic resulted in large standard errors for all the variance estimators studied. Under stratified random sampling, the combination of the inverse sample size weighting strategy and the meta-analysis variance estimator resulted in confidence intervals which either excluded the population mean or were very wide. For the most part, combining the surveys resulted in improved standard errors, with the weighted and robust meta-analysis variance estimators. Even though the average variance estimator did not work well with most of the weighting strategies, it yielded standard errors which were smaller than the standard errors obtained from one sample.

Chapter 8

Conclusion

In this thesis, we initially replicated Fox's study comparing the convergence properties of combined estimates using three different weighting strategies, that is a weight of 1, the inverse of sample size and the inverse of variance. Our results agreed with the results Fox obtained. The combined mean estimate converged towards the finite population mean, and not the superpopulation mean. We investigated the use of different weighting strategies to combine surveys under different sampling strategies, that is, simple random sampling, stratified random sampling and cluster random sampling. The variance, skewness and kurtosis of the superpopulation were also investigated for influence on the choice of the weighting strategy when calculating a combined point estimate. Different finite population sizes were used in the simulations.

Comparing the MSE across population sizes for the same sample size, under simple random sampling, showed that the variance of the superpopulation distribution, did not influence the choice of the weighting strategy. However, increasing the variance of the superpopulation distribution increased MSE values for all the weighting strategies. This agrees with classical statistical theory. Increasing the number of samples combined also decreased the magnitude of the MSE values. The skewness of the superpopulation appeared to affect the choice of the weighting strategy. In general, the symmetric superpopulations appeared to work well with all the weighting strategies except the inverse of skewness and the inverse of the D'Agostino Pearson statistic. This is expected as skewness is not expected to play an important role when the data are symmetrical. The skewed superpopulations appeared to work best with the inverse of sample size for all the sample sizes considered.

The variance of the superpopulation distribution was not investigated for influence on the choice of the weighting strategy under stratified random sampling because we had already seen its lack of influence under simple random sampling. The superpopulation appeared to influence the choice of the weighting strategy. Combining estimates yielded better estimates with lower MSE values compared to estimates from one sample regardless of weighting strategy. MSE values improved with increasing sample size and number of samples combined. All the weighting strategies worked well with the symmetric superpopulation distributions, except the inverse of skewness. This is expected of a symmetric superpopulation as skewness is not expected to play a role. The inverse of Kurtosis did not work well with the superpopulation distributed as the t-distribution with three degrees of freedom. This was not expected and requires further investigation with larger sample sizes.

Similar to simple random sampling and stratified sampling, the inverse of skewness did not work well with the symmetric superpopulation distributions under cluster random sampling. The inverse of sample size performed best for the t distribution superpopulations. All the other weighting strategies performed similarly across the superpopulation distributions. Combining samples appeared to improve estimates, with decreasing MSE values as the number of samples combined was increased.

The meta-analysis variance estimator performed best when the inverse of variance was used to weight the surveys under simple random sampling. The same observation was made for all numbers of samples combined and all the sample sizes investigated. This is an ideal result because in meta-analysis, using the inverse of the variance to weight studies results in the minimum variance for the combined effect size estimate. The weighted meta-analysis variance estimator and the robust meta-analysis variance estimators worked well with all the weighting strategies when ten or more samples were combined, with the exception of the inverse of skewness and the inverse of the D'Agostino-Pearson statistic. In fact, none of the variance estimators worked well with

these two weighting strategies. In general, the average variance estimator did not work well with all the superpopulation distributions and would not be recommended for use under simple random sampling and stratified random sampling.

Simulation results for variance estimators under stratified random sampling and cluster random sampling were not in line with our expectations. This was also the case with simulation results for variance estimators assuming non-normal superpopulation distributions. Most of the relative differences observed between the variance obtained by simulation and the variances calculated using the studied variance estimators were unacceptably high. We have decided not to present these results in this thesis and defer them for further studies. Instead, empirical variances were used to provide a measure of precision for the point estimators.

Under stratified random sampling and cluster random sampling, design consistent weighting strategies resulted in MSE values different from the naïve weighting strategies. We recommend that design consistent weighting strategies be used to yield more accurate combined estimates.

We applied the studied weighting strategies and variance estimators to two South African national surveys and concluded that combining the surveys resulted in improved estimates and smaller standard errors than would be observed when one survey is used, under both simple and stratified random sampling.

Under complex sampling, our simulations generated only one finite populations from where the sample samples to be combined were selected. We would like to continue this study with different finite populations being generated. Future studies may also include replication methods of variance estimation like the bootstrap and jackknife variance estimators.

References

1. Ehrlich R, Joubert G. *Epidemiology: A research manual for South Africa*. 3rd ed. Cape Town: ABA Press; 2014.
2. Rothman KJ, Greenland S, Lash T. *Modern epidemiology*. 3rd ed. Wolters Kluwer Health; 2008.
3. Dhir A, Chen S, Nieminen M. Computers & Education A repeat cross-sectional analysis of the psychometric properties of the Compulsive Internet Use Scale (CIUS) with adolescents from public and private schools. *Comput Educ* [Internet] 2015;86:172–81. Available from: <http://dx.doi.org/10.1016/j.compedu.2015.03.011>
4. Steel D, McLaren C. *Design and Analysis of Repeated Surveys* [Internet]. Wollongong: 2008. Available from: <http://ro.uow.edu.au/cssmwp/10/>
5. Sharot T. Attrition and Rotation in Panel Surveys. *J R Stat Soc Ser D (The Stat* [Internet] 1991;40(3):325–31. Available from: <http://www.jstor.org.ezproxy.puc.cl/stable/2348285>
6. Porta M. *A dictionary of Epidemiology*. 5th ed. Oxford University Press; 2008.
7. Niemeier JL, Yee, Debbie. *Advantages and disadvantages: Longitudinal vs Repeated cross-sectional surveys*. 1996.
8. Caplan LS, Lane DS, Grimson R. The use of cohort vs repeated cross-sectional sample survey data in monitoring changing breast cancer screening practices. *Prev Med (Baltim)* [Internet] 1995;24(6):553–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8610077>
9. Pahkinen R, Lehtonen E. *Practical methods for design and analysis of complex surveys*. 2nd ed. Chichester: John Wiley & Sons; 2004.
10. Elkasabi MA. *Weighting Procedures for Dual Frame Telephone Surveys : A Case Study in Egypt* | S ... Page 1 of 11 *Weighting Procedures for Dual Frame Telephone Surveys : A Case Study in Egypt Special issue Weighting Procedures for Dual Frame Telephone Surveys : A Case S. Surv Insights Methods from F* [Internet] 2015;1–11. Available from: <http://surveyinsights.org/?p=5291>
11. Graubard BI, Korn EL. *Analyzing Health Surveys for Cancer-Related*

- Objectives. *J Natl cancer Inst* 1999;19(12).
12. Korn EL, Graubard BI. Analysis of large Health Surveys: Accounting for sampling design. *R Stat Soc* 1995;158(2):263–95.
 13. Heeringa S, West B, Berglund P. Applied Survey data Analysis. New York: CRC Press, Taylor & Francis group; 2010.
 14. Ghana M of H. National Malaria Control Monitoring: GHANA. 2008.
 15. Floyd J, Fowler J. Survey Research Methods. 3rd ed. California: Sage Publications; 2002.
 16. Cochran WG. Sampling Techniques. 3rd ed. Wiley; 1977.
 17. Nemeth R. Sampling design of health surveys: household as a sampling unit. 2003.
 18. Kish L. Survey Sampling. Classics. New York: Wiley; 1995.
 19. Korn EL, Graubard BI. Analysis of Health surveys. John Wiley & Sons; 2011.
 20. de Toledo Vieira M. Analysis of Longitudinal Survey Data: Allowing for complex Survey Design in covariance structure models [Internet]. VDM Verlag; 2009 [cited 2016 Mar 14]. Available from:
[http://www.amazon.com/s?ie=UTF8&page=1&rh=n%3A283155%2Cp_27%3AMarcel de Toledo Vieira](http://www.amazon.com/s?ie=UTF8&page=1&rh=n%3A283155%2Cp_27%3AMarcel+de+Toledo+Vieira)
 21. de Toledo Vieira M. Modelling Complex Longitudinal Survey Data. 2005;(September).
 22. Lohr SL. Sampling design and analysis. 2nd ed. Boston: Richard Stratton; 2010.
 23. Skinner CJ, Holt D, Smith TMF. Analysis of complex surveys. Wiley Ser Probab Math Stat [Internet] 1989 [cited 2016 Mar 14]; Available from:
https://www.mendeley.com/research/analysis-complex-surveys-1/?utm_source=desktop&utm_medium=1.16&utm_campaign=open_catalog&userDocumentId=%7Bbecd0e8c-70c7-4f16-a8dc-7013d214bb40%7D
 24. DMDC Report No. 2007-003. 2005 Workplace and Equal Opportunity Survey of Active Duty Members : Statistical Methodology Report. Arlington: 2007.
 25. Merkouris T. Combining Independent Regression Estimators From Multiple Surveys. *J Am Stat Assoc* 2004;99(468):1131–9.
 26. Kish L. Rolling samples and censuses. *Surv Methodol* 1990;16(1):63–79.
 27. Kish L. Multipopulation Survey Designs : Five Types with Seven Shared

- Aspects *. *Int Stat Rev* 1994;62(2):167–86.
28. Kish L. Combining multipopulation statistics. *J Stat Plan Inference* 2002;102(1):109–18.
 29. Fox K. A Framework for the Meta-Analysis of Survey Data. In: *JSM Proceedings, Survey Research Methods Section*. 2010. page 527–38.
 30. Bailar J, Mosteller F. Combining results from independent investigations. In: *Medical uses of statistics*. Waltham MA: NEJM Press; 1986.
 31. Borenstein M, Hedges L V., Higgins JPT, Rothstein HR. *Introduction to meta-analysis*. Wiley; 2009.
 32. Rao SR, Graubard BI, Schmid CH, Morton SC, Louis AETA, Zaslavsky AEM, et al. Meta-analysis of survey data : application to health services research. *Heal Serv Outcomes Res Methodol* 2008;8:98–114.
 33. Kish L. Cumulating / Combining Population Surveys. *Surv Methodol* 1999;25(2):129–38.
 34. Schenker N, Gentleman JF, Rose D, Hing E, Shimizu IM. Combining estimates from complementary surveys: a case study using prevalence estimates from national health surveys of households and nursing homes. *Public Health Rep* 2002;117(4):393–407.
 35. Botman S, Moore T, Moriarity C, Parsons VL. Design and Estimation for the National Health Interview Survey, 1995-2004. National Center for Health Statistics [Internet]. 2004. Available from:
http://www.cdc.gov/nchs/data/series/sr_02/sr02_130.pdf
 36. Jones A. The National Nursing Home Survey: 1999 summary. [Internet]. In: *Vital and health statistics. Series 13, Data from the National Health Survey*. 2002. page 1–116. Available from:
<http://www.ncbi.nlm.nih.gov/pubmed/12071118>
 37. Alexander CH. Still Rolling : Leslie Kish’S “Rolling Samples” and the American Community Survey. *Proc Stat Canada Symp* 2001 2001;1–10.
 38. Bethlehem J. Solving the Nonresponse Problem With Sample Matching ? *Soc Sci Comput Rev* 2016;34(1):59–77.
 39. Sarndal C, Swensson B, Wretman J. *Model assisted survey sampling*. New York: Springer; 1992.
 40. Kish L. *Survey Sampling*. 1st ed. New York: John Wiley & Sons; 1965.

41. Brannick MT, Yang L-Q, Cafri G. Comparison of Weights for Meta-Analysis of r and d Under Realistic Conditions. *Organ Res Methods* [Internet] 2011;14(4):587–607. Available from: <http://orm.sagepub.com/cgi/doi/10.1177/1094428110368725>
42. Wannell B, Thomas S. Combining cycles of the Canadian Community Health Survey. *Heal Reports* 2009;20(1).
43. Roberts G, Binder D. Analyses Based on Combining Similar Information from Multiple Surveys. In: *JSM Proceedings, Survey Research Methods Section*. 2009. page 2138–47.
44. Latouche M, Dufor J, Merkouris T. *Cross-Sectional Weighting : Combining Two or More Panels*. Ottawa: 2000.
45. Lavigne M, Michaud S. *General aspects of the survey*. 1998;
46. Schenker N, Raghunathan TE. Combining information from multiple surveys to enhance estimation of measures of health ‡. *Stat Med* 2007;26(8):1802–11.
47. Wu C. Combining information from multiple surveys through the empirical likelihood method. *Can J Stat* 2004;32(1):1–12.
48. O’Muircheartaigh C, Pedlow S. COMBINING SAMPLES VS. CUMULATING CASES: A COMPARISON OF TWO WEIGHTING STRATEGIES IN NLSY97. *Jt Stat Meet - Sect Surv Res Methods* 2003;3987–92.
49. Merkouris T. Efficient Small-Domain Estimation by Combining Information from Multiple Surveys. 2003;3419–25.
50. Dong Q, Elliott MR, Raghunathan TE. Combining information from multiple complex surveys. *Surv Methodol* 2014;40(2):347–54.
51. Park S, Kwang J, Stukel D. A measurement error model approach to survey data integration : combining information from two surveys. *METRON* 2017;75(3):345–57.
52. Lehtonen R, Pahkinen E. *Practical Methods for Design and Analysis of Complex Surveys*. In: Barnett Vic (Nottingham Trent University U, editor. *Statistics in Practice*. Chichester, England: John Wiley & Sons; 2004. page 3–338.
53. Rao JNK, Kim J. Estimatimation of Totals and Regression parameters by combining data from two independent surveys.
54. Lu Y. Regression coefficient estimation in dual frame surveys. *Commun Stat*

- Comput [Internet] 2014;43(7):1675–1684. Available from:
<http://www.tandfonline.com/doi/abs/10.1080/03610918.2012.752835>
55. Elliott MR, Davis WW. Obtaining Cancer Risk Factor Prevalence Estimates in Small Areas : Combining Data from Two Surveys. *J R Stat Soc Ser C(Applied Stat)* 2005;54(3):595–609.
 56. Lotfi F, Motlagh SN, Mahdavi G, Keshavarz K, Hadian M, Gorji HA. Factors Affecting the Utilization of Outpatient Health Services and Importance of Health Insurance. 2017;18(8).
 57. Morakinyo OM, Adebowale SA, Oloruntoba EO. Wealth status and sex differential of household head : implication for source of drinking water in Nigeria. *Arch Public Heal* [Internet] 2015;1–9. Available from:
<http://dx.doi.org/10.1186/s13690-015-0105-9>
 58. Thomas S. Combinining cycles of the Canadian Health Survey. In: *Proceedings of statistics Canada symposium; Methodology issues in measuring population health*. 2006.
 59. Chu A, Brick JM, Kalton G. Weights for Combining Surveys across Time or Space. In: *Bulletin of the International Statistical institute*. 1999. page 103–4.
 60. Kish L. *Statistical design for research*. John Wiley & Sons Inc.; 1987.
 61. ASPE. Appendix B : Data Set Aggregation B-1 B-2 [Internet]. :1–5. Available from: <https://aspe.hhs.gov/sites/default/files/pdf/74856/apb.pdf>
 62. Thomas S, Wannell B. Combining cycles of the Canadian Community Health Survey. *Heal Reports* [Internet] 2009;20(1):1–6. Available from:
http://publications.gc.ca/collections/collection_2009/statcan/82-003-X/82-003-x2009001-eng.pdf#page=55
 63. Mukhopadhyay P. *Complex surveys: Analysis of categorical data*. 2016.
 64. Fuller WA. *Sampling Statistics*. New Jersey: John Wiley & Sons; 2009.
 65. Sorensen JB. The Use and Misuse of the Coefficient of Variation in Organizational Demography Research. *Sociol methods Res* 2002;30(4):475–91.
 66. Crawley M. *The R Book*. Second. Chichester, England: John Wiley & Sons; 2013.
 67. D’agostino RB, Belanger A, D’Agostino Jr RB. A Suggestion for Using Powerful and Informative Tests of Normality. *Am Stat* 1990;44(4):316–21.

68. Wolter KW. Introduction to variance estimation. Second. New York: Springer; 2007.
69. Nguyen N, Burke O, Murphy P. A Simulation Study of Weighting Methods to Improve Labour Force Estimates of Immigrants in Ireland A Simulation Study of Weighting Methods to Improve Labour Force Estimates of Immigrants in Ireland. *J Off Stat* 2016;(September).
70. Kolenikov S, Pitblado J. Analysis of Complex Health Survey Data. In: Johnson T, editor. *Handbook of Health Survey Methods*. New Jersey: John Wiley & Sons Inc.; 2015. page 755–804.
71. Lumley AT. Package ‘survey’ Documentation. 2018;
72. Rimoldini L. Weighted skewness and kurtosis unbiased by sample size. 2013;1–33.
73. Parchami A, Bahonar S. Package ‘Weighted.Desc.Stat.’ 2016;
74. Lee S (UCLA C for HPR, Davis WW (National CI, Nguyen H (UCLA C for HPR, McNeel T (Information MSI., Brick JM (Westat), Flores-Cervantes I (Westat). Examining Trends and averages using combined cross-sectional survey data from multiple years [Internet]. 2007 [cited 2016 Feb 16]. Available from: [file:///C:/Users/User/Desktop/PhD/surveys_articles/multiple surveys/CHIS methodology paper 2007.pdf](file:///C:/Users/User/Desktop/PhD/surveys_articles/multiple%20surveys/CHIS%20methodology%20paper%202007.pdf)
75. Morton SC. Combining Surveys From A Meta-analysis Perspective. In: International Statistical Institute (ISI) Annual Conference. Helsinki: 1999.
76. Sidik K, Jonkman JN. Simple heterogeneity variance estimation for. *J R Stat Soc Ser C* 2005;54(2):367–84.
77. Sidik K, Jonkman JN. Robust variance estimation for random effects meta-analysis. *Comput Stat Data Anal* 2006;50(12):3681–701.
78. Efron B. Bootstrap methods: Another look at the Jackknife. *Ann Stat* 1979;7(1):1–26.
79. Quenouille MH. Problems in Plane Sampling. *Ann Math Stat* [Internet] 1949;20(3):355–75. Available from: <http://www.jstor.org/stable/2238700>
80. Rust K. Variance estimation for complex estimators in sample surveys. *J Off Stat* 1985;1(4):381–97.
81. Efron B, Stein C. Efron1978_The Jackknife estimate of variance.pdf. 1978.
82. Miller RG. Miller1974_The Jackknife--A review. *Biometrika* 1974;61(1):1–15.

83. Calaway R, Microsoft corporation, Weston S, Tenenbaum D. doParallel [Internet]. 2018; Available from: <https://cran.r-project.org/web/packages/doParallel/index.html>
84. Grama A, Gupta A, Karypis G, Kumar V. Introduction to parallel computing. 2nd ed. Pearson Education Limited; 2003.
85. Statistics South Africa. Community Survey 2016: Technical Report [Internet]. 2016. Available from: chrome-extension://oemmndcbldboiebfnladdacbfmadadm/http://cs2016.statssa.gov.za/wp-content/uploads/2016/06/CS-2016-Technical-report_Web.pdf
86. Statistics South Africa. Community survey 2016 in brief [Internet]. 2016. Available from: http://cs2016.statssa.gov.za/wp-content/uploads/2016/07/NT-30-06-2016-RELEASE-for-CS-2016-_Statistical-releas_1-July-2016.pdf
87. Statistics South Africa. General Household Survey 2016. 2018.

Appendix A1

Comparing weighting strategies: Simple Random Sampling: Normal distribution

Scenario 1: N=1 00 000 N(100,25) n=1000 $\theta = 99.988$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,983	-4,856	3,541	1,584	99,913	100,052	2,509
2	1/n	99,985	-2,204	2,546	1,139	99,935	100,035	1,297
	$(Var(\hat{\theta}_a))^{-1}$	99,986	-1,958	2,549	1,140	99,936	100,036	1,300
	$(CV(\hat{\theta}_a))^{-1}$	99,986	-1,955	2,547	1,139	99,936	100,036	1,298
	$(skewness(s_{dc}))^{-1}$	99,983	-4,128	2,829	1,265	99,928	100,039	1,602
	$(kurtosis(k_{dc}))^{-1}$	99,985	-2,371	2,547	1,139	99,935	100,035	1,298
	$(K^2)^{-1}$	99,985	-2,670	2,898	1,296	99,928	100,042	1,680
3	1/n	99,987	-0,621	2,058	0,921	99,947	100,027	0,847
	$(Var(\hat{\theta}_a))^{-1}$	99,987	-0,555	2,061	0,922	99,947	100,027	0,849
	$(CV(\hat{\theta}_a))^{-1}$	99,987	-0,415	2,059	0,921	99,947	100,028	0,848
	$(skewness(s_{dc}))^{-1}$	99,984	-3,353	2,467	1,103	99,936	100,033	1,218
	$(kurtosis(k_{dc}))^{-1}$	99,987	-0,597	2,060	0,921	99,947	100,027	0,849
	$(K^2)^{-1}$	99,986	-1,790	2,584	1,156	99,935	100,036	1,336
5	1/n	99,988	0,336	1,553	0,694	99,957	100,018	0,482
	$(Var(\hat{\theta}_a))^{-1}$	99,988	0,338	1,555	0,696	99,957	100,018	0,484
	$(CV(\hat{\theta}_a))^{-1}$	99,988	0,545	1,553	0,695	99,958	100,019	0,483
	$(skewness(s_{dc}))^{-1}$	99,987	-0,945	2,060	0,922	99,946	100,027	0,849
	$(kurtosis(k_{dc}))^{-1}$	99,988	0,208	1,551	0,694	99,957	100,018	0,481
	$(K^2)^{-1}$	99,986	-1,517	2,217	0,992	99,943	100,030	0,983
10	1/n	99,990	1,995	1,109	0,496	99,968	100,011	0,246
	$(Var(\hat{\theta}_a))^{-1}$	99,990	1,948	1,109	0,496	99,968	100,011	0,246
	$(CV(\hat{\theta}_a))^{-1}$	99,990	2,203	1,109	0,496	99,968	100,012	0,246
	$(skewness(s_{dc}))^{-1}$	99,988	0,719	1,850	0,828	99,952	100,025	0,685
	$(kurtosis(k_{dc}))^{-1}$	99,989	1,901	1,111	0,497	99,968	100,011	0,247
	$(K^2)^{-1}$	99,986	-1,962	1,913	0,856	99,948	100,023	0,732
15	1/n	99,989	1,312	0,904	0,404	99,971	100,007	0,164
	$(Var(\hat{\theta}_a))^{-1}$	99,989	1,286	0,905	0,405	99,971	100,007	0,164
	$(CV(\hat{\theta}_a))^{-1}$	99,989	1,536	0,904	0,404	99,971	100,007	0,164
	$(skewness(s_{dc}))^{-1}$	99,988	0,892	1,725	0,771	99,955	100,022	0,595
	$(kurtosis(k_{dc}))^{-1}$	99,989	1,274	0,905	0,405	99,971	100,007	0,164
	$(K^2)^{-1}$	99,986	-1,529	1,771	0,792	99,951	100,021	0,627
20	1/n	99,989	0,955	0,800	0,358	99,973	100,004	0,128
	$(Var(\hat{\theta}_a))^{-1}$	99,989	0,932	0,802	0,359	99,973	100,004	0,129
	$(CV(\hat{\theta}_a))^{-1}$	99,989	1,185	0,801	0,358	99,973	100,004	0,128
	$(skewness(s_{dc}))^{-1}$	99,988	0,831	1,627	0,728	99,957	100,020	0,530
	$(kurtosis(k_{dc}))^{-1}$	99,989	0,944	0,803	0,359	99,973	100,004	0,129
	$(K^2)^{-1}$	99,986	-2,014	1,627	0,728	99,954	100,017	0,530

$\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 2: N=1 00 000 N(100,25) n=5000 $\theta = 99.967$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		99,970	3,027	1,556	6,962	99,940	100,001	4,853
2	1/n	99,969	1,698	1,114	4,982	99,947	99,991	2,484
	$(Var(\hat{\theta}_a))^{-1}$	99,969	1,713	1,114	4,984	99,947	99,991	2,485
	$(CV(\hat{\theta}_a))^{-1}$	99,969	1,730	1,114	4,983	99,947	99,991	2,484
	$(skewness(s_{dc}))^{-1}$	99,968	0,333	1,241	5,553	99,943	99,992	3,081
	$(kurtosis(k_{dc}))^{-1}$	99,969	1,685	1,113	4,980	99,947	99,991	2,482
3	1/n	99,968	0,478	0,910	4,070	99,950	99,986	1,655
	$(Var(\hat{\theta}_a))^{-1}$	99,968	0,488	0,910	4,071	99,950	99,986	1,656
	$(CV(\hat{\theta}_a))^{-1}$	99,968	0,517	0,910	4,070	99,950	99,986	1,656
	$(skewness(s_{dc}))^{-1}$	99,966	-1,602	1,106	4,946	99,944	99,987	2,447
	$(kurtosis(k_{dc}))^{-1}$	99,968	0,458	0,908	4,063	99,950	99,986	1,650
5	1/n	99,968	0,127	0,690	3,088	99,954	99,981	0,953
	$(Var(\hat{\theta}_a))^{-1}$	99,968	0,128	0,691	3,089	99,954	99,981	0,954
	$(CV(\hat{\theta}_a))^{-1}$	99,968	0,167	0,690	3,089	99,954	99,981	0,953
	$(skewness(s_{dc}))^{-1}$	99,965	-2,501	0,980	4,382	99,946	99,984	1,925
	$(kurtosis(k_{dc}))^{-1}$	99,968	0,116	0,690	3,085	99,954	99,981	0,951
10	1/n	99,968	0,475	0,503	2,250	99,958	99,978	0,506
	$(Var(\hat{\theta}_a))^{-1}$	99,968	0,468	0,503	2,252	99,958	99,978	0,507
	$(CV(\hat{\theta}_a))^{-1}$	99,968	0,516	0,503	2,251	99,958	99,978	0,507
	$(skewness(s_{dc}))^{-1}$	99,966	-1,132	0,858	3,840	99,949	99,983	1,475
	$(kurtosis(k_{dc}))^{-1}$	99,968	0,479	0,503	2,252	99,958	99,978	0,507
15	1/n	99,968	0,208	0,416	1,862	99,959	99,976	0,347
	$(Var(\hat{\theta}_a))^{-1}$	99,968	0,205	0,417	1,864	99,959	99,976	0,347
	$(CV(\hat{\theta}_a))^{-1}$	99,968	0,253	0,416	1,863	99,960	99,976	0,347
	$(skewness(s_{dc}))^{-1}$	99,965	-2,203	0,820	3,667	99,949	99,981	1,349
	$(kurtosis(k_{dc}))^{-1}$	99,968	0,214	0,417	1,864	99,959	99,976	0,347
20	1/n	99,968	0,277	0,358	1,603	99,961	99,975	0,257
	$(Var(\hat{\theta}_a))^{-1}$	99,968	0,275	0,358	1,604	99,961	99,975	0,257
	$(CV(\hat{\theta}_a))^{-1}$	99,968	0,323	0,358	1,603	99,961	99,975	0,257
	$(skewness(s_{dc}))^{-1}$	99,967	-0,235	0,789	3,527	99,952	99,983	1,244
	$(kurtosis(k_{dc}))^{-1}$	99,968	0,289	0,359	1,605	99,961	99,975	0,257

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 3: N=1 00 000 N(100,25) n=10 000 $\theta = 99.987$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-3}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		99,986	-1,384	4,806	1,520	99,983	99,989	2,311
2	1/n	99,986	-1,369	3,509	1,110	99,984	99,988	1,233
	$(Var(\hat{\theta}_a))^{-1}$	99,986	-1,359	3,511	1,110	99,984	99,988	1,234
	$(CV(\hat{\theta}_a))^{-1}$	99,986	-1,353	3,510	1,110	99,984	99,988	1,233
	$(skewness(s_{dc}))^{-1}$	99,987	-0,039	3,973	1,256	99,985	99,990	1,578
	$(kurtosis(k_{dc}))^{-1}$	99,986	-1,357	3,510	1,110	99,984	99,988	1,233
3	1/n	99,986	-0,867	2,787	0,881	99,985	99,988	0,777
	$(Var(\hat{\theta}_a))^{-1}$	99,986	-0,871	2,788	0,882	99,985	99,988	0,778
	$(CV(\hat{\theta}_a))^{-1}$	99,986	-0,854	2,788	0,881	99,985	99,988	0,778
	$(skewness(s_{dc}))^{-1}$	99,988	0,556	3,541	1,120	99,985	99,990	1,254
	$(kurtosis(k_{dc}))^{-1}$	99,986	-0,864	2,788	0,881	99,985	99,988	0,778
5	1/n	99,987	-0,564	2,136	0,675	99,985	99,988	0,456
	$(Var(\hat{\theta}_a))^{-1}$	99,987	-0,562	2,136	0,675	99,985	99,988	0,456
	$(CV(\hat{\theta}_a))^{-1}$	99,987	-0,545	2,136	0,675	99,985	99,988	0,456
	$(skewness(s_{dc}))^{-1}$	99,988	0,626	3,083	0,975	99,986	99,990	0,950
	$(kurtosis(k_{dc}))^{-1}$	99,987	-0,563	2,135	0,675	99,985	99,988	0,456
10	1/n	99,987	-0,252	1,510	0,477	99,986	99,988	0,228
	$(Var(\hat{\theta}_a))^{-1}$	99,987	-0,245	1,510	0,478	99,986	99,988	0,228
	$(CV(\hat{\theta}_a))^{-1}$	99,987	-0,228	1,510	0,478	99,986	99,988	0,228
	$(skewness(s_{dc}))^{-1}$	99,988	0,662	2,742	0,867	99,986	99,989	0,752
	$(kurtosis(k_{dc}))^{-1}$	99,987	-0,238	1,510	0,477	99,986	99,988	0,228
15	1/n	99,987	-0,210	1,255	0,397	99,986	99,988	0,157
	$(Var(\hat{\theta}_a))^{-1}$	99,987	-0,209	1,255	0,397	99,986	99,988	0,158
	$(CV(\hat{\theta}_a))^{-1}$	99,987	-0,188	1,255	0,397	99,986	99,988	0,157
	$(skewness(s_{dc}))^{-1}$	99,988	0,429	2,502	0,791	99,986	99,989	0,626
	$(kurtosis(k_{dc}))^{-1}$	99,987	-0,201	1,255	0,397	99,986	99,988	0,157
20	1/n	99,987	-0,246	1,101	0,348	99,986	99,988	0,121
	$(Var(\hat{\theta}_a))^{-1}$	99,987	-0,247	1,102	0,348	99,986	99,988	0,121
	$(CV(\hat{\theta}_a))^{-1}$	99,987	-0,225	1,102	0,348	99,986	99,988	0,121
	$(skewness(s_{dc}))^{-1}$	99,988	0,401	2,338	0,739	99,986	99,989	0,547
	$(kurtosis(k_{dc}))^{-1}$	99,987	-0,234	1,101	0,348	99,986	99,988	0,121

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 7: N=5 000 000 N(100,25) n=1000 $\theta = 100.002$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		100,012	9,638	5,145	1,627	100,002	100,022	2,656
2	1/n	100,004	2,351	3,628	1,147	99,997	100,011	1,317
	$(Var(\hat{\theta}_a))^{-1}$	100,004	2,434	3,638	1,150	99,997	100,012	1,324
	$(CV(\hat{\theta}_a))^{-1}$	100,004	2,523	3,633	1,149	99,997	100,012	1,320
	$(skewness(s_{dc}))^{-1}$	100,007	5,036	4,162	1,316	99,999	100,015	1,735
	$(kurtosis(k_{dc}))^{-1}$	100,004	2,380	3,625	1,146	99,997	100,011	1,314
	$(K^2)^{-1}$	100,006	4,496	4,185	1,323	99,998	100,015	1,753
3	1/n	100,005	3,235	2,964	0,937	99,999	100,011	0,880
	$(Var(\hat{\theta}_a))^{-1}$	100,005	3,226	2,970	0,939	99,999	100,011	0,883
	$(CV(\hat{\theta}_a))^{-1}$	100,005	3,407	2,967	0,938	100,000	100,011	0,881
	$(skewness(s_{dc}))^{-1}$	100,004	1,919	3,692	1,167	99,997	100,011	1,363
	$(kurtosis(k_{dc}))^{-1}$	100,005	3,289	2,962	0,937	99,999	100,011	0,879
	$(K^2)^{-1}$	100,006	3,795	3,705	1,171	99,999	100,013	1,374
5	1/n	100,006	3,815	2,259	0,714	100,001	100,010	0,512
	$(Var(\hat{\theta}_a))^{-1}$	100,006	3,910	2,262	0,715	100,001	100,010	0,513
	$(CV(\hat{\theta}_a))^{-1}$	100,006	4,069	2,260	0,715	100,002	100,010	0,512
	$(skewness(s_{dc}))^{-1}$	100,003	1,204	3,316	1,048	99,997	100,010	1,100
	$(kurtosis(k_{dc}))^{-1}$	100,006	3,859	2,261	0,715	100,001	100,010	0,513
	$(K^2)^{-1}$	100,004	1,975	3,241	1,025	99,998	100,010	1,051
10	1/n	100,003	0,936	1,556	0,492	100,000	100,006	0,242
	$(Var(\hat{\theta}_a))^{-1}$	100,003	0,946	1,558	0,493	100,000	100,006	0,243
	$(CV(\hat{\theta}_a))^{-1}$	100,003	1,174	1,556	0,492	100,000	100,006	0,242
	$(skewness(s_{dc}))^{-1}$	100,004	2,489	2,770	0,876	99,999	100,010	0,768
	$(kurtosis(k_{dc}))^{-1}$	100,003	1,048	1,557	0,492	100,000	100,006	0,242
	$(K^2)^{-1}$	100,004	1,956	2,707	0,856	99,999	100,009	0,733
15	1/n	100,002	-0,154	1,281	0,405	99,999	100,004	0,164
	$(Var(\hat{\theta}_a))^{-1}$	100,002	-0,102	1,283	0,406	99,999	100,004	0,165
	$(CV(\hat{\theta}_a))^{-1}$	100,002	0,114	1,281	0,405	100,000	100,005	0,164
	$(skewness(s_{dc}))^{-1}$	100,003	0,957	2,651	0,838	99,998	100,008	0,703
	$(kurtosis(k_{dc}))^{-1}$	100,002	-0,116	1,282	0,405	99,999	100,004	0,164
	$(K^2)^{-1}$	100,002	-0,159	2,464	0,779	99,997	100,007	0,607
20	1/n	100,002	-0,161	1,144	0,362	100,000	100,004	0,131
	$(Var(\hat{\theta}_a))^{-1}$	100,002	-0,140	1,146	0,362	100,000	100,004	0,131
	$(CV(\hat{\theta}_a))^{-1}$	100,002	0,094	1,145	0,362	100,000	100,004	0,131
	$(skewness(s_{dc}))^{-1}$	100,003	1,085	2,588	0,818	99,998	100,008	0,670
	$(kurtosis(k_{dc}))^{-1}$	100,002	-0,122	1,144	0,362	100,000	100,004	0,131
	$(K^2)^{-1}$	100,001	-1,259	2,259	0,714	99,996	100,005	0,510

$\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$) = standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$) = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 8: N=5 000 000 N(100,25) n=5000 $\theta = 100.002$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	$SE(\hat{\theta}_c)$ $\times 10^{-2}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-3}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-3}$
						LB	UB	
1		100,001	-0,853	7,068	2,235	99,997	100,005	4,996
2	1/n	100,001	-1,251	4,994	1,579	99,998	100,004	2,496
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-1,236	4,994	1,579	99,998	100,004	2,496
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-1,217	4,994	1,579	99,998	100,004	2,496
	$(skewness(s_{dc}))^{-1}$	100,000	-1,741	5,612	1,775	99,997	100,004	3,153
	$(kurtosis(k_{dc}))^{-1}$	100,001	-1,249	4,995	1,580	99,998	100,004	2,497
3	1/n	100,000	-2,140	4,108	1,299	99,997	100,002	1,692
	$(Var(\hat{\theta}_a))^{-1}$	100,000	-2,124	4,107	1,299	99,997	100,002	1,692
	$(CV(\hat{\theta}_a))^{-1}$	100,000	-2,098	4,108	1,299	99,997	100,002	1,692
	$(skewness(s_{dc}))^{-1}$	99,999	-2,624	4,975	1,573	99,996	100,002	2,482
	$(kurtosis(k_{dc}))^{-1}$	100,000	-2,130	4,107	1,299	99,997	100,002	1,691
5	1/n	100,001	-0,628	3,206	1,014	99,999	100,003	1,028
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-0,601	3,205	1,014	99,999	100,003	1,028
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-0,574	3,205	1,014	99,999	100,003	1,028
	$(skewness(s_{dc}))^{-1}$	100,001	-1,443	4,433	1,402	99,998	100,003	1,967
	$(kurtosis(k_{dc}))^{-1}$	100,001	-0,641	3,205	1,014	99,999	100,003	1,028
10	1/n	100,000	-1,760	2,320	0,734	99,999	100,002	0,541
	$(Var(\hat{\theta}_a))^{-1}$	100,000	-1,738	2,319	0,733	99,999	100,002	0,541
	$(CV(\hat{\theta}_a))^{-1}$	100,000	-1,704	2,319	0,733	99,999	100,002	0,541
	$(skewness(s_{dc}))^{-1}$	100,000	-2,239	3,821	1,208	99,997	100,002	1,465
	$(kurtosis(k_{dc}))^{-1}$	100,000	-1,763	2,318	0,733	99,999	100,002	0,540
15	1/n	100,001	-1,190	1,871	0,592	100,000	100,002	0,352
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-1,163	1,870	0,591	100,000	100,002	0,351
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-1,130	1,871	0,592	100,000	100,002	0,351
	$(skewness(s_{dc}))^{-1}$	100,000	-2,245	3,520	1,113	99,998	100,002	1,244
	$(kurtosis(k_{dc}))^{-1}$	100,001	-1,197	1,871	0,592	100,000	100,002	0,351
20	1/n	100,001	-1,206	1,607	0,508	100,000	100,002	0,260
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-1,189	1,605	0,508	100,000	100,002	0,259
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-1,150	1,606	0,508	100,000	100,002	0,259
	$(skewness(s_{dc}))^{-1}$	100,000	-2,075	3,324	1,051	99,998	100,002	1,109
	$(kurtosis(k_{dc}))^{-1}$	100,001	-1,220	1,606	0,508	100,000	100,002	0,260

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, $SE(\hat{\theta}_c)$ =standard error of the combined estimate, $CV(\hat{\theta}_c)$ (%) = coefficient of variation, $CI_{95\%}(\hat{\theta}_c)$ = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 9: N=5 000 000 N(100,25) n=10 000 $\theta = 100.002$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	$SE(\hat{\theta}_c)$ $\times 10^{-2}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-3}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-3}$
						LB	UB	
1		100,003	0,559	5,106	1,615	99,999	100,006	2,608
2	1/n	100,001	-1,186	3,600	1,138	99,999	100,003	1,297
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-1,174	3,598	1,138	99,999	100,003	1,296
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-1,168	3,599	1,138	99,999	100,003	1,297
	$(skewness(s_{dc}))^{-1}$	100,001	-0,552	3,956	1,251	99,999	100,004	1,565
	$(kurtosis(k_{dc}))^{-1}$	100,001	-1,174	3,597	1,138	99,999	100,003	1,295
3	1/n	100,001	-1,249	2,920	0,923	99,999	100,003	0,854
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-1,249	2,919	0,923	99,999	100,003	0,853
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-1,232	2,919	0,923	99,999	100,003	0,854
	$(skewness(s_{dc}))^{-1}$	100,001	-0,723	3,536	1,118	99,999	100,003	1,251
	$(kurtosis(k_{dc}))^{-1}$	100,001	-1,246	2,917	0,923	99,999	100,003	0,853
5	1/n	100,001	-1,208	2,233	0,706	99,999	100,002	0,500
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-1,212	2,233	0,706	99,999	100,002	0,500
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-1,190	2,233	0,706	99,999	100,002	0,500
	$(skewness(s_{dc}))^{-1}$	100,002	-0,173	3,157	0,998	100,000	100,004	0,997
	$(kurtosis(k_{dc}))^{-1}$	100,001	-1,220	2,232	0,706	99,999	100,002	0,500
10	1/n	100,002	-0,392	1,601	0,506	100,001	100,003	0,256
	$(Var(\hat{\theta}_a))^{-1}$	100,002	-0,394	1,601	0,506	100,001	100,003	0,256
	$(CV(\hat{\theta}_a))^{-1}$	100,002	-0,370	1,601	0,506	100,001	100,003	0,256
	$(skewness(s_{dc}))^{-1}$	100,002	-0,204	2,658	0,840	100,000	100,003	0,706
	$(kurtosis(k_{dc}))^{-1}$	100,002	-0,406	1,601	0,506	100,001	100,003	0,256
15	1/n	100,002	-0,117	1,313	0,415	100,001	100,003	0,172
	$(Var(\hat{\theta}_a))^{-1}$	100,002	-0,115	1,313	0,415	100,001	100,003	0,172
	$(CV(\hat{\theta}_a))^{-1}$	100,002	-0,093	1,313	0,415	100,001	100,003	0,172
	$(skewness(s_{dc}))^{-1}$	100,002	0,306	2,468	0,780	100,001	100,004	0,609
	$(kurtosis(k_{dc}))^{-1}$	100,002	-0,129	1,313	0,415	100,001	100,003	0,172
20	1/n	100,002	0,085	1,109	0,351	100,001	100,003	0,123
	$(Var(\hat{\theta}_a))^{-1}$	100,002	0,086	1,109	0,351	100,001	100,003	0,123
	$(CV(\hat{\theta}_a))^{-1}$	100,002	0,109	1,109	0,351	100,001	100,003	0,123
	$(skewness(s_{dc}))^{-1}$	100,003	0,812	2,373	0,751	100,001	100,004	0,564
	$(kurtosis(k_{dc}))^{-1}$	100,002	0,075	1,109	0,351	100,001	100,003	0,123

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, $SE(\hat{\theta}_c)$ =standard error of the combined estimate, $CV(\hat{\theta}_c)$ (%) = coefficient of variation, $CI_{95\%}(\hat{\theta}_c)$ = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 10: N=100 000 N(100,49) n=1000 $\theta = 100.023$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		100,027	4,430	7,027	2,222	100,014	100,041	4,940
2	1/n	100,024	1,353	4,840	1,530	100,015	100,034	2,343
	$(Var(\hat{\theta}_a))^{-1}$	100,024	1,347	4,833	1,528	100,015	100,034	2,336
	$(CV(\hat{\theta}_a))^{-1}$	100,025	1,603	4,836	1,529	100,015	100,034	2,339
	$(skewness(s_{dc}))^{-1}$	100,025	2,439	5,562	1,758	100,014	100,036	3,094
	$(kurtosis(k_{dc}))^{-1}$	100,024	1,022	4,854	1,535	100,014	100,033	2,356
	$(K^2)^{-1}$	100,025	1,882	5,530	1,748	100,014	100,036	3,058
3	1/n	100,025	2,427	4,025	1,272	100,017	100,033	1,620
	$(Var(\hat{\theta}_a))^{-1}$	100,025	2,432	4,032	1,275	100,017	100,033	1,626
	$(CV(\hat{\theta}_a))^{-1}$	100,026	2,760	4,027	1,273	100,018	100,034	1,623
	$(skewness(s_{dc}))^{-1}$	100,027	4,090	4,914	1,553	100,017	100,037	2,416
	$(kurtosis(k_{dc}))^{-1}$	100,025	2,199	4,035	1,276	100,017	100,033	1,628
	$(K^2)^{-1}$	100,028	5,184	4,968	1,571	100,018	100,038	2,471
5	1/n	100,022	-0,595	3,166	1,001	100,016	100,029	1,003
	$(Var(\hat{\theta}_a))^{-1}$	100,022	-0,474	3,169	1,002	100,016	100,029	1,004
	$(CV(\hat{\theta}_a))^{-1}$	100,023	-0,142	3,167	1,001	100,017	100,029	1,003
	$(skewness(s_{dc}))^{-1}$	100,023	-0,251	4,346	1,374	100,014	100,031	1,889
	$(kurtosis(k_{dc}))^{-1}$	100,022	-0,779	3,168	1,002	100,016	100,028	1,004
	$(K^2)^{-1}$	100,024	0,797	4,296	1,358	100,015	100,032	1,846
10	1/n	100,022	-1,204	2,197	0,695	100,017	100,026	0,483
	$(Var(\hat{\theta}_a))^{-1}$	100,022	-1,164	2,199	0,695	100,017	100,026	0,484
	$(CV(\hat{\theta}_a))^{-1}$	100,022	-0,748	2,198	0,695	100,018	100,026	0,483
	$(skewness(s_{dc}))^{-1}$	100,023	-0,320	3,644	1,152	100,015	100,030	1,328
	$(kurtosis(k_{dc}))^{-1}$	100,022	-1,376	2,197	0,695	100,017	100,026	0,483
	$(K^2)^{-1}$	100,025	1,954	3,535	1,118	100,018	100,032	1,250
15	1/n	100,022	-0,621	1,795	0,568	100,019	100,026	0,322
	$(Var(\hat{\theta}_a))^{-1}$	100,022	-0,612	1,796	0,568	100,019	100,026	0,323
	$(CV(\hat{\theta}_a))^{-1}$	100,023	-0,167	1,795	0,568	100,019	100,026	0,322
	$(skewness(s_{dc}))^{-1}$	100,023	-0,182	3,426	1,083	100,016	100,029	1,174
	$(kurtosis(k_{dc}))^{-1}$	100,022	-0,816	1,795	0,568	100,019	100,026	0,322
	$(K^2)^{-1}$	100,023	-0,181	3,179	1,005	100,017	100,029	1,010
20	1/n	100,023	-0,118	1,542	0,488	100,020	100,026	0,238
	$(Var(\hat{\theta}_a))^{-1}$	100,023	-0,109	1,544	0,488	100,020	100,026	0,238
	$(CV(\hat{\theta}_a))^{-1}$	100,023	0,351	1,543	0,488	100,020	100,026	0,238
	$(skewness(s_{dc}))^{-1}$	100,024	1,020	3,296	1,042	100,017	100,030	1,087
	$(kurtosis(k_{dc}))^{-1}$	100,023	-0,327	1,542	0,488	100,020	100,026	0,238
	$(K^2)^{-1}$	100,022	-0,553	3,010	0,952	100,016	100,028	0,906

$\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 11: N=100 000 N(100,49) n=5000 $\theta = 100.023$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	$SE(\hat{\theta}_c)$ $\times 10^{-2}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-3}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-3}$
						LB	UB	
1		100,024	1,086	9,465	2,994	100,018	100,030	8,964
2	1/n	100,023	0,210	6,646	2,102	100,019	100,027	4,419
	$(Var(\hat{\theta}_a))^{-1}$	100,023	0,231	6,644	2,101	100,019	100,027	4,416
	$(CV(\hat{\theta}_a))^{-1}$	100,023	0,265	6,644	2,102	100,019	100,027	4,417
	$(skewness(s_{dc}))^{-1}$	100,023	-0,277	7,583	2,399	100,018	100,027	5,753
	$(kurtosis(k_{dc}))^{-1}$	100,023	0,191	6,649	2,103	100,019	100,027	4,423
3	1/n	100,024	0,904	5,434	1,719	100,020	100,027	2,955
	$(Var(\hat{\theta}_a))^{-1}$	100,024	0,923	5,433	1,719	100,020	100,027	2,954
	$(CV(\hat{\theta}_a))^{-1}$	100,024	0,975	5,433	1,719	100,021	100,027	2,954
	$(skewness(s_{dc}))^{-1}$	100,024	0,949	6,893	2,180	100,020	100,028	4,754
	$(kurtosis(k_{dc}))^{-1}$	100,024	0,803	5,434	1,719	100,020	100,027	2,955
5	1/n	100,024	0,815	4,232	1,339	100,021	100,026	1,793
	$(Var(\hat{\theta}_a))^{-1}$	100,024	0,856	4,232	1,339	100,021	100,026	1,793
	$(CV(\hat{\theta}_a))^{-1}$	100,024	0,910	4,232	1,339	100,021	100,026	1,793
	$(skewness(s_{dc}))^{-1}$	100,025	1,572	6,090	1,926	100,021	100,028	3,714
	$(kurtosis(k_{dc}))^{-1}$	100,024	0,764	4,237	1,340	100,021	100,026	1,797
10	1/n	100,023	0,185	2,994	0,947	100,021	100,025	0,897
	$(Var(\hat{\theta}_a))^{-1}$	100,023	0,176	2,996	0,948	100,021	100,025	0,898
	$(CV(\hat{\theta}_a))^{-1}$	100,023	0,265	2,995	0,947	100,021	100,025	0,897
	$(skewness(s_{dc}))^{-1}$	100,021	-2,013	5,225	1,653	100,018	100,024	2,735
	$(kurtosis(k_{dc}))^{-1}$	100,023	0,157	2,997	0,948	100,021	100,025	0,899
15	1/n	100,024	0,582	2,490	0,788	100,022	100,025	0,621
	$(Var(\hat{\theta}_a))^{-1}$	100,024	0,594	2,494	0,789	100,022	100,025	0,623
	$(CV(\hat{\theta}_a))^{-1}$	100,024	0,675	2,492	0,788	100,022	100,025	0,622
	$(skewness(s_{dc}))^{-1}$	100,021	-1,623	4,748	1,502	100,018	100,024	2,258
	$(kurtosis(k_{dc}))^{-1}$	100,023	0,543	2,492	0,788	100,022	100,025	0,621
20	1/n	100,023	-0,168	2,226	0,704	100,021	100,024	0,496
	$(Var(\hat{\theta}_a))^{-1}$	100,023	-0,153	2,229	0,705	100,021	100,024	0,497
	$(CV(\hat{\theta}_a))^{-1}$	100,023	-0,071	2,227	0,705	100,021	100,024	0,496
	$(skewness(s_{dc}))^{-1}$	100,021	-1,547	4,680	1,480	100,018	100,024	2,193
	$(kurtosis(k_{dc}))^{-1}$	100,023	-0,204	2,227	0,704	100,021	100,024	0,496

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, $SE(\hat{\theta}_c)$ =standard error of the combined estimate, $CV(\hat{\theta}_c)$ (%) = coefficient of variation, $CI_{95\%}(\hat{\theta}_c)$ = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 12: N=100 000 N(100,49) n=10000 $\theta = 100.023$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-3}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		100,025	1,770	6,557	2,074	100,021	100,029	4,305
2	1/n	100,022	-0,704	4,690	1,483	100,019	100,025	2,201
	$(Var(\hat{\theta}_a))^{-1}$	100,022	-0,717	4,691	1,484	100,019	100,025	2,202
	$(CV(\hat{\theta}_a))^{-1}$	100,022	-0,689	4,691	1,484	100,019	100,025	2,202
	$(skewness(s_{dc}))^{-1}$	100,023	-0,240	5,292	1,674	100,019	100,026	2,802
	$(kurtosis(k_{dc}))^{-1}$	100,022	-0,702	4,689	1,483	100,019	100,025	2,200
3	1/n	100,022	-1,012	3,812	1,206	100,020	100,024	1,455
	$(Var(\hat{\theta}_a))^{-1}$	100,022	-1,004	3,811	1,205	100,020	100,024	1,454
	$(CV(\hat{\theta}_a))^{-1}$	100,022	-0,980	3,812	1,206	100,020	100,024	1,454
	$(skewness(s_{dc}))^{-1}$	100,022	-0,539	4,651	1,471	100,020	100,025	2,164
	$(kurtosis(k_{dc}))^{-1}$	100,022	-1,009	3,813	1,206	100,020	100,024	1,456
5	1/n	100,022	-0,580	2,927	0,926	100,021	100,024	0,857
	$(Var(\hat{\theta}_a))^{-1}$	100,022	-0,569	2,926	0,925	100,021	100,024	0,857
	$(CV(\hat{\theta}_a))^{-1}$	100,022	-0,540	2,926	0,926	100,021	100,024	0,857
	$(skewness(s_{dc}))^{-1}$	100,023	0,216	4,141	1,310	100,021	100,026	1,716
	$(kurtosis(k_{dc}))^{-1}$	100,022	-0,584	2,927	0,926	100,021	100,024	0,858
10	1/n	100,023	-0,201	2,098	0,664	100,021	100,024	0,440
	$(Var(\hat{\theta}_a))^{-1}$	100,023	-0,192	2,097	0,663	100,021	100,024	0,440
	$(CV(\hat{\theta}_a))^{-1}$	100,023	-0,157	2,097	0,663	100,021	100,024	0,440
	$(skewness(s_{dc}))^{-1}$	100,024	0,983	3,548	1,122	100,022	100,026	1,261
	$(kurtosis(k_{dc}))^{-1}$	100,023	-0,210	2,098	0,663	100,021	100,024	0,440
15	1/n	100,023	-0,280	1,719	0,544	100,022	100,024	0,296
	$(Var(\hat{\theta}_a))^{-1}$	100,023	-0,275	1,719	0,544	100,022	100,024	0,296
	$(CV(\hat{\theta}_a))^{-1}$	100,023	-0,236	1,719	0,544	100,022	100,024	0,296
	$(skewness(s_{dc}))^{-1}$	100,024	1,338	3,371	1,066	100,022	100,026	1,139
	$(kurtosis(k_{dc}))^{-1}$	100,023	-0,285	1,719	0,544	100,022	100,024	0,296
20	1/n	100,023	-0,216	1,487	0,470	100,022	100,024	0,221
	$(Var(\hat{\theta}_a))^{-1}$	100,023	-0,207	1,487	0,470	100,022	100,024	0,221
	$(CV(\hat{\theta}_a))^{-1}$	100,023	-0,170	1,487	0,470	100,022	100,024	0,221
	$(skewness(s_{dc}))^{-1}$	100,025	1,954	3,158	0,999	100,023	100,027	1,002
	$(kurtosis(k_{dc}))^{-1}$	100,023	-0,224	1,486	0,470	100,022	100,024	0,221

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 13: N=1 000 000 N(100,49) n=1000 $\theta = 99.996$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		100,003	6,812	6,746	2,133	99,990	100,016	4,556
2	1/n	100,004	7,979	5,146	1,627	99,994	100,014	2,655
	$(Var(\hat{\theta}_a))^{-1}$	100,004	8,085	5,137	1,624	99,994	100,014	2,645
	$(CV(\hat{\theta}_a))^{-1}$	100,004	8,246	5,141	1,626	99,994	100,014	2,650
	$(skewness(s_{dc}))^{-1}$	99,999	3,439	5,721	1,809	99,988	100,011	3,274
	$(kurtosis(k_{dc}))^{-1}$	100,004	8,003	5,137	1,624	99,994	100,014	2,645
	$(K^2)^{-1}$	100,001	5,196	5,754	1,820	99,990	100,012	3,314
3	1/n	100,005	9,120	4,166	1,317	99,997	100,013	1,744
	$(Var(\hat{\theta}_a))^{-1}$	100,005	9,389	4,170	1,319	99,997	100,014	1,748
	$(CV(\hat{\theta}_a))^{-1}$	100,006	9,584	4,167	1,318	99,997	100,014	1,746
	$(skewness(s_{dc}))^{-1}$	100,000	3,721	5,023	1,588	99,990	100,010	2,524
	$(kurtosis(k_{dc}))^{-1}$	100,005	8,992	4,158	1,315	99,997	100,013	1,737
	$(K^2)^{-1}$	100,002	5,523	5,081	1,607	99,992	100,011	2,584
5	1/n	99,998	1,941	3,257	1,030	99,992	100,004	1,061
	$(Var(\hat{\theta}_a))^{-1}$	99,998	2,152	3,264	1,032	99,992	100,005	1,066
	$(CV(\hat{\theta}_a))^{-1}$	99,998	2,447	3,260	1,031	99,992	100,005	1,064
	$(skewness(s_{dc}))^{-1}$	99,994	-1,770	4,567	1,444	99,985	100,003	2,086
	$(kurtosis(k_{dc}))^{-1}$	99,998	1,734	3,258	1,030	99,991	100,004	1,062
	$(K^2)^{-1}$	99,992	-3,906	4,448	1,407	99,983	100,001	1,980
10	1/n	99,999	2,961	2,254	0,713	99,995	100,003	0,509
	$(Var(\hat{\theta}_a))^{-1}$	99,999	3,054	2,257	0,714	99,995	100,003	0,510
	$(CV(\hat{\theta}_a))^{-1}$	99,999	3,451	2,255	0,713	99,995	100,004	0,510
	$(skewness(s_{dc}))^{-1}$	99,994	-2,426	3,832	1,212	99,986	100,001	1,469
	$(kurtosis(k_{dc}))^{-1}$	99,999	2,896	2,252	0,712	99,994	100,003	0,508
	$(K^2)^{-1}$	99,992	-3,572	3,767	1,191	99,985	100,000	1,421
15	1/n	99,999	2,893	1,784	0,564	99,995	100,002	0,319
	$(Var(\hat{\theta}_a))^{-1}$	99,999	2,996	1,784	0,564	99,996	100,003	0,319
	$(CV(\hat{\theta}_a))^{-1}$	99,999	3,407	1,784	0,564	99,996	100,003	0,319
	$(skewness(s_{dc}))^{-1}$	99,993	-3,351	3,384	1,070	99,986	99,999	1,146
	$(kurtosis(k_{dc}))^{-1}$	99,999	2,718	1,785	0,565	99,995	100,002	0,319
	$(K^2)^{-1}$	99,996	-0,459	3,294	1,042	99,989	100,002	1,085
20	1/n	99,998	2,371	1,577	0,499	99,995	100,001	0,249
	$(Var(\hat{\theta}_a))^{-1}$	99,998	2,417	1,578	0,499	99,995	100,002	0,249
	$(CV(\hat{\theta}_a))^{-1}$	99,999	2,862	1,577	0,499	99,996	100,002	0,249
	$(skewness(s_{dc}))^{-1}$	99,991	-4,691	3,362	1,063	99,985	99,998	1,132
	$(kurtosis(k_{dc}))^{-1}$	99,998	2,252	1,576	0,498	99,995	100,001	0,249
	$(K^2)^{-1}$	99,996	0,340	3,024	0,956	99,990	100,002	0,914

$\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 14 N=1 000 000 N(100,49) n=5000 $\theta = 99.996$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	$SE(\hat{\theta}_c)$ $\times 10^{-1}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-2}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-2}$
						LB	UB	
1		99,999	2,538	1,019	0,322	99,992	100,005	1,040
2	1/n	99,996	-0,290	0,723	0,229	99,991	100,000	0,522
	$(Var(\hat{\theta}_a))^{-1}$	99,996	-0,297	0,723	0,229	99,991	100,000	0,522
	$(CV(\hat{\theta}_a))^{-1}$	99,996	-0,244	0,723	0,229	99,991	100,000	0,522
	$(skewness(s_{dc}))^{-1}$	99,996	-0,400	0,817	0,258	99,991	100,001	0,668
	$(kurtosis(k_{dc}))^{-1}$	99,996	-0,301	0,723	0,229	99,991	100,000	0,523
3	1/n	99,996	0,150	0,584	0,185	99,993	100,000	0,341
	$(Var(\hat{\theta}_a))^{-1}$	99,996	0,120	0,584	0,185	99,993	100,000	0,341
	$(CV(\hat{\theta}_a))^{-1}$	99,996	0,203	0,584	0,185	99,993	100,000	0,341
	$(skewness(s_{dc}))^{-1}$	99,997	0,511	0,717	0,227	99,992	100,001	0,514
	$(kurtosis(k_{dc}))^{-1}$	99,996	0,178	0,585	0,185	99,993	100,000	0,342
5	1/n	99,997	0,707	0,433	0,137	99,994	99,999	0,188
	$(Var(\hat{\theta}_a))^{-1}$	99,997	0,704	0,433	0,137	99,994	99,999	0,188
	$(CV(\hat{\theta}_a))^{-1}$	99,997	0,784	0,433	0,137	99,994	99,999	0,188
	$(skewness(s_{dc}))^{-1}$	99,997	1,196	0,617	0,195	99,993	100,001	0,381
	$(kurtosis(k_{dc}))^{-1}$	99,997	0,725	0,434	0,137	99,994	99,999	0,188
10	1/n	99,996	-0,022	0,304	0,096	99,994	99,998	0,092
	$(Var(\hat{\theta}_a))^{-1}$	99,996	0,008	0,304	0,096	99,994	99,998	0,092
	$(CV(\hat{\theta}_a))^{-1}$	99,996	0,081	0,304	0,096	99,994	99,998	0,092
	$(skewness(s_{dc}))^{-1}$	99,995	-0,889	0,518	0,164	99,992	99,998	0,268
	$(kurtosis(k_{dc}))^{-1}$	99,996	-0,016	0,304	0,096	99,994	99,998	0,093
15	1/n	99,997	0,707	0,252	0,080	99,995	99,998	0,063
	$(Var(\hat{\theta}_a))^{-1}$	99,997	0,730	0,252	0,080	99,995	99,998	0,063
	$(CV(\hat{\theta}_a))^{-1}$	99,997	0,809	0,252	0,080	99,995	99,998	0,063
	$(skewness(s_{dc}))^{-1}$	99,995	-0,537	0,478	0,151	99,993	99,998	0,228
	$(kurtosis(k_{dc}))^{-1}$	99,997	0,709	0,252	0,080	99,995	99,998	0,064
20	1/n	99,997	0,584	0,214	0,068	99,995	99,998	0,046
	$(Var(\hat{\theta}_a))^{-1}$	99,997	0,607	0,214	0,068	99,995	99,998	0,046
	$(CV(\hat{\theta}_a))^{-1}$	99,997	0,687	0,214	0,068	99,995	99,998	0,046
	$(skewness(s_{dc}))^{-1}$	99,996	0,019	0,463	0,147	99,993	99,999	0,215
	$(kurtosis(k_{dc}))^{-1}$	99,997	0,586	0,215	0,068	99,995	99,998	0,046

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, $SE(\hat{\theta}_c)$ =standard error of the combined estimate, $CV(\hat{\theta}_c)$ (%) = coefficient of variation, $CI_{95\%}(\hat{\theta}_c)$ = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 15: N=1 000 000 N(100,49) n=10000 $\theta = 99.996$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	$SE(\hat{\theta}_c)$ $\times 10^{-3}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-2}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-3}$
						LB	UB	
1		99,993	-3,388	2,204	6,970	99,988	99,997	4,868
2	1/n	99,995	-1,288	1,598	5,053	99,992	99,998	2,554
	$(Var(\hat{\theta}_a))^{-1}$	99,995	-1,273	1,598	5,054	99,992	99,998	2,556
	$(CV(\hat{\theta}_a))^{-1}$	99,995	-1,255	1,598	5,053	99,992	99,998	2,555
	$(skewness(s_{dc}))^{-1}$	99,995	-1,028	1,747	5,524	99,992	99,998	3,053
	$(kurtosis(k_{dc}))^{-1}$	99,995	-1,276	1,598	5,055	99,992	99,998	2,556
3	1/n	99,996	-0,482	1,285	4,063	99,993	99,998	1,651
	$(Var(\hat{\theta}_a))^{-1}$	99,996	-0,458	1,285	4,063	99,993	99,998	1,651
	$(CV(\hat{\theta}_a))^{-1}$	99,996	-0,437	1,285	4,063	99,993	99,998	1,651
	$(skewness(s_{dc}))^{-1}$	99,995	-1,375	1,523	4,816	99,992	99,998	2,321
	$(kurtosis(k_{dc}))^{-1}$	99,996	-0,454	1,285	4,063	99,993	99,998	1,651
5	1/n	99,996	0,231	1,007	3,186	99,994	99,998	1,015
	$(Var(\hat{\theta}_a))^{-1}$	99,996	0,252	1,007	3,184	99,994	99,998	1,014
	$(CV(\hat{\theta}_a))^{-1}$	99,996	0,280	1,007	3,185	99,994	99,998	1,014
	$(skewness(s_{dc}))^{-1}$	99,995	-0,535	1,366	4,319	99,993	99,998	1,866
	$(kurtosis(k_{dc}))^{-1}$	99,996	0,240	1,007	3,185	99,994	99,998	1,014
10	1/n	99,996	0,051	0,689	2,177	99,995	99,997	0,474
	$(Var(\hat{\theta}_a))^{-1}$	99,996	0,056	0,688	2,177	99,995	99,997	0,474
	$(CV(\hat{\theta}_a))^{-1}$	99,996	0,097	0,688	2,177	99,995	99,997	0,474
	$(skewness(s_{dc}))^{-1}$	99,995	-0,573	1,166	3,686	99,993	99,998	1,359
	$(kurtosis(k_{dc}))^{-1}$	99,996	0,065	0,688	2,177	99,995	99,997	0,474
15	1/n	99,996	0,272	0,565	1,786	99,995	99,997	0,319
	$(Var(\hat{\theta}_a))^{-1}$	99,996	0,277	0,564	1,785	99,995	99,997	0,319
	$(CV(\hat{\theta}_a))^{-1}$	99,996	0,319	0,565	1,785	99,995	99,997	0,319
	$(skewness(s_{dc}))^{-1}$	99,996	0,089	1,127	3,565	99,994	99,998	1,271
	$(kurtosis(k_{dc}))^{-1}$	99,996	0,281	0,565	1,786	99,995	99,997	0,319
20	1/n	99,996	0,146	0,495	1,565	99,995	99,997	0,245
	$(Var(\hat{\theta}_a))^{-1}$	99,996	0,150	0,495	1,564	99,995	99,997	0,245
	$(CV(\hat{\theta}_a))^{-1}$	99,996	0,194	0,495	1,564	99,995	99,997	0,245
	$(skewness(s_{dc}))^{-1}$	99,996	-0,335	1,085	3,431	99,994	99,998	1,177
	$(kurtosis(k_{dc}))^{-1}$	99,996	0,155	0,495	1,565	99,995	99,997	0,245

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, $SE(\hat{\theta}_c)$ =standard error of the combined estimate, $CV(\hat{\theta}_c)$ (%) = coefficient of variation, $CI_{95\%}(\hat{\theta}_c)$ = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 16: N=5 000 000 N(100,49) n=1000 $\theta = 100.003$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		100,016	1,349	7,203	2,277	100,002	100,030	5,207
2	1/n	100,006	0,329	5,080	1,606	99,996	100,016	2,581
	$(Var(\hat{\theta}_d))^{-1}$	100,006	0,341	5,093	1,610	99,996	100,016	2,595
	$(CV(\hat{\theta}_d))^{-1}$	100,006	0,329	5,080	1,606	99,996	100,016	2,581
	$(skewness(s_{dc}))^{-1}$	100,006	0,361	5,086	1,608	99,996	100,016	2,588
	$(kurtosis(k_{dc}))^{-1}$	100,010	0,705	5,827	1,842	99,998	100,021	3,400
	$(K^2)^{-1}$	100,006	0,333	5,075	1,605	99,996	100,016	2,576
3	1/n	100,009	0,629	5,858	1,852	99,998	100,021	3,436
	$(Var(\hat{\theta}_d))^{-1}$	100,007	0,453	4,150	1,312	99,999	100,015	1,724
	$(CV(\hat{\theta}_d))^{-1}$	100,007	0,452	4,157	1,315	99,999	100,015	1,730
	$(skewness(s_{dc}))^{-1}$	100,007	0,453	4,150	1,312	99,999	100,015	1,724
	$(kurtosis(k_{dc}))^{-1}$	100,008	0,487	4,153	1,313	99,999	100,016	1,727
	$(K^2)^{-1}$	100,005	0,269	5,168	1,634	99,995	100,016	2,672
5	1/n	100,007	0,460	4,147	1,311	99,999	100,015	1,722
	$(Var(\hat{\theta}_d))^{-1}$	100,008	0,531	5,186	1,640	99,998	100,018	2,693
	$(CV(\hat{\theta}_d))^{-1}$	100,008	0,534	3,163	1,000	100,002	100,014	1,003
	$(skewness(s_{dc}))^{-1}$	100,008	0,547	3,166	1,001	100,002	100,014	1,005
	$(kurtosis(k_{dc}))^{-1}$	100,008	0,534	3,163	1,000	100,002	100,014	1,003
	$(K^2)^{-1}$	100,009	0,581	3,164	1,000	100,002	100,015	1,004
10	1/n	100,004	0,169	4,642	1,468	99,995	100,014	2,155
	$(Var(\hat{\theta}_d))^{-1}$	100,008	0,540	3,165	1,001	100,002	100,014	1,005
	$(CV(\hat{\theta}_d))^{-1}$	100,006	0,277	4,537	1,435	99,997	100,014	2,059
	$(skewness(s_{dc}))^{-1}$	100,004	0,131	2,178	0,689	100,000	100,008	0,474
	$(kurtosis(k_{dc}))^{-1}$	100,004	0,132	2,181	0,690	100,000	100,008	0,476
	$(K^2)^{-1}$	100,004	0,131	2,178	0,689	100,000	100,008	0,474
15	1/n	100,005	0,178	2,179	0,689	100,000	100,009	0,475
	$(Var(\hat{\theta}_d))^{-1}$	100,006	0,348	3,878	1,226	99,999	100,014	1,505
	$(CV(\hat{\theta}_d))^{-1}$	100,004	0,147	2,179	0,689	100,000	100,008	0,475
	$(skewness(s_{dc}))^{-1}$	100,005	0,274	3,790	1,198	99,998	100,013	1,437
	$(kurtosis(k_{dc}))^{-1}$	100,003	-0,022	1,793	0,567	99,999	100,006	0,322
	$(K^2)^{-1}$	100,003	-0,014	1,796	0,568	99,999	100,006	0,322
20	1/n	100,003	-0,022	1,793	0,567	99,999	100,006	0,322
	$(Var(\hat{\theta}_d))^{-1}$	100,003	0,030	1,794	0,567	100,000	100,007	0,322
	$(CV(\hat{\theta}_d))^{-1}$	100,004	0,134	3,711	1,173	99,997	100,011	1,377
	$(skewness(s_{dc}))^{-1}$	100,003	-0,016	1,795	0,568	99,999	100,006	0,322
	$(kurtosis(k_{dc}))^{-1}$	100,003	-0,022	3,449	1,091	99,996	100,009	1,190
	$(K^2)^{-1}$	100,003	-0,022	1,602	0,506	99,999	100,006	0,256

$\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 17 N=5 000 000 N(100,49) n=5000 $\theta = 100.003$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	$SE(\hat{\theta}_c)$ $\times 10^{-2}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-3}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-3}$
						LB	UB	
1		100,002	-1,194	9,895	3,129	99,995	100,008	9,793
2	1/n	100,001	-1,752	6,992	2,211	99,997	100,005	4,892
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-1,730	6,992	2,211	99,997	100,005	4,891
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-1,689	6,992	2,211	99,997	100,005	4,891
	$(skewness(s_{dc}))^{-1}$	100,000	-2,437	7,857	2,485	99,995	100,005	6,180
	$(kurtosis(k_{dc}))^{-1}$	100,001	-1,748	6,993	2,211	99,997	100,005	4,893
3	1/n	100,000	-2,996	5,751	1,819	99,996	100,003	3,316
	$(Var(\hat{\theta}_a))^{-1}$	100,000	-2,974	5,750	1,818	99,996	100,003	3,316
	$(CV(\hat{\theta}_a))^{-1}$	100,000	-2,919	5,751	1,819	99,996	100,003	3,315
	$(skewness(s_{dc}))^{-1}$	99,999	-3,674	6,965	2,202	99,995	100,003	4,864
	$(kurtosis(k_{dc}))^{-1}$	100,000	-2,982	5,749	1,818	99,996	100,003	3,314
5	1/n	100,002	-0,879	4,488	1,419	99,999	100,005	2,015
	$(Var(\hat{\theta}_a))^{-1}$	100,002	-0,841	4,487	1,419	99,999	100,005	2,014
	$(CV(\hat{\theta}_a))^{-1}$	100,002	-0,782	4,487	1,419	99,999	100,005	2,014
	$(skewness(s_{dc}))^{-1}$	100,001	-2,020	6,206	1,963	99,997	100,005	3,855
	$(kurtosis(k_{dc}))^{-1}$	100,002	-0,898	4,487	1,419	99,999	100,005	2,014
10	1/n	100,000	-2,463	3,247	1,027	99,998	100,002	1,061
	$(Var(\hat{\theta}_a))^{-1}$	100,000	-2,433	3,247	1,027	99,998	100,002	1,060
	$(CV(\hat{\theta}_a))^{-1}$	100,000	-2,361	3,247	1,027	99,998	100,002	1,060
	$(skewness(s_{dc}))^{-1}$	100,000	-3,135	5,349	1,692	99,996	100,003	2,871
	$(kurtosis(k_{dc}))^{-1}$	100,000	-2,468	3,245	1,026	99,998	100,002	1,059
15	1/n	100,001	-1,665	2,620	0,829	99,999	100,003	0,689
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-1,629	2,618	0,828	100,000	100,003	0,688
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-1,556	2,619	0,828	100,000	100,003	0,688
	$(skewness(s_{dc}))^{-1}$	100,000	-3,143	4,928	1,558	99,997	100,003	2,438
	$(kurtosis(k_{dc}))^{-1}$	100,001	-1,676	2,619	0,828	99,999	100,003	0,689
20	1/n	100,001	-1,689	2,250	0,711	100,000	100,002	0,509
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-1,664	2,248	0,711	100,000	100,002	0,508
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-1,584	2,248	0,711	100,000	100,003	0,508
	$(skewness(s_{dc}))^{-1}$	100,000	-2,906	4,654	1,472	99,997	100,003	2,174
	$(kurtosis(k_{dc}))^{-1}$	100,001	-1,708	2,249	0,711	100,000	100,002	0,509

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, $SE(\hat{\theta}_c)$ =standard error of the combined estimate, $CV(\hat{\theta}_c)$ (%) = coefficient of variation, $CI_{95\%}(\hat{\theta}_c)$ = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 18: N=5 000 000 N(100,49) n=10000 $\theta = 100.003$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	$SE(\hat{\theta}_c)$ $\times 10^{-2}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-3}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-3}$
						LB	UB	
1		100,004	0,782	7,148	2,261	99,999	100,008	5,111
2	1/n	100,001	-1,660	5,040	1,594	99,998	100,004	2,543
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-1,644	5,037	1,593	99,998	100,004	2,540
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-1,628	5,038	1,593	99,998	100,004	2,541
	$(skewness(s_{dc}))^{-1}$	100,002	-0,772	5,538	1,751	99,999	100,005	3,068
	$(kurtosis(k_{dc}))^{-1}$	100,001	-1,644	5,036	1,593	99,998	100,004	2,539
3	1/n	100,001	-1,749	4,088	1,293	99,998	100,004	1,674
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-1,749	4,086	1,292	99,998	100,004	1,673
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-1,715	4,087	1,292	99,999	100,004	1,673
	$(skewness(s_{dc}))^{-1}$	100,002	-1,012	4,950	1,565	99,999	100,005	2,451
	$(kurtosis(k_{dc}))^{-1}$	100,001	-1,745	4,084	1,292	99,998	100,004	1,671
5	1/n	100,001	-1,692	3,126	0,989	99,999	100,003	0,980
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-1,696	3,127	0,989	99,999	100,003	0,981
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-1,654	3,126	0,989	99,999	100,003	0,980
	$(skewness(s_{dc}))^{-1}$	100,003	-0,242	4,420	1,398	100,000	100,005	1,954
	$(kurtosis(k_{dc}))^{-1}$	100,001	-1,707	3,125	0,988	99,999	100,003	0,980
10	1/n	100,002	-0,549	2,241	0,709	100,001	100,004	0,503
	$(Var(\hat{\theta}_a))^{-1}$	100,002	-0,551	2,241	0,709	100,001	100,004	0,503
	$(CV(\hat{\theta}_a))^{-1}$	100,002	-0,506	2,241	0,709	100,001	100,004	0,502
	$(skewness(s_{dc}))^{-1}$	100,002	-0,285	3,721	1,177	100,000	100,005	1,384
	$(kurtosis(k_{dc}))^{-1}$	100,002	-0,569	2,241	0,709	100,001	100,004	0,503
15	1/n	100,003	-0,164	1,838	0,581	100,001	100,004	0,338
	$(Var(\hat{\theta}_a))^{-1}$	100,003	-0,162	1,838	0,581	100,001	100,004	0,338
	$(CV(\hat{\theta}_a))^{-1}$	100,003	-0,117	1,838	0,581	100,002	100,004	0,338
	$(skewness(s_{dc}))^{-1}$	100,003	0,428	3,455	1,093	100,001	100,005	1,194
	$(kurtosis(k_{dc}))^{-1}$	100,003	-0,180	1,838	0,581	100,001	100,004	0,338
20	1/n	100,003	0,119	1,552	0,491	100,002	100,004	0,241
	$(Var(\hat{\theta}_a))^{-1}$	100,003	0,120	1,553	0,491	100,002	100,004	0,241
	$(CV(\hat{\theta}_a))^{-1}$	100,003	0,166	1,552	0,491	100,002	100,004	0,241
	$(skewness(s_{dc}))^{-1}$	100,004	1,136	3,323	1,051	100,002	100,006	1,105
	$(kurtosis(k_{dc}))^{-1}$	100,003	0,105	1,552	0,491	100,002	100,004	0,241

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, $SE(\hat{\theta}_c)$ =standard error of the combined estimate, $CV(\hat{\theta}_c)$ (%) = coefficient of variation, $CI_{95\%}(\hat{\theta}_c)$ = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 19: N=100 000 N(100,81) n=1000 $\theta = 100.03$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		100,035	5,696	9,035	2,856	100,017	100,053	8,166
2	1/n	100,031	1,740	6,223	1,967	100,019	100,043	3,873
	$(Var(\hat{\theta}_a))^{-1}$	100,031	1,732	6,214	1,965	100,019	100,043	3,862
	$(CV(\hat{\theta}_a))^{-1}$	100,032	2,153	6,217	1,965	100,019	100,044	3,866
	$(skewness(s_{dc}))^{-1}$	100,033	3,136	7,151	2,260	100,019	100,047	5,114
	$(kurtosis(k_{dc}))^{-1}$	100,031	1,314	6,241	1,973	100,019	100,043	3,895
	$(K^2)^{-1}$	100,032	2,420	7,109	2,247	100,018	100,046	5,055
3	1/n	100,033	3,120	5,174	1,636	100,022	100,043	2,678
	$(Var(\hat{\theta}_a))^{-1}$	100,033	3,127	5,184	1,639	100,022	100,043	2,688
	$(CV(\hat{\theta}_a))^{-1}$	100,033	3,669	5,178	1,637	100,023	100,043	2,682
	$(skewness(s_{dc}))^{-1}$	100,035	5,258	6,317	1,997	100,022	100,047	3,994
	$(kurtosis(k_{dc}))^{-1}$	100,032	2,828	5,187	1,640	100,022	100,042	2,692
	$(K^2)^{-1}$	100,036	6,665	6,388	2,019	100,024	100,049	4,085
5	1/n	100,029	-0,765	4,071	1,287	100,021	100,037	1,657
	$(Var(\hat{\theta}_a))^{-1}$	100,029	-0,609	4,074	1,288	100,021	100,037	1,660
	$(CV(\hat{\theta}_a))^{-1}$	100,029	-0,039	4,072	1,287	100,021	100,037	1,658
	$(skewness(s_{dc}))^{-1}$	100,029	-0,323	5,588	1,767	100,018	100,040	3,123
	$(kurtosis(k_{dc}))^{-1}$	100,028	-1,002	4,074	1,288	100,021	100,036	1,659
	$(K^2)^{-1}$	100,031	1,024	5,523	1,746	100,020	100,041	3,051
10	1/n	100,028	-1,548	2,825	0,893	100,022	100,033	0,798
	$(Var(\hat{\theta}_a))^{-1}$	100,028	-1,496	2,827	0,894	100,022	100,034	0,800
	$(CV(\hat{\theta}_a))^{-1}$	100,029	-0,801	2,826	0,893	100,023	100,034	0,799
	$(skewness(s_{dc}))^{-1}$	100,029	-0,412	4,686	1,481	100,020	100,038	2,196
	$(kurtosis(k_{dc}))^{-1}$	100,028	-1,769	2,825	0,893	100,022	100,033	0,798
	$(K^2)^{-1}$	100,032	2,511	4,545	1,437	100,023	100,041	2,067
15	1/n	100,029	-0,798	2,308	0,730	100,024	100,033	0,533
	$(Var(\hat{\theta}_a))^{-1}$	100,029	-0,786	2,309	0,730	100,024	100,033	0,533
	$(CV(\hat{\theta}_a))^{-1}$	100,029	-0,049	2,308	0,730	100,025	100,034	0,533
	$(skewness(s_{dc}))^{-1}$	100,029	-0,234	4,405	1,393	100,021	100,038	1,941
	$(kurtosis(k_{dc}))^{-1}$	100,028	-1,049	2,308	0,730	100,024	100,033	0,533
	$(K^2)^{-1}$	100,029	-0,232	4,087	1,292	100,021	100,037	1,670
20	1/n	100,029	-0,152	1,983	0,627	100,025	100,033	0,393
	$(Var(\hat{\theta}_a))^{-1}$	100,029	-0,140	1,985	0,628	100,025	100,033	0,394
	$(CV(\hat{\theta}_a))^{-1}$	100,030	0,622	1,984	0,627	100,026	100,034	0,394
	$(skewness(s_{dc}))^{-1}$	100,031	1,312	4,238	1,340	100,022	100,039	1,796
	$(kurtosis(k_{dc}))^{-1}$	100,029	-0,421	1,983	0,627	100,025	100,033	0,393
	$(K^2)^{-1}$	100,029	-0,712	3,870	1,224	100,021	100,036	1,498

$\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 20: N=100 000 N(100,81) n=5000 $\theta = 100.03$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	$SE(\hat{\theta}_c)$ $\times 10^{-1}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-2}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-2}$
						LB	UB	
1		100,031	1,396	1,217	3,849	100,023	100,038	1,482
2	1/n	100,030	0,270	0,854	2,703	100,024	100,035	0,730
	$(Var(\hat{\theta}_a))^{-1}$	100,030	0,297	0,854	2,702	100,024	100,035	0,730
	$(CV(\hat{\theta}_a))^{-1}$	100,030	0,357	0,854	2,702	100,025	100,035	0,730
	$(skewness(s_{dc}))^{-1}$	100,029	-0,356	0,975	3,084	100,023	100,035	0,951
	$(kurtosis(k_{dc}))^{-1}$	100,030	0,245	0,855	2,704	100,024	100,035	0,731
3	1/n	100,031	1,162	0,699	2,210	100,026	100,035	0,489
	$(Var(\hat{\theta}_a))^{-1}$	100,031	1,187	0,699	2,210	100,026	100,035	0,488
	$(CV(\hat{\theta}_a))^{-1}$	100,031	1,276	0,698	2,210	100,026	100,035	0,488
	$(skewness(s_{dc}))^{-1}$	100,031	1,220	0,886	2,803	100,025	100,036	0,786
	$(kurtosis(k_{dc}))^{-1}$	100,031	1,033	0,699	2,210	100,026	100,035	0,489
5	1/n	100,031	1,047	0,544	1,721	100,027	100,034	0,296
	$(Var(\hat{\theta}_a))^{-1}$	100,031	1,100	0,544	1,721	100,027	100,034	0,296
	$(CV(\hat{\theta}_a))^{-1}$	100,031	1,198	0,544	1,721	100,027	100,034	0,296
	$(skewness(s_{dc}))^{-1}$	100,032	2,021	0,783	2,477	100,027	100,036	0,614
	$(kurtosis(k_{dc}))^{-1}$	100,030	0,982	0,545	1,723	100,027	100,034	0,297
10	1/n	100,030	0,238	0,385	1,218	100,027	100,032	0,148
	$(Var(\hat{\theta}_a))^{-1}$	100,030	0,227	0,385	1,218	100,027	100,032	0,148
	$(CV(\hat{\theta}_a))^{-1}$	100,030	0,372	0,385	1,218	100,027	100,032	0,148
	$(skewness(s_{dc}))^{-1}$	100,027	-2,587	0,672	2,125	100,023	100,031	0,452
	$(kurtosis(k_{dc}))^{-1}$	100,030	0,202	0,385	1,219	100,027	100,032	0,149
15	1/n	100,030	0,749	0,320	1,013	100,028	100,032	0,103
	$(Var(\hat{\theta}_a))^{-1}$	100,030	0,764	0,321	1,014	100,028	100,032	0,103
	$(CV(\hat{\theta}_a))^{-1}$	100,030	0,900	0,320	1,013	100,028	100,032	0,103
	$(skewness(s_{dc}))^{-1}$	100,027	-2,087	0,610	1,931	100,024	100,031	0,373
	$(kurtosis(k_{dc}))^{-1}$	100,030	0,698	0,320	1,013	100,028	100,032	0,103
20	1/n	100,029	-0,216	0,286	0,905	100,027	100,031	0,082
	$(Var(\hat{\theta}_a))^{-1}$	100,029	-0,196	0,287	0,906	100,028	100,031	0,082
	$(CV(\hat{\theta}_a))^{-1}$	100,029	-0,059	0,286	0,906	100,028	100,031	0,082
	$(skewness(s_{dc}))^{-1}$	100,028	-1,988	0,602	1,903	100,024	100,031	0,363
	$(kurtosis(k_{dc}))^{-1}$	100,029	-0,262	0,286	0,906	100,027	100,031	0,082

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, $SE(\hat{\theta}_c)$ =standard error of the combined estimate, $CV(\hat{\theta}_c)$ (%) = coefficient of variation, $CI_{95\%}(\hat{\theta}_c)$ = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 21: N=100 000 N(10,81) n=5000 $\theta = 100.03$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	$SE(\hat{\theta}_c)$ $\times 10^{-2}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-2}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-2}$
						LB	UB	
1		100,032	2,276	8,430	2,667	100,027	100,037	0,712
2	1/n	100,029	-0,905	6,030	1,907	100,025	100,032	0,364
	$(Var(\hat{\theta}_a))^{-1}$	100,029	-0,921	6,031	1,908	100,025	100,032	0,364
	$(CV(\hat{\theta}_a))^{-1}$	100,029	-0,877	6,030	1,908	100,025	100,032	0,364
	$(skewness(s_{dc}))^{-1}$	100,029	-0,308	6,804	2,152	100,025	100,033	0,463
	$(kurtosis(k_{dc}))^{-1}$	100,029	-0,902	6,028	1,907	100,025	100,032	0,364
3	1/n	100,028	-1,301	4,901	1,550	100,025	100,031	0,241
	$(Var(\hat{\theta}_a))^{-1}$	100,028	-1,291	4,900	1,550	100,025	100,031	0,240
	$(CV(\hat{\theta}_a))^{-1}$	100,028	-1,249	4,900	1,550	100,025	100,031	0,240
	$(skewness(s_{dc}))^{-1}$	100,029	-0,693	5,979	1,891	100,025	100,033	0,358
	$(kurtosis(k_{dc}))^{-1}$	100,028	-1,297	4,903	1,551	100,025	100,031	0,241
5	1/n	100,029	-0,746	3,763	1,190	100,026	100,031	0,142
	$(Var(\hat{\theta}_a))^{-1}$	100,029	-0,732	3,762	1,190	100,026	100,031	0,142
	$(CV(\hat{\theta}_a))^{-1}$	100,029	-0,681	3,762	1,190	100,026	100,031	0,142
	$(skewness(s_{dc}))^{-1}$	100,030	0,278	5,324	1,684	100,026	100,033	0,284
	$(kurtosis(k_{dc}))^{-1}$	100,029	-0,751	3,764	1,191	100,026	100,031	0,142
10	1/n	100,029	-0,259	2,697	0,853	100,028	100,031	0,073
	$(Var(\hat{\theta}_a))^{-1}$	100,029	-0,247	2,696	0,853	100,028	100,031	0,073
	$(CV(\hat{\theta}_a))^{-1}$	100,029	-0,187	2,696	0,853	100,028	100,031	0,073
	$(skewness(s_{dc}))^{-1}$	100,031	1,264	4,562	1,443	100,028	100,034	0,208
	$(kurtosis(k_{dc}))^{-1}$	100,029	-0,270	2,697	0,853	100,028	100,031	0,073
15	1/n	100,029	-0,360	2,210	0,699	100,028	100,031	0,049
	$(Var(\hat{\theta}_a))^{-1}$	100,029	-0,354	2,210	0,699	100,028	100,031	0,049
	$(CV(\hat{\theta}_a))^{-1}$	100,029	-0,288	2,210	0,699	100,028	100,031	0,049
	$(skewness(s_{dc}))^{-1}$	100,031	1,720	4,334	1,371	100,029	100,034	0,188
	$(kurtosis(k_{dc}))^{-1}$	100,029	-0,366	2,210	0,699	100,028	100,030	0,049
20	1/n	100,029	-0,278	1,911	0,605	100,028	100,030	0,037
	$(Var(\hat{\theta}_a))^{-1}$	100,029	-0,266	1,911	0,605	100,028	100,030	0,037
	$(CV(\hat{\theta}_a))^{-1}$	100,029	-0,203	1,911	0,605	100,028	100,030	0,037
	$(skewness(s_{dc}))^{-1}$	100,032	2,512	4,060	1,284	100,029	100,035	0,166
	$(kurtosis(k_{dc}))^{-1}$	100,029	-0,288	1,911	0,604	100,028	100,030	0,037

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, $SE(\hat{\theta}_c)$ =standard error of the combined estimate, $CV(\hat{\theta}_c)$ (%) = coefficient of variation, $CI_{95\%}(\hat{\theta}_c)$ = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 22: N=1 000 000 N(100,81) n=1000 $\theta = 99.995$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		100,004	0,876	8,674	2,743	99,987	100,021	7,531
2	1/n	100,005	1,026	6,617	2,092	99,992	100,018	4,389
	$(Var(\hat{\theta}_a))^{-1}$	100,005	1,040	6,604	2,088	99,992	100,018	4,372
	$(CV(\hat{\theta}_a))^{-1}$	100,006	1,068	6,610	2,090	99,993	100,019	4,380
	$(skewness(s_{dc}))^{-1}$	99,999	0,442	7,356	2,326	99,985	100,014	5,412
	$(kurtosis(k_{dc}))^{-1}$	100,005	1,029	6,604	2,088	99,992	100,018	4,372
	$(K^2)^{-1}$	100,002	0,668	7,398	2,339	99,987	100,016	5,478
3	1/n	100,007	1,173	5,356	1,694	99,996	100,017	2,883
	$(Var(\hat{\theta}_a))^{-1}$	100,007	1,207	5,361	1,695	99,996	100,017	2,889
	$(CV(\hat{\theta}_a))^{-1}$	100,007	1,244	5,358	1,694	99,997	100,018	2,886
	$(skewness(s_{dc}))^{-1}$	100,000	0,478	6,458	2,042	99,987	100,012	4,173
	$(kurtosis(k_{dc}))^{-1}$	100,006	1,156	5,347	1,691	99,996	100,017	2,872
	$(K^2)^{-1}$	100,002	0,710	6,532	2,066	99,989	100,015	4,272
5	1/n	99,997	0,250	4,188	1,324	99,989	100,006	1,754
	$(Var(\hat{\theta}_a))^{-1}$	99,998	0,277	4,197	1,327	99,989	100,006	1,762
	$(CV(\hat{\theta}_a))^{-1}$	99,998	0,329	4,192	1,326	99,990	100,006	1,758
	$(skewness(s_{dc}))^{-1}$	99,993	-0,228	5,872	1,857	99,981	100,004	3,449
	$(kurtosis(k_{dc}))^{-1}$	99,997	0,223	4,189	1,325	99,989	100,005	1,755
	$(K^2)^{-1}$	99,990	-0,502	5,719	1,809	99,979	100,001	3,274
10	1/n	99,999	0,381	2,898	0,917	99,993	100,004	0,841
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,393	2,901	0,918	99,993	100,004	0,843
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,460	2,900	0,917	99,994	100,005	0,843
	$(skewness(s_{dc}))^{-1}$	99,992	-0,312	4,926	1,558	99,982	100,001	2,428
	$(kurtosis(k_{dc}))^{-1}$	99,999	0,372	2,895	0,916	99,993	100,004	0,840
	$(K^2)^{-1}$	99,990	-0,459	4,844	1,532	99,981	100,000	2,348
15	1/n	99,999	0,372	2,294	0,725	99,994	100,003	0,528
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,385	2,293	0,725	99,994	100,003	0,527
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,455	2,293	0,725	99,995	100,004	0,528
	$(skewness(s_{dc}))^{-1}$	99,991	-0,431	4,351	1,376	99,982	99,999	1,895
	$(kurtosis(k_{dc}))^{-1}$	99,998	0,349	2,295	0,726	99,994	100,003	0,528
	$(K^2)^{-1}$	99,994	-0,059	4,235	1,339	99,986	100,003	1,793
20	1/n	99,998	0,305	2,027	0,641	99,994	100,002	0,412
	$(Var(\hat{\theta}_a))^{-1}$	99,998	0,311	2,028	0,641	99,994	100,002	0,412
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,385	2,027	0,641	99,995	100,003	0,412
	$(skewness(s_{dc}))^{-1}$	99,989	-0,603	4,322	1,367	99,980	99,997	1,872
	$(kurtosis(k_{dc}))^{-1}$	99,998	0,290	2,026	0,641	99,994	100,002	0,411
	$(K^2)^{-1}$	99,995	0,044	3,888	1,229	99,988	100,003	1,511

$\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 23 N=1 000 000 N(100,81) n=5000 $\theta = 99.995$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	$SE(\hat{\theta}_c)$ $\times 10^{-1}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-3}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-2}$
						LB	UB	
1		99,998	3,264	1,310	4,144	99,990	100,006	1,718
2	1/n	99,995	-0,372	0,929	2,938	99,989	100,000	0,863
	$(Var(\hat{\theta}_a))^{-1}$	99,994	-0,382	0,929	2,938	99,989	100,000	0,863
	$(CV(\hat{\theta}_a))^{-1}$	99,995	-0,296	0,929	2,938	99,989	100,000	0,863
	$(skewness(s_{dc}))^{-1}$	99,994	-0,514	1,051	3,322	99,988	100,001	1,104
	$(kurtosis(k_{dc}))^{-1}$	99,994	-0,388	0,930	2,939	99,989	100,000	0,864
3	1/n	99,995	0,192	0,751	2,373	99,990	100,000	0,563
	$(Var(\hat{\theta}_a))^{-1}$	99,995	0,155	0,751	2,373	99,990	100,000	0,563
	$(CV(\hat{\theta}_a))^{-1}$	99,995	0,287	0,750	2,373	99,991	100,000	0,563
	$(skewness(s_{dc}))^{-1}$	99,996	0,657	0,922	2,915	99,990	100,001	0,849
	$(kurtosis(k_{dc}))^{-1}$	99,995	0,229	0,752	2,377	99,990	100,000	0,565
5	1/n	99,996	0,909	0,557	1,762	99,992	99,999	0,311
	$(Var(\hat{\theta}_a))^{-1}$	99,996	0,905	0,557	1,762	99,992	99,999	0,310
	$(CV(\hat{\theta}_a))^{-1}$	99,996	1,037	0,557	1,762	99,992	99,999	0,310
	$(skewness(s_{dc}))^{-1}$	99,996	1,537	0,793	2,508	99,991	100,001	0,629
	$(kurtosis(k_{dc}))^{-1}$	99,996	0,932	0,558	1,764	99,992	99,999	0,311
10	1/n	99,995	-0,028	0,391	1,235	99,992	99,997	0,153
	$(Var(\hat{\theta}_a))^{-1}$	99,995	0,011	0,391	1,236	99,992	99,997	0,153
	$(CV(\hat{\theta}_a))^{-1}$	99,995	0,136	0,391	1,236	99,993	99,997	0,153
	$(skewness(s_{dc}))^{-1}$	99,994	-1,143	0,666	2,105	99,990	99,998	0,443
	$(kurtosis(k_{dc}))^{-1}$	99,995	-0,021	0,391	1,237	99,992	99,997	0,153
15	1/n	99,996	0,909	0,323	1,023	99,994	99,998	0,105
	$(Var(\hat{\theta}_a))^{-1}$	99,996	0,939	0,323	1,023	99,994	99,998	0,105
	$(CV(\hat{\theta}_a))^{-1}$	99,996	1,073	0,323	1,023	99,994	99,998	0,105
	$(skewness(s_{dc}))^{-1}$	99,994	-0,690	0,614	1,942	99,990	99,998	0,377
	$(kurtosis(k_{dc}))^{-1}$	99,996	0,912	0,324	1,024	99,994	99,998	0,105
20	1/n	99,996	0,750	0,276	0,872	99,994	99,997	0,076
	$(Var(\hat{\theta}_a))^{-1}$	99,996	0,780	0,275	0,871	99,994	99,997	0,076
	$(CV(\hat{\theta}_a))^{-1}$	99,996	0,917	0,276	0,871	99,994	99,997	0,076
	$(skewness(s_{dc}))^{-1}$	99,995	0,024	0,596	1,884	99,991	99,999	0,355
	$(kurtosis(k_{dc}))^{-1}$	99,996	0,753	0,276	0,873	99,994	99,997	0,076

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, $SE(\hat{\theta}_c)$ =standard error of the combined estimate, $CV(\hat{\theta}_c)$ (%) = coefficient of variation, $CI_{95\%}(\hat{\theta}_c)$ = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 24: N=1 000 000 N(100,81) n=10000 $\theta = 99.995$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		99,991	-4,356	2,833	8,961	99,985	99,996	8,048
2	1/n	99,993	-1,656	2,054	6,497	99,989	99,997	4,223
	$(Var(\hat{\theta}_a))^{-1}$	99,993	-1,636	2,055	6,498	99,989	99,997	4,225
	$(CV(\hat{\theta}_a))^{-1}$	99,993	-1,605	2,054	6,497	99,989	99,997	4,224
	$(skewness(s_{dc}))^{-1}$	99,994	-1,321	2,246	7,103	99,989	99,998	5,046
	$(kurtosis(k_{dc}))^{-1}$	99,993	-1,641	2,055	6,499	99,989	99,997	4,226
3	1/n	99,994	-0,620	1,652	5,224	99,991	99,997	2,729
	$(Var(\hat{\theta}_a))^{-1}$	99,994	-0,589	1,652	5,224	99,991	99,998	2,729
	$(CV(\hat{\theta}_a))^{-1}$	99,994	-0,550	1,652	5,224	99,991	99,998	2,729
	$(skewness(s_{dc}))^{-1}$	99,993	-1,768	1,958	6,193	99,989	99,997	3,837
	$(kurtosis(k_{dc}))^{-1}$	99,994	-0,583	1,652	5,225	99,991	99,998	2,730
5	1/n	99,995	0,297	1,295	4,096	99,993	99,998	1,678
	$(Var(\hat{\theta}_a))^{-1}$	99,995	0,324	1,295	4,094	99,993	99,998	1,676
	$(CV(\hat{\theta}_a))^{-1}$	99,995	0,374	1,295	4,095	99,993	99,998	1,677
	$(skewness(s_{dc}))^{-1}$	99,994	-0,687	1,756	5,553	99,991	99,998	3,084
	$(kurtosis(k_{dc}))^{-1}$	99,995	0,309	1,295	4,095	99,993	99,998	1,677
10	1/n	99,995	0,066	0,885	2,800	99,993	99,997	0,784
	$(Var(\hat{\theta}_a))^{-1}$	99,995	0,072	0,885	2,799	99,993	99,997	0,783
	$(CV(\hat{\theta}_a))^{-1}$	99,995	0,140	0,885	2,799	99,993	99,997	0,783
	$(skewness(s_{dc}))^{-1}$	99,994	-0,737	1,499	4,740	99,991	99,997	2,247
	$(kurtosis(k_{dc}))^{-1}$	99,995	0,083	0,885	2,799	99,993	99,997	0,784
15	1/n	99,995	0,349	0,726	2,296	99,994	99,997	0,527
	$(Var(\hat{\theta}_a))^{-1}$	99,995	0,356	0,726	2,295	99,994	99,997	0,527
	$(CV(\hat{\theta}_a))^{-1}$	99,995	0,427	0,726	2,295	99,994	99,997	0,527
	$(skewness(s_{dc}))^{-1}$	99,995	0,115	1,449	4,584	99,992	99,998	2,101
	$(kurtosis(k_{dc}))^{-1}$	99,995	0,362	0,726	2,296	99,994	99,997	0,527
20	1/n	99,995	0,188	0,636	2,012	99,994	99,996	0,405
	$(Var(\hat{\theta}_a))^{-1}$	99,995	0,192	0,636	2,011	99,994	99,996	0,404
	$(CV(\hat{\theta}_a))^{-1}$	99,995	0,267	0,636	2,011	99,994	99,996	0,404
	$(skewness(s_{dc}))^{-1}$	99,994	-0,431	1,395	4,412	99,992	99,997	1,946
	$(kurtosis(k_{dc}))^{-1}$	99,995	0,199	0,636	2,013	99,994	99,996	0,405

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE($\hat{\theta}_c$)=standard error of the combined estimate, CV($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 25: N=5 000 000 N(100,81) n=1000 $\theta = 100.004$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		100,021	1,735	9,261	2,928	100,003	100,039	8,607
2	1/n	100,008	0,423	6,531	2,065	99,995	100,021	4,267
	$(Var(\hat{\theta}_a))^{-1}$	100,008	0,438	6,548	2,071	99,995	100,021	4,290
	$(CV(\hat{\theta}_a))^{-1}$	100,008	0,473	6,539	2,068	99,995	100,021	4,279
	$(skewness(s_{dc}))^{-1}$	100,013	0,906	7,492	2,369	99,998	100,027	5,621
	$(kurtosis(k_{dc}))^{-1}$	100,008	0,428	6,525	2,063	99,995	100,021	4,259
	$(K^2)^{-1}$	100,012	0,809	7,532	2,382	99,997	100,026	5,680
3	1/n	100,009	0,582	5,335	1,687	99,999	100,020	2,850
	$(Var(\hat{\theta}_a))^{-1}$	100,009	0,581	5,345	1,690	99,999	100,020	2,860
	$(CV(\hat{\theta}_a))^{-1}$	100,010	0,639	5,340	1,689	99,999	100,020	2,856
	$(skewness(s_{dc}))^{-1}$	100,007	0,345	6,645	2,101	99,994	100,020	4,417
	$(kurtosis(k_{dc}))^{-1}$	100,009	0,592	5,332	1,686	99,999	100,020	2,847
	$(K^2)^{-1}$	100,010	0,683	6,668	2,108	99,997	100,023	4,451
5	1/n	100,010	0,687	4,066	1,286	100,002	100,018	1,658
	$(Var(\hat{\theta}_a))^{-1}$	100,011	0,704	4,071	1,287	100,003	100,019	1,662
	$(CV(\hat{\theta}_a))^{-1}$	100,011	0,762	4,068	1,286	100,003	100,019	1,660
	$(skewness(s_{dc}))^{-1}$	100,006	0,217	5,968	1,887	99,994	100,017	3,563
	$(kurtosis(k_{dc}))^{-1}$	100,010	0,695	4,069	1,287	100,003	100,018	1,661
	$(K^2)^{-1}$	100,007	0,356	5,833	1,845	99,996	100,019	3,404
10	1/n	100,005	0,168	2,800	0,885	100,000	100,011	0,784
	$(Var(\hat{\theta}_a))^{-1}$	100,005	0,170	2,804	0,887	100,000	100,011	0,787
	$(CV(\hat{\theta}_a))^{-1}$	100,006	0,245	2,802	0,886	100,001	100,011	0,786
	$(skewness(s_{dc}))^{-1}$	100,008	0,448	4,986	1,577	99,998	100,018	2,488
	$(kurtosis(k_{dc}))^{-1}$	100,005	0,189	2,802	0,886	100,000	100,011	0,785
	$(K^2)^{-1}$	100,007	0,352	4,873	1,541	99,998	100,017	2,376
15	1/n	100,003	-0,028	2,305	0,729	99,999	100,008	0,531
	$(Var(\hat{\theta}_a))^{-1}$	100,003	-0,018	2,309	0,730	99,999	100,008	0,533
	$(CV(\hat{\theta}_a))^{-1}$	100,004	0,055	2,307	0,729	100,000	100,009	0,532
	$(skewness(s_{dc}))^{-1}$	100,005	0,172	4,771	1,509	99,996	100,015	2,277
	$(kurtosis(k_{dc}))^{-1}$	100,003	-0,021	2,307	0,730	99,999	100,008	0,532
	$(K^2)^{-1}$	100,003	-0,029	4,435	1,402	99,995	100,012	1,967
20	1/n	100,003	-0,029	2,059	0,651	99,999	100,007	0,424
	$(Var(\hat{\theta}_a))^{-1}$	100,003	-0,025	2,062	0,652	99,999	100,007	0,425
	$(CV(\hat{\theta}_a))^{-1}$	100,004	0,052	2,060	0,652	100,000	100,008	0,425
	$(skewness(s_{dc}))^{-1}$	100,005	0,195	4,659	1,473	99,996	100,015	2,171
	$(kurtosis(k_{dc}))^{-1}$	100,003	-0,022	2,059	0,651	99,999	100,007	0,424
	$(K^2)^{-1}$	100,001	-0,227	4,066	1,286	99,993	100,009	1,654

$\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 26 N=5 000 000 N(100,81) n=5000 $\theta = 100.004$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	$SE(\hat{\theta}_c)$ $\times 10^{-1}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-3}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-2}$
						LB	UB	
1		100,002	-1,535	1,272	4,023	99,994	100,010	1,619
2	1/n	100,001	-2,252	0,899	2,843	99,996	100,007	0,809
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-2,225	0,899	2,843	99,996	100,007	0,809
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-2,153	0,899	2,843	99,996	100,007	0,809
	$(skewness(s_{dc}))^{-1}$	100,000	-3,133	1,010	3,195	99,994	100,007	1,022
	$(kurtosis(k_{dc}))^{-1}$	100,001	-2,247	0,899	2,843	99,996	100,007	0,809
3	1/n	100,000	-3,852	0,739	2,338	99,995	100,004	0,548
	$(Var(\hat{\theta}_a))^{-1}$	100,000	-3,823	0,739	2,338	99,995	100,004	0,548
	$(CV(\hat{\theta}_a))^{-1}$	100,000	-3,729	0,739	2,338	99,995	100,004	0,548
	$(skewness(s_{dc}))^{-1}$	99,999	-4,723	0,895	2,832	99,993	100,004	0,804
	$(kurtosis(k_{dc}))^{-1}$	100,000	-3,834	0,739	2,337	99,995	100,004	0,548
5	1/n	100,002	-1,130	0,577	1,825	99,999	100,006	0,333
	$(Var(\hat{\theta}_a))^{-1}$	100,002	-1,082	0,577	1,825	99,999	100,006	0,333
	$(CV(\hat{\theta}_a))^{-1}$	100,003	-0,976	0,577	1,825	99,999	100,006	0,333
	$(skewness(s_{dc}))^{-1}$	100,001	-2,597	0,798	2,523	99,996	100,006	0,637
	$(kurtosis(k_{dc}))^{-1}$	100,002	-1,155	0,577	1,824	99,999	100,006	0,333
10	1/n	100,000	-3,167	0,418	1,320	99,998	100,003	0,175
	$(Var(\hat{\theta}_a))^{-1}$	100,000	-3,128	0,417	1,320	99,998	100,003	0,175
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-3,003	0,417	1,320	99,998	100,003	0,175
	$(skewness(s_{dc}))^{-1}$	100,000	-4,031	0,688	2,175	99,995	100,004	0,475
	$(kurtosis(k_{dc}))^{-1}$	100,000	-3,173	0,417	1,319	99,998	100,003	0,175
15	1/n	100,001	-2,141	0,337	1,065	99,999	100,003	0,114
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-2,094	0,337	1,065	99,999	100,004	0,114
	$(CV(\hat{\theta}_a))^{-1}$	100,002	-1,968	0,337	1,065	99,999	100,004	0,114
	$(skewness(s_{dc}))^{-1}$	100,000	-4,041	0,634	2,003	99,996	100,003	0,403
	$(kurtosis(k_{dc}))^{-1}$	100,001	-2,154	0,337	1,065	99,999	100,003	0,114
20	1/n	100,001	-2,171	0,289	0,915	100,000	100,003	0,084
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-2,140	0,289	0,914	100,000	100,003	0,084
	$(CV(\hat{\theta}_a))^{-1}$	100,002	-2,003	0,289	0,914	100,000	100,003	0,084
	$(skewness(s_{dc}))^{-1}$	100,000	-3,736	0,598	1,892	99,996	100,004	0,359
	$(kurtosis(k_{dc}))^{-1}$	100,001	-2,196	0,289	0,914	100,000	100,003	0,084

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, $SE(\hat{\theta}_c)$ =standard error of the combined estimate, $CV(\hat{\theta}_c)$ (%) = coefficient of variation, $CI_{95\%}(\hat{\theta}_c)$ = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 27: N=5 000 000 N(100,81) n=10000 $\theta = 100.004$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	$SE(\hat{\theta}_c)$ $\times 10^{-2}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-3}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-3}$
						LB	UB	
1		100,005	1,006	9,191	2,906	99,999	100,010	8,449
2	1/n	100,001	-2,135	6,480	2,049	99,997	100,005	4,203
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-2,113	6,476	2,048	99,997	100,005	4,199
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-2,084	6,478	2,049	99,997	100,005	4,201
	$(skewness(s_{dc}))^{-1}$	100,003	-0,993	7,120	2,252	99,998	100,007	5,071
	$(kurtosis(k_{dc}))^{-1}$	100,001	-2,114	6,475	2,048	99,997	100,005	4,197
3	1/n	100,001	-2,248	5,255	1,662	99,998	100,005	2,767
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-2,249	5,254	1,661	99,998	100,005	2,765
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-2,193	5,254	1,662	99,998	100,005	2,766
	$(skewness(s_{dc}))^{-1}$	100,002	-1,301	6,364	2,013	99,998	100,006	4,052
	$(kurtosis(k_{dc}))^{-1}$	100,001	-2,243	5,251	1,661	99,998	100,005	2,763
5	1/n	100,001	-2,175	4,020	1,271	99,999	100,004	1,621
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-2,181	4,020	1,271	99,999	100,004	1,621
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-2,112	4,020	1,271	99,999	100,004	1,620
	$(skewness(s_{dc}))^{-1}$	100,003	-0,311	5,682	1,797	100,000	100,007	3,229
	$(kurtosis(k_{dc}))^{-1}$	100,001	-2,195	4,018	1,271	99,999	100,004	1,619
10	1/n	100,003	-0,706	2,881	0,911	100,001	100,005	0,831
	$(Var(\hat{\theta}_a))^{-1}$	100,003	-0,708	2,881	0,911	100,001	100,005	0,831
	$(CV(\hat{\theta}_a))^{-1}$	100,003	-0,634	2,881	0,911	100,001	100,005	0,831
	$(skewness(s_{dc}))^{-1}$	100,003	-0,367	4,783	1,513	100,000	100,006	2,288
	$(kurtosis(k_{dc}))^{-1}$	100,003	-0,731	2,882	0,911	100,001	100,005	0,831
15	1/n	100,003	-0,211	2,363	0,747	100,002	100,005	0,558
	$(Var(\hat{\theta}_a))^{-1}$	100,003	-0,208	2,363	0,747	100,002	100,005	0,558
	$(CV(\hat{\theta}_a))^{-1}$	100,003	-0,134	2,363	0,747	100,002	100,005	0,558
	$(skewness(s_{dc}))^{-1}$	100,004	0,551	4,442	1,405	100,001	100,007	1,973
	$(kurtosis(k_{dc}))^{-1}$	100,003	-0,231	2,364	0,747	100,002	100,005	0,559
20	1/n	100,004	0,153	1,995	0,631	100,002	100,005	0,398
	$(Var(\hat{\theta}_a))^{-1}$	100,004	0,155	1,996	0,631	100,002	100,005	0,399
	$(CV(\hat{\theta}_a))^{-1}$	100,004	0,230	1,996	0,631	100,003	100,005	0,398
	$(skewness(s_{dc}))^{-1}$	100,005	1,461	4,272	1,351	100,002	100,008	1,827
	$(kurtosis(k_{dc}))^{-1}$	100,004	0,135	1,996	0,631	100,002	100,005	0,398

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, $SE(\hat{\theta}_c)$ =standard error of the combined estimate, $CV(\hat{\theta}_c)$ (%) = coefficient of variation, $CI_{95\%}(\hat{\theta}_c)$ = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Appendix A2

Comparing weighting strategies: Stratified Random Sampling: Normal distribution

An asterisk before a weighting strategy indicates that the estimator is not design consistent. That is, a naïve estimator that does not account for the sample design was used to weight surveys.

Scenario 1: N=1 000 000 N(100,25) n=1000 D=20, $\theta = 99.999$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,996	-3,524	3,471	1,098	99,989	100,002	1,206
2	1/n	99,997	-2,406	2,432	0,769	99,992	100,001	0,592
	$(Var(\hat{\theta}_a))^{-1}$	99,997	-1,839	2,457	0,777	99,992	100,002	0,604
	$*(Var(\hat{\theta}_a))^{-1}$	99,997	-1,982	2,454	0,776	99,992	100,002	0,603
	$(CV(\hat{\theta}_a))^{-1}$	99,997	-1,824	2,454	0,776	99,992	100,002	0,603
	$*(CV(\hat{\theta}_a))^{-1}$	99,997	-1,898	2,453	0,776	99,992	100,002	0,602
	$(skewness(s_{dc}))^{-1}$	99,998	-1,424	2,761	0,873	99,992	100,003	0,762
	$*(skewness(s_{dc}))^{-1}$	99,997	-1,630	2,748	0,869	99,992	100,003	0,755
	$(kurtosis(K_{dc}))^{-1}$	99,997	-1,899	2,456	0,777	99,992	100,002	0,604
	$*(kurtosis(K_{dc}))^{-1}$	99,997	-1,882	2,456	0,777	99,992	100,002	0,603
	$meff/Var(y)$	99,997	-1,984	2,454	0,776	99,992	100,002	0,603
	$(meff)^{-1}$	99,997	-1,792	2,456	0,777	99,992	100,002	0,604
3	1/n	100,000	0,982	1,912	0,605	99,996	100,004	0,366
	$(Var(\hat{\theta}_a))^{-1}$	100,001	1,559	1,924	0,608	99,997	100,004	0,370
	$*(Var(\hat{\theta}_a))^{-1}$	100,001	1,431	1,920	0,607	99,997	100,004	0,369
	$(CV(\hat{\theta}_a))^{-1}$	100,001	1,606	1,921	0,608	99,997	100,004	0,369
	$*(CV(\hat{\theta}_a))^{-1}$	100,001	1,540	1,919	0,607	99,997	100,004	0,369
	$(skewness(s_{dc}))^{-1}$	100,001	1,611	2,422	0,766	99,996	100,005	0,587
	$*(skewness(s_{dc}))^{-1}$	100,000	0,525	2,375	0,751	99,995	100,004	0,564
	$(kurtosis(K_{dc}))^{-1}$	100,001	1,489	1,919	0,607	99,997	100,004	0,368
	$*(kurtosis(K_{dc}))^{-1}$	100,001	1,503	1,919	0,607	99,997	100,004	0,368
	$meff/Var(y)$	100,000	1,374	1,926	0,609	99,997	100,004	0,371
	$(meff)^{-1}$	100,001	1,608	1,925	0,609	99,997	100,004	0,371
5	1/n	100,000	1,088	1,553	0,491	99,997	100,003	0,241
	$(Var(\hat{\theta}_a))^{-1}$	100,000	1,213	1,555	0,492	99,997	100,003	0,242
	$*(Var(\hat{\theta}_a))^{-1}$	100,000	1,108	1,553	0,491	99,997	100,003	0,241
	$(CV(\hat{\theta}_a))^{-1}$	100,000	1,245	1,554	0,491	99,997	100,003	0,241
	$*(CV(\hat{\theta}_a))^{-1}$	100,000	1,190	1,553	0,491	99,997	100,003	0,241
	$(skewness(s_{dc}))^{-1}$	100,000	0,516	2,177	0,688	99,995	100,004	0,474
	$*(skewness(s_{dc}))^{-1}$	99,999	-0,126	2,094	0,662	99,995	100,003	0,439
	$(kurtosis(K_{dc}))^{-1}$	100,000	1,102	1,556	0,492	99,997	100,003	0,242
	$*(kurtosis(K_{dc}))^{-1}$	100,000	1,100	1,555	0,492	99,997	100,003	0,242
	$meff/Var(y)$	100,000	1,027	1,555	0,492	99,997	100,003	0,242
	$(meff)^{-1}$	100,000	1,075	1,561	0,494	99,997	100,003	0,244
10	1/n	100,001	2,338	1,110	0,351	99,999	100,004	0,124
	$(Var(\hat{\theta}_a))^{-1}$	100,001	2,382	1,112	0,351	99,999	100,004	0,124
	$*(Var(\hat{\theta}_a))^{-1}$	100,001	2,310	1,109	0,351	99,999	100,004	0,124

	$(CV(\hat{\theta}_d))^{-1}$	100,002	2,467	1,111	0,351	99,999	100,004	0,124
	$*(CV(\hat{\theta}_d))^{-1}$	100,002	2,429	1,110	0,351	99,999	100,004	0,124
	$(\text{skewness}(s_{dc}))^{-1}$	100,001	2,101	1,853	0,586	99,998	100,005	0,344
	$*(\text{skewness}(s_{dc}))^{-1}$	100,000	0,552	1,843	0,583	99,996	100,003	0,340
	$(\text{kurtosis}(K_{dc}))^{-1}$	100,001	2,380	1,111	0,351	99,999	100,004	0,124
	$*(\text{kurtosis}(K_{dc}))^{-1}$	100,001	2,374	1,111	0,351	99,999	100,004	0,124
	$meff/Var(y)$	100,001	2,225	1,111	0,351	99,999	100,003	0,124
	$(meff)^{-1}$	100,001	2,372	1,114	0,352	99,999	100,004	0,125
15	$1/n$	100,000	0,870	0,919	0,291	99,998	100,002	0,085
	$(Var(\hat{\theta}_d))^{-1}$	100,000	0,891	0,918	0,290	99,998	100,002	0,084
	$*(Var(\hat{\theta}_d))^{-1}$	100,000	0,873	0,918	0,290	99,998	100,002	0,084
	$(CV(\hat{\theta}_d))^{-1}$	100,000	0,993	0,918	0,290	99,998	100,002	0,084
	$*(CV(\hat{\theta}_d))^{-1}$	100,000	0,982	0,918	0,290	99,998	100,002	0,084
	$(\text{skewness}(s_{dc}))^{-1}$	99,999	0,134	1,764	0,558	99,996	100,003	0,311
	$*(\text{skewness}(s_{dc}))^{-1}$	99,997	-1,733	1,750	0,553	99,994	100,001	0,307
	$(\text{kurtosis}(K_{dc}))^{-1}$	100,000	0,898	0,919	0,291	99,998	100,002	0,085
	$*(\text{kurtosis}(K_{dc}))^{-1}$	100,000	0,889	0,919	0,291	99,998	100,002	0,085
	$meff/Var(y)$	100,000	0,896	0,917	0,290	99,998	100,002	0,084
	$(meff)^{-1}$	100,000	0,950	0,920	0,291	99,998	100,002	0,085
20	$1/n$	99,999	0,347	0,791	0,250	99,998	100,001	0,063
	$(Var(\hat{\theta}_d))^{-1}$	99,999	0,347	0,793	0,251	99,998	100,001	0,063
	$*(Var(\hat{\theta}_d))^{-1}$	99,999	0,352	0,790	0,250	99,998	100,001	0,062
	$(CV(\hat{\theta}_d))^{-1}$	100,000	0,461	0,792	0,250	99,998	100,001	0,063
	$*(CV(\hat{\theta}_d))^{-1}$	100,000	0,462	0,791	0,250	99,998	100,001	0,063
	$(\text{skewness}(s_{dc}))^{-1}$	99,998	-0,625	1,673	0,529	99,995	100,002	0,280
	$*(\text{skewness}(s_{dc}))^{-1}$	99,997	-1,797	1,718	0,543	99,994	100,001	0,296
	$(\text{kurtosis}(K_{dc}))^{-1}$	99,999	0,366	0,791	0,250	99,998	100,001	0,063
	$*(\text{kurtosis}(K_{dc}))^{-1}$	99,999	0,357	0,790	0,250	99,998	100,001	0,063
	$meff/Var(y)$	99,999	0,360	0,790	0,250	99,998	100,001	0,062
	$(meff)^{-1}$	99,999	0,325	0,794	0,251	99,998	100,001	0,063

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 2: N=1000 000 N(100,25) n=5000 D=20 $\theta = 99.999$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		100,001	2,212	1,591	5,032	99,998	100,004	2,537
2	1/n	100,001	1,666	1,094	3,461	99,999	100,003	1,200
	$(Var(\hat{\theta}_a))^{-1}$	100,001	1,848	1,100	3,478	99,999	100,003	1,213
	$*(Var(\hat{\theta}_a))^{-1}$	100,001	1,835	1,100	3,477	99,999	100,003	1,213
	$(CV(\hat{\theta}_a))^{-1}$	100,001	1,860	1,100	3,477	99,999	100,003	1,212
	$*(CV(\hat{\theta}_a))^{-1}$	100,001	1,854	1,100	3,477	99,999	100,003	1,212
	$(skewness(s_{dc}))^{-1}$	100,001	1,798	1,199	3,792	99,999	100,003	1,441
	$*(skewness(s_{dc}))^{-1}$	100,001	1,687	1,209	3,824	99,998	100,003	1,465
	$(kurtosis(K_{dc}))^{-1}$	100,001	1,830	1,099	3,475	99,999	100,003	1,211
	$*(kurtosis(K_{dc}))^{-1}$	100,001	1,832	1,099	3,475	99,999	100,003	1,211
	$meff/Var(y)$	100,001	1,835	1,100	3,477	99,999	100,003	1,213
	$(meff)^{-1}$	100,001	1,861	1,100	3,477	99,999	100,003	1,213
3	1/n	100,000	1,208	0,911	2,882	99,999	100,002	0,832
	$(Var(\hat{\theta}_a))^{-1}$	100,000	1,341	0,916	2,896	99,999	100,002	0,841
	$*(Var(\hat{\theta}_a))^{-1}$	100,000	1,332	0,915	2,893	99,999	100,002	0,839
	$(CV(\hat{\theta}_a))^{-1}$	100,000	1,354	0,916	2,895	99,999	100,002	0,840
	$*(CV(\hat{\theta}_a))^{-1}$	100,000	1,350	0,915	2,894	99,999	100,002	0,839
	$(skewness(s_{dc}))^{-1}$	100,001	1,736	1,098	3,471	99,999	100,003	1,208
	$*(skewness(s_{dc}))^{-1}$	100,000	1,223	1,102	3,486	99,998	100,002	1,216
	$(kurtosis(K_{dc}))^{-1}$	100,000	1,324	0,915	2,893	99,999	100,002	0,839
	$*(kurtosis(K_{dc}))^{-1}$	100,000	1,329	0,915	2,893	99,999	100,002	0,839
	$meff/Var(y)$	100,000	1,341	0,915	2,894	99,999	100,002	0,839
	$(meff)^{-1}$	100,000	1,345	0,916	2,898	99,999	100,002	0,841
5	1/n	100,000	0,815	0,703	2,223	99,999	100,001	0,495
	$(Var(\hat{\theta}_a))^{-1}$	100,000	0,811	0,703	2,223	99,999	100,001	0,495
	$*(Var(\hat{\theta}_a))^{-1}$	100,000	0,813	0,702	2,221	99,999	100,001	0,494
	$(CV(\hat{\theta}_a))^{-1}$	100,000	0,832	0,703	2,223	99,999	100,001	0,495
	$*(CV(\hat{\theta}_a))^{-1}$	100,000	0,832	0,703	2,222	99,999	100,001	0,494
	$(skewness(s_{dc}))^{-1}$	100,001	1,637	0,978	3,092	99,999	100,003	0,959
	$*(skewness(s_{dc}))^{-1}$	100,000	0,592	0,953	3,013	99,998	100,002	0,908
	$(kurtosis(K_{dc}))^{-1}$	100,000	0,797	0,703	2,222	99,998	100,001	0,495
	$*(kurtosis(K_{dc}))^{-1}$	100,000	0,799	0,703	2,223	99,998	100,001	0,495
	$meff/Var(y)$	100,000	0,831	0,703	2,222	99,999	100,001	0,494
	$(meff)^{-1}$	100,000	0,831	0,704	2,225	99,999	100,001	0,496
10	1/n	99,999	-0,195	0,496	1,567	99,998	100,000	0,246
	$(Var(\hat{\theta}_a))^{-1}$	99,999	-0,204	0,496	1,567	99,998	100,000	0,246
	$*(Var(\hat{\theta}_a))^{-1}$	99,999	-0,201	0,496	1,568	99,998	100,000	0,246

	$(CV(\hat{\theta}_d))^{-1}$	99,999	-0,178	0,496	1,567	99,998	100,000	0,246
	$*(CV(\hat{\theta}_d))^{-1}$	99,999	-0,177	0,496	1,567	99,998	100,000	0,246
	$(skewness(s_{dc}))^{-1}$	100,000	0,850	0,867	2,741	99,998	100,002	0,752
	$*(skewness(s_{dc}))^{-1}$	99,998	-1,237	0,812	2,567	99,996	99,999	0,661
	$(kurtosis(K_{dc}))^{-1}$	99,999	-0,193	0,496	1,568	99,998	100,000	0,246
	$*(kurtosis(K_{dc}))^{-1}$	99,999	-0,192	0,496	1,568	99,998	100,000	0,246
	$meff/Var(y)$	99,999	-0,220	0,496	1,569	99,998	100,000	0,246
	$(meff)^{-1}$	99,999	-0,190	0,496	1,567	99,998	100,000	0,246
15	$1/n$	99,999	-0,297	0,400	1,266	99,998	100,000	0,160
	$(Var(\hat{\theta}_d))^{-1}$	99,999	-0,308	0,401	1,267	99,998	100,000	0,161
	$*(Var(\hat{\theta}_d))^{-1}$	99,999	-0,299	0,400	1,266	99,998	100,000	0,160
	$(CV(\hat{\theta}_d))^{-1}$	99,999	-0,280	0,400	1,266	99,998	100,000	0,160
	$*(CV(\hat{\theta}_d))^{-1}$	99,999	-0,276	0,400	1,266	99,998	100,000	0,160
	$(skewness(s_{dc}))^{-1}$	100,000	0,918	0,816	2,580	99,998	100,002	0,667
	$*(skewness(s_{dc}))^{-1}$	99,998	-0,664	0,752	2,379	99,997	100,000	0,566
	$(kurtosis(K_{dc}))^{-1}$	99,999	-0,297	0,400	1,266	99,998	100,000	0,160
	$*(kurtosis(K_{dc}))^{-1}$	99,999	-0,298	0,400	1,266	99,998	100,000	0,160
	$meff/Var(y)$	99,999	-0,293	0,400	1,266	99,998	100,000	0,160
	$(meff)^{-1}$	99,999	-0,310	0,401	1,267	99,998	100,000	0,161
20	$1/n$	99,999	-0,136	0,352	1,113	99,998	100,000	0,124
	$(Var(\hat{\theta}_d))^{-1}$	99,999	-0,149	0,352	1,114	99,998	100,000	0,124
	$*(Var(\hat{\theta}_d))^{-1}$	99,999	-0,137	0,352	1,113	99,998	100,000	0,124
	$(CV(\hat{\theta}_d))^{-1}$	99,999	-0,120	0,352	1,113	99,998	100,000	0,124
	$*(CV(\hat{\theta}_d))^{-1}$	99,999	-0,115	0,352	1,113	99,998	100,000	0,124
	$(skewness(s_{dc}))^{-1}$	100,000	0,837	0,789	2,494	99,998	100,001	0,623
	$*(skewness(s_{dc}))^{-1}$	99,999	-0,133	0,743	2,350	99,997	100,000	0,552
	$(kurtosis(K_{dc}))^{-1}$	99,999	-0,138	0,352	1,113	99,998	100,000	0,124
	$*(kurtosis(K_{dc}))^{-1}$	99,999	-0,138	0,352	1,113	99,998	100,000	0,124
	$meff/Var(y)$	99,999	-0,192	0,353	1,117	99,998	100,000	0,125
	$(meff)^{-1}$	99,999	-0,169	0,352	1,112	99,998	100,000	0,124

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 3: N=1000 000 N(100,25) n=10000 D=20 $\theta = 99.999$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		99,999	-0,251	1,086	3,434	99,997	100,001	1,179
2	1/n	100,000	1,047	0,778	2,459	99,999	100,002	0,606
	$(Var(\hat{\theta}_a))^{-1}$	100,000	1,057	0,782	2,474	99,999	100,002	0,613
	$*(Var(\hat{\theta}_a))^{-1}$	100,000	1,069	0,782	2,473	99,999	100,002	0,613
	$(CV(\hat{\theta}_a))^{-1}$	100,000	1,074	0,782	2,473	99,999	100,002	0,613
	$*(CV(\hat{\theta}_a))^{-1}$	100,000	1,080	0,782	2,472	99,999	100,002	0,612
	$(skewness(s_{dc}))^{-1}$	100,001	1,560	0,881	2,786	99,999	100,002	0,779
	$*(skewness(s_{dc}))^{-1}$	100,001	1,459	0,870	2,751	99,999	100,002	0,759
	$(kurtosis(K_{dc}))^{-1}$	100,000	1,082	0,782	2,472	99,999	100,002	0,612
	$*(kurtosis(K_{dc}))^{-1}$	100,000	1,079	0,782	2,472	99,999	100,002	0,612
	$meff/Var(y)$	100,000	1,069	0,782	2,473	99,999	100,002	0,613
	$(meff)^{-1}$	100,000	1,068	0,782	2,474	99,999	100,002	0,613
3	1/n	99,999	0,044	0,646	2,042	99,998	100,000	0,417
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,204	0,648	2,049	99,998	100,001	0,420
	$*(Var(\hat{\theta}_a))^{-1}$	99,999	0,209	0,647	2,047	99,998	100,001	0,419
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,217	0,648	2,048	99,998	100,001	0,419
	$*(CV(\hat{\theta}_a))^{-1}$	99,999	0,219	0,647	2,047	99,998	100,001	0,419
	$(skewness(s_{dc}))^{-1}$	100,000	0,689	0,778	2,459	99,998	100,001	0,605
	$*(skewness(s_{dc}))^{-1}$	100,000	0,459	0,764	2,417	99,998	100,001	0,585
	$(kurtosis(K_{dc}))^{-1}$	99,999	0,211	0,647	2,047	99,998	100,001	0,419
	$*(kurtosis(K_{dc}))^{-1}$	99,999	0,210	0,647	2,047	99,998	100,001	0,419
	$meff/Var(y)$	99,999	0,209	0,647	2,047	99,998	100,001	0,419
	$(meff)^{-1}$	99,999	0,208	0,648	2,049	99,998	100,001	0,420
5	1/n	100,000	0,637	0,489	1,546	99,999	100,001	0,239
	$(Var(\hat{\theta}_a))^{-1}$	100,000	0,635	0,489	1,548	99,999	100,001	0,240
	$*(Var(\hat{\theta}_a))^{-1}$	100,000	0,638	0,489	1,546	99,999	100,001	0,239
	$(CV(\hat{\theta}_a))^{-1}$	100,000	0,646	0,489	1,547	99,999	100,001	0,240
	$*(CV(\hat{\theta}_a))^{-1}$	100,000	0,647	0,489	1,546	99,999	100,001	0,239
	$(skewness(s_{dc}))^{-1}$	100,000	1,341	0,673	2,129	99,999	100,002	0,455
	$*(skewness(s_{dc}))^{-1}$	100,000	0,661	0,669	2,114	99,998	100,001	0,447
	$(kurtosis(K_{dc}))^{-1}$	100,000	0,642	0,489	1,545	99,999	100,001	0,239
	$*(kurtosis(K_{dc}))^{-1}$	100,000	0,640	0,489	1,545	99,999	100,001	0,239
	$meff/Var(y)$	100,000	0,622	0,489	1,548	99,999	100,001	0,240
	$(meff)^{-1}$	100,000	0,627	0,490	1,549	99,999	100,001	0,240
10	1/n	100,000	0,499	0,348	1,100	99,999	100,000	0,121
	$(Var(\hat{\theta}_a))^{-1}$	100,000	0,496	0,348	1,102	99,999	100,000	0,122
	$*(Var(\hat{\theta}_a))^{-1}$	100,000	0,498	0,348	1,100	99,999	100,000	0,121

	$(CV(\hat{\theta}_d))^{-1}$	100,000	0,508	0,348	1,101	99,999	100,000	0,122
	$*(CV(\hat{\theta}_d))^{-1}$	100,000	0,509	0,348	1,100	99,999	100,000	0,121
	$(\text{skewness}(s_{dc}))^{-1}$	100,000	0,959	0,595	1,882	99,999	100,001	0,355
	$*(\text{skewness}(s_{dc}))^{-1}$	100,000	0,604	0,596	1,884	99,999	100,001	0,355
	$(\text{kurtosis}(K_{dc}))^{-1}$	100,000	0,507	0,348	1,101	99,999	100,000	0,121
	$*(\text{kurtosis}(K_{dc}))^{-1}$	100,000	0,506	0,348	1,101	99,999	100,000	0,121
	$meff/Var(y)$	100,000	0,474	0,350	1,106	99,999	100,000	0,122
	$(meff)^{-1}$	100,000	0,497	0,349	1,102	99,999	100,000	0,122
15	$1/n$	99,999	0,405	0,282	0,893	99,999	100,000	0,080
	$(Var(\hat{\theta}_d))^{-1}$	99,999	0,405	0,283	0,894	99,999	100,000	0,080
	$*(Var(\hat{\theta}_d))^{-1}$	99,999	0,400	0,282	0,893	99,999	100,000	0,080
	$(CV(\hat{\theta}_d))^{-1}$	99,999	0,416	0,283	0,894	99,999	100,000	0,080
	$*(CV(\hat{\theta}_d))^{-1}$	99,999	0,413	0,282	0,893	99,999	100,000	0,080
	$(\text{skewness}(s_{dc}))^{-1}$	100,000	0,579	0,550	1,740	99,999	100,001	0,303
	$*(\text{skewness}(s_{dc}))^{-1}$	100,000	0,839	0,536	1,695	99,999	100,001	0,288
	$(\text{kurtosis}(K_{dc}))^{-1}$	99,999	0,411	0,283	0,894	99,999	100,000	0,080
	$*(\text{kurtosis}(K_{dc}))^{-1}$	99,999	0,412	0,283	0,894	99,999	100,000	0,080
	$meff/Var(y)$	99,999	0,389	0,283	0,894	99,999	100,000	0,080
	$(meff)^{-1}$	99,999	0,417	0,283	0,894	99,999	100,000	0,080
20	$1/n$	100,000	0,512	0,243	0,769	99,999	100,000	0,059
	$(Var(\hat{\theta}_d))^{-1}$	100,000	0,513	0,243	0,769	99,999	100,000	0,059
	$*(Var(\hat{\theta}_d))^{-1}$	100,000	0,510	0,243	0,769	99,999	100,000	0,059
	$(CV(\hat{\theta}_d))^{-1}$	100,000	0,524	0,243	0,769	99,999	100,000	0,059
	$*(CV(\hat{\theta}_d))^{-1}$	100,000	0,522	0,243	0,769	99,999	100,000	0,059
	$(\text{skewness}(s_{dc}))^{-1}$	100,000	0,517	0,525	1,659	99,999	100,001	0,275
	$*(\text{skewness}(s_{dc}))^{-1}$	100,000	1,186	0,511	1,616	99,999	100,001	0,263
	$(\text{kurtosis}(K_{dc}))^{-1}$	100,000	0,518	0,243	0,769	99,999	100,000	0,059
	$*(\text{kurtosis}(K_{dc}))^{-1}$	100,000	0,518	0,243	0,769	99,999	100,000	0,059
	$meff/Var(y)$	100,000	0,464	0,246	0,779	99,999	100,000	0,061
	$(meff)^{-1}$	100,000	0,475	0,245	0,776	99,999	100,000	0,060

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Appendix A3

Comparing weighting strategies: Cluster Random Sampling: Normal distribution

An asterisk before a weighting strategy indicates that the estimator is not design consistent. That is, a naïve estimator that does not account for the sample design was used to weight surveys.

Scenario 1: N=1 000 000 N(100,25) n=1000 D=20, $\theta = 100.031$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	$SE(\hat{\theta}_c)$ $\times 10^{-2}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-1}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-2}$
						LB	UB	
1		99,975	-5,665	2,565	2,566	99,924	100,025	6,902
2	1/n	99,996	-3,477	1,371	1,371	99,970	100,023	2,002
	$(Var(\hat{\theta}_a))^{-1}$	99,985	-4,601	1,672	1,672	99,952	100,018	3,006
	$*(Var(\hat{\theta}_a))^{-1}$	99,986	-4,513	1,655	1,655	99,954	100,019	2,942
	$(CV(\hat{\theta}_a))^{-1}$	99,986	-4,528	1,662	1,662	99,953	100,019	2,966
	$*(CV(\hat{\theta}_a))^{-1}$	99,986	-4,484	1,656	1,657	99,954	100,019	2,945
	$(skewness(s_{dc}))^{-1}$	100,006	-2,533	1,794	1,794	99,971	100,041	3,282
	$*(skewness(s_{dc}))^{-1}$	100,006	-2,533	1,794	1,794	99,971	100,041	3,282
	$(kurtosis(K_{dc}))^{-1}$	99,986	-4,519	1,662	1,662	99,953	100,019	2,967
	$*(kurtosis(K_{dc}))^{-1}$	99,986	-4,519	1,662	1,662	99,953	100,019	2,967
	$meff/Var(y)$	99,986	-4,516	1,655	1,655	99,954	100,019	2,943
	$(meff)^{-1}$	99,985	-4,590	1,673	1,673	99,953	100,018	3,010
3	1/n	99,990	-4,168	1,175	1,175	99,967	100,013	1,553
	$(Var(\hat{\theta}_a))^{-1}$	99,987	-4,397	1,340	1,341	99,961	100,014	1,990
	$*(Var(\hat{\theta}_a))^{-1}$	99,987	-4,417	1,312	1,312	99,961	100,013	1,916
	$(CV(\hat{\theta}_a))^{-1}$	99,987	-4,377	1,321	1,321	99,962	100,013	1,937
	$*(CV(\hat{\theta}_a))^{-1}$	99,987	-4,389	1,311	1,311	99,962	100,013	1,910
	$(skewness(s_{dc}))^{-1}$	100,008	-2,360	1,681	1,681	99,975	100,041	2,882
	$*(skewness(s_{dc}))^{-1}$	100,008	-2,360	1,681	1,681	99,975	100,041	2,882
	$(kurtosis(K_{dc}))^{-1}$	99,986	-4,486	1,318	1,318	99,961	100,012	1,938
	$*(kurtosis(K_{dc}))^{-1}$	99,986	-4,486	1,318	1,318	99,961	100,012	1,938
	$meff/Var(y)$	99,989	-4,238	1,338	1,339	99,963	100,015	1,971
	$(meff)^{-1}$	99,987	-4,449	1,333	1,333	99,961	100,013	1,974
5	1/n	99,995	-3,579	1,103	1,103	99,974	100,017	1,345
	$(Var(\hat{\theta}_a))^{-1}$	99,996	-3,510	1,108	1,108	99,974	100,018	1,352
	$*(Var(\hat{\theta}_a))^{-1}$	99,996	-3,496	1,101	1,101	99,975	100,018	1,334
	$(CV(\hat{\theta}_a))^{-1}$	99,996	-3,501	1,102	1,103	99,975	100,018	1,338
	$*(CV(\hat{\theta}_a))^{-1}$	99,996	-3,498	1,102	1,102	99,975	100,018	1,336
	$(skewness(s_{dc}))^{-1}$	99,999	-3,191	1,527	1,527	99,969	100,029	2,434
	$*(skewness(s_{dc}))^{-1}$	99,999	-3,191	1,527	1,527	99,969	100,029	2,434
	$(kurtosis(K_{dc}))^{-1}$	99,995	-3,610	1,108	1,109	99,973	100,017	1,359
	$*(kurtosis(K_{dc}))^{-1}$	99,995	-3,610	1,108	1,109	99,973	100,017	1,359
	$meff/Var(y)$	99,999	-3,207	1,213	1,213	99,975	100,023	1,575
	$(meff)^{-1}$	99,997	-3,428	1,161	1,161	99,974	100,020	1,466
10	1/n	99,996	-3,501	0,728	0,728	99,982	100,010	0,653
	$(Var(\hat{\theta}_a))^{-1}$	99,997	-3,438	0,727	0,727	99,983	100,011	0,647
	$*(Var(\hat{\theta}_a))^{-1}$	99,997	-3,456	0,724	0,724	99,982	100,011	0,644

	$(CV(\hat{\theta}_d))^{-1}$	99,997	-3,419	0,724	0,724	99,983	100,011	0,641
	$*(CV(\hat{\theta}_d))^{-1}$	99,997	-3,431	0,726	0,726	99,983	100,011	0,645
	$(\text{skewness}(s_{dc}))^{-1}$	100,002	-2,926	1,370	1,370	99,975	100,029	1,961
	$*(\text{skewness}(s_{dc}))^{-1}$	100,002	-2,926	1,370	1,370	99,975	100,029	1,961
	$(\text{kurtosis}(K_{dc}))^{-1}$	99,996	-3,530	0,727	0,727	99,982	100,010	0,653
	$*(\text{kurtosis}(K_{dc}))^{-1}$	99,996	-3,530	0,727	0,727	99,982	100,010	0,653
	$meff/Var(y)$	99,995	-3,644	0,759	0,759	99,980	100,010	0,708
	$(meff)^{-1}$	99,996	-3,564	0,740	0,740	99,981	100,010	0,674
15	$1/n$	99,999	-3,242	0,611	0,611	99,987	100,011	0,478
	$(Var(\hat{\theta}_d))^{-1}$	99,998	-3,286	0,594	0,594	99,987	100,010	0,460
	$*(Var(\hat{\theta}_d))^{-1}$	99,999	-3,229	0,606	0,606	99,987	100,011	0,471
	$(CV(\hat{\theta}_d))^{-1}$	99,999	-3,217	0,600	0,600	99,987	100,011	0,463
	$*(CV(\hat{\theta}_d))^{-1}$	99,999	-3,186	0,608	0,608	99,987	100,011	0,472
	$(\text{skewness}(s_{dc}))^{-1}$	100,001	-3,036	1,355	1,355	99,974	100,027	1,929
	$*(\text{skewness}(s_{dc}))^{-1}$	100,001	-3,036	1,355	1,355	99,974	100,027	1,929
	$(\text{kurtosis}(K_{dc}))^{-1}$	99,999	-3,272	0,613	0,613	99,986	100,011	0,483
	$*(\text{kurtosis}(K_{dc}))^{-1}$	99,999	-3,272	0,613	0,613	99,986	100,011	0,483
	$meff/Var(y)$	99,997	-3,429	0,646	0,646	99,984	100,010	0,535
	$(meff)^{-1}$	99,997	-3,419	0,610	0,610	99,985	100,009	0,489
20	$1/n$	100,000	-3,098	0,537	0,537	99,990	100,011	0,384
	$(Var(\hat{\theta}_d))^{-1}$	100,000	-3,129	0,528	0,528	99,990	100,010	0,377
	$*(Var(\hat{\theta}_d))^{-1}$	100,000	-3,091	0,535	0,535	99,990	100,011	0,382
	$(CV(\hat{\theta}_d))^{-1}$	100,001	-3,061	0,530	0,530	99,990	100,011	0,375
	$*(CV(\hat{\theta}_d))^{-1}$	100,001	-3,042	0,536	0,536	99,990	100,011	0,380
	$(\text{skewness}(s_{dc}))^{-1}$	100,006	-2,536	1,304	1,303	99,980	100,031	1,764
	$*(\text{skewness}(s_{dc}))^{-1}$	100,006	-2,536	1,304	1,303	99,980	100,031	1,764
	$(\text{kurtosis}(K_{dc}))^{-1}$	100,000	-3,131	0,540	0,540	99,989	100,010	0,389
	$*(\text{kurtosis}(K_{dc}))^{-1}$	100,000	-3,131	0,540	0,540	99,989	100,010	0,389
	$meff/Var(y)$	100,001	-3,055	0,539	0,539	99,990	100,011	0,384
	$(meff)^{-1}$	100,000	-3,125	0,529	0,529	99,990	100,010	0,377

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 2: N=1000 000 N(100,25) n=5000 D=20 $\theta = 100.031$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		100,035	4,140	1,404	1,403	100,008	100,063	1,972
2	1/n	100,024	-7,345	0,908	0,908	100,006	100,042	0,830
	$(Var(\hat{\theta}_a))^{-1}$	100,024	-6,854	1,041	1,041	100,004	100,045	1,089
	$*(Var(\hat{\theta}_a))^{-1}$	100,024	-6,786	1,043	1,043	100,004	100,045	1,093
	$(CV(\hat{\theta}_a))^{-1}$	100,024	-6,775	1,043	1,042	100,004	100,045	1,092
	$*(CV(\hat{\theta}_a))^{-1}$	100,024	-6,741	1,044	1,044	100,004	100,045	1,095
	$(skewness(s_{dc}))^{-1}$	100,032	0,862	1,314	1,314	100,006	100,058	1,727
	$*(skewness(s_{dc}))^{-1}$	100,032	0,862	1,314	1,314	100,006	100,058	1,727
	$(kurtosis(K_{dc}))^{-1}$	100,024	-7,210	1,048	1,048	100,003	100,045	1,104
	$*(kurtosis(K_{dc}))^{-1}$	100,024	-7,210	1,048	1,048	100,003	100,045	1,104
	$mef/Var(y)$	100,024	-6,793	1,043	1,043	100,004	100,045	1,093
	$(mef)^{-1}$	100,024	-7,018	1,043	1,043	100,004	100,045	1,093
3	1/n	100,028	-3,697	0,711	0,710	100,014	100,041	0,506
	$(Var(\hat{\theta}_a))^{-1}$	100,029	-2,419	0,816	0,816	100,013	100,045	0,667
	$*(Var(\hat{\theta}_a))^{-1}$	100,029	-2,665	0,823	0,822	100,012	100,045	0,678
	$(CV(\hat{\theta}_a))^{-1}$	100,029	-2,445	0,820	0,820	100,013	100,045	0,673
	$*(CV(\hat{\theta}_a))^{-1}$	100,029	-2,575	0,824	0,824	100,012	100,045	0,679
	$(skewness(s_{dc}))^{-1}$	100,032	0,807	1,158	1,158	100,009	100,055	1,342
	$*(skewness(s_{dc}))^{-1}$	100,032	0,807	1,158	1,158	100,009	100,055	1,342
	$(kurtosis(K_{dc}))^{-1}$	100,028	-3,239	0,826	0,826	100,012	100,044	0,684
	$*(kurtosis(K_{dc}))^{-1}$	100,028	-3,239	0,826	0,826	100,012	100,044	0,684
	$mef/Var(y)$	100,029	-2,539	0,822	0,822	100,013	100,045	0,676
	$(mef)^{-1}$	100,029	-2,520	0,819	0,819	100,013	100,045	0,671
5	1/n	100,031	-0,506	0,614	0,613	100,019	100,043	0,377
	$(Var(\hat{\theta}_a))^{-1}$	100,031	0,247	0,610	0,610	100,020	100,043	0,372
	$*(Var(\hat{\theta}_a))^{-1}$	100,031	-0,281	0,611	0,611	100,019	100,043	0,373
	$(CV(\hat{\theta}_a))^{-1}$	100,031	0,047	0,611	0,611	100,019	100,043	0,373
	$*(CV(\hat{\theta}_a))^{-1}$	100,031	-0,218	0,612	0,612	100,019	100,043	0,375
	$(skewness(s_{dc}))^{-1}$	100,038	6,994	1,076	1,075	100,017	100,059	1,162
	$*(skewness(s_{dc}))^{-1}$	100,038	6,994	1,076	1,075	100,017	100,059	1,162
	$(kurtosis(K_{dc}))^{-1}$	100,030	-0,854	0,613	0,613	100,018	100,042	0,376
	$*(kurtosis(K_{dc}))^{-1}$	100,030	-0,854	0,613	0,613	100,018	100,042	0,376
	$mef/Var(y)$	100,030	-0,980	0,621	0,620	100,018	100,042	0,385
	$(mef)^{-1}$	100,031	0,199	0,611	0,610	100,019	100,043	0,373
10	1/n	100,027	-4,396	0,460	0,460	100,018	100,036	0,214
	$(Var(\hat{\theta}_a))^{-1}$	100,027	-4,376	0,460	0,460	100,018	100,036	0,213
	$*(Var(\hat{\theta}_a))^{-1}$	100,027	-4,368	0,459	0,459	100,018	100,036	0,213

	$(CV(\hat{\theta}_d))^{-1}$	100,027	-4,188	0,459	0,459	100,018	100,036	0,213
	$*(CV(\hat{\theta}_d))^{-1}$	100,027	-4,180	0,460	0,459	100,018	100,036	0,213
	$(\text{skewness}(s_{dc}))^{-1}$	100,037	5,317	0,982	0,982	100,017	100,056	0,968
	$*(\text{skewness}(s_{dc}))^{-1}$	100,037	5,317	0,982	0,982	100,017	100,056	0,968
	$(\text{kurtosis}(K_{dc}))^{-1}$	100,027	-4,529	0,461	0,461	100,018	100,036	0,215
	$*(\text{kurtosis}(K_{dc}))^{-1}$	100,027	-4,529	0,461	0,461	100,018	100,036	0,215
	$meff/Var(y)$	100,027	-4,077	0,462	0,462	100,018	100,036	0,215
	$(meff)^{-1}$	100,027	-4,361	0,460	0,460	100,018	100,036	0,214
15	$1/n$	100,028	-2,724	0,389	0,389	100,021	100,036	0,152
	$(Var(\hat{\theta}_d))^{-1}$	100,028	-2,792	0,387	0,387	100,021	100,036	0,151
	$*(Var(\hat{\theta}_d))^{-1}$	100,028	-2,765	0,389	0,389	100,021	100,036	0,152
	$(CV(\hat{\theta}_d))^{-1}$	100,029	-2,553	0,388	0,388	100,021	100,036	0,151
	$*(CV(\hat{\theta}_d))^{-1}$	100,029	-2,538	0,389	0,389	100,021	100,036	0,152
	$(\text{skewness}(s_{dc}))^{-1}$	100,036	4,938	0,848	0,848	100,020	100,053	0,722
	$*(\text{skewness}(s_{dc}))^{-1}$	100,036	4,938	0,848	0,848	100,020	100,053	0,722
	$(\text{kurtosis}(K_{dc}))^{-1}$	100,028	-2,840	0,391	0,391	100,021	100,036	0,153
	$*(\text{kurtosis}(K_{dc}))^{-1}$	100,028	-2,840	0,391	0,391	100,021	100,036	0,153
	$meff/Var(y)$	100,026	-5,200	0,468	0,468	100,017	100,035	0,222
	$(meff)^{-1}$	100,027	-4,270	0,422	0,422	100,019	100,035	0,180
20	$1/n$	100,026	-4,955	0,327	0,327	100,020	100,033	0,110
	$(Var(\hat{\theta}_d))^{-1}$	100,026	-4,810	0,326	0,325	100,020	100,033	0,108
	$*(Var(\hat{\theta}_d))^{-1}$	100,026	-4,927	0,327	0,327	100,020	100,033	0,109
	$(CV(\hat{\theta}_d))^{-1}$	100,027	-4,663	0,326	0,326	100,020	100,033	0,109
	$*(CV(\hat{\theta}_d))^{-1}$	100,027	-4,720	0,327	0,327	100,020	100,033	0,109
	$(\text{skewness}(s_{dc}))^{-1}$	100,036	4,864	0,753	0,753	100,021	100,051	0,570
	$*(\text{skewness}(s_{dc}))^{-1}$	100,036	4,864	0,753	0,753	100,021	100,051	0,570
	$(\text{kurtosis}(K_{dc}))^{-1}$	100,026	-4,954	0,328	0,328	100,020	100,033	0,110
	$*(\text{kurtosis}(K_{dc}))^{-1}$	100,026	-4,954	0,328	0,328	100,020	100,033	0,110
	$meff/Var(y)$	100,026	-5,582	0,339	0,339	100,019	100,032	0,118
	$(meff)^{-1}$	100,027	-4,374	0,324	0,324	100,020	100,033	0,107

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 3: N=1000 000 N(100,25) n=10000 D=20 $\theta = 100.031$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		100,020	-1,093	9,764	9,762	100,001	100,039	9,653
2	1/n	100,016	-1,476	6,104	6,103	100,004	100,028	3,944
	$(Var(\hat{\theta}_d))^{-1}$	100,028	-0,364	6,167	6,165	100,015	100,040	3,816
	$*(Var(\hat{\theta}_d))^{-1}$	100,027	-0,436	6,180	6,178	100,015	100,039	3,838
	$(CV(\hat{\theta}_d))^{-1}$	100,027	-0,403	6,170	6,168	100,015	100,039	3,823
	$*(CV(\hat{\theta}_d))^{-1}$	100,027	-0,439	6,178	6,177	100,015	100,039	3,837
	$(skewness(s_{dc}))^{-1}$	100,027	-0,414	6,607	6,605	100,014	100,040	4,383
	$*(skewness(s_{dc}))^{-1}$	100,027	-0,414	6,607	6,605	100,014	100,040	4,383
	$(kurtosis(K_{dc}))^{-1}$	100,027	-0,453	6,175	6,173	100,015	100,039	3,834
	$*(kurtosis(K_{dc}))^{-1}$	100,027	-0,453	6,175	6,173	100,015	100,039	3,834
	$meff/Var(y)$	100,027	-0,435	6,180	6,178	100,015	100,039	3,838
	$(meff)^{-1}$	100,027	-0,381	6,162	6,161	100,015	100,039	3,812
3	1/n	100,017	-1,465	5,223	5,223	100,006	100,027	2,943
	$(Var(\hat{\theta}_d))^{-1}$	100,028	-0,311	5,021	5,019	100,018	100,038	2,530
	$*(Var(\hat{\theta}_d))^{-1}$	100,027	-0,380	5,025	5,023	100,018	100,037	2,539
	$(CV(\hat{\theta}_d))^{-1}$	100,028	-0,350	5,020	5,019	100,018	100,038	2,532
	$*(CV(\hat{\theta}_d))^{-1}$	100,027	-0,384	5,025	5,023	100,018	100,037	2,540
	$(skewness(s_{dc}))^{-1}$	100,019	-1,202	6,199	6,197	100,007	100,031	3,987
	$*(skewness(s_{dc}))^{-1}$	100,019	-1,202	6,199	6,197	100,007	100,031	3,987
	$(kurtosis(K_{dc}))^{-1}$	100,027	-0,410	5,015	5,014	100,017	100,037	2,532
	$*(kurtosis(K_{dc}))^{-1}$	100,027	-0,410	5,015	5,014	100,017	100,037	2,532
	$meff/Var(y)$	100,028	-0,344	5,016	5,015	100,018	100,038	2,528
	$(meff)^{-1}$	100,028	-0,337	5,018	5,016	100,018	100,038	2,529
5	1/n	100,030	-0,078	4,067	4,066	100,022	100,038	1,655
	$(Var(\hat{\theta}_d))^{-1}$	100,031	-0,049	4,087	4,086	100,023	100,039	1,671
	$*(Var(\hat{\theta}_d))^{-1}$	100,030	-0,072	4,067	4,066	100,023	100,038	1,654
	$(CV(\hat{\theta}_d))^{-1}$	100,031	-0,055	4,076	4,075	100,023	100,039	1,662
	$*(CV(\hat{\theta}_d))^{-1}$	100,031	-0,067	4,068	4,066	100,023	100,039	1,655
	$(skewness(s_{dc}))^{-1}$	100,029	-0,211	5,433	5,431	100,018	100,040	2,956
	$*(skewness(s_{dc}))^{-1}$	100,029	-0,211	5,433	5,431	100,018	100,040	2,956
	$(kurtosis(K_{dc}))^{-1}$	100,030	-0,086	4,063	4,062	100,022	100,038	1,652
	$*(kurtosis(K_{dc}))^{-1}$	100,030	-0,086	4,063	4,062	100,022	100,038	1,652
	$meff/Var(y)$	100,030	-0,103	4,011	4,010	100,022	100,038	1,610
	$(meff)^{-1}$	100,031	-0,071	4,052	4,050	100,023	100,038	1,642
10	1/n	100,028	-0,340	2,808	2,807	100,022	100,033	0,800
	$(Var(\hat{\theta}_d))^{-1}$	100,028	-0,320	2,822	2,821	100,022	100,034	0,807
	$*(Var(\hat{\theta}_d))^{-1}$	100,028	-0,340	2,808	2,808	100,022	100,033	0,800

	$(CV(\hat{\theta}_d))^{-1}$	100,028	-0,320	2,814	2,813	100,023	100,034	0,802
	$*(CV(\hat{\theta}_d))^{-1}$	100,028	-0,330	2,809	2,808	100,022	100,033	0,800
	$(skewness(s_{dc}))^{-1}$	100,032	0,107	4,409	4,408	100,024	100,041	1,945
	$*(skewness(s_{dc}))^{-1}$	100,032	0,107	4,409	4,408	100,024	100,041	1,945
	$(kurtosis(K_{dc}))^{-1}$	100,028	-0,349	2,810	2,809	100,022	100,033	0,802
	$*(kurtosis(K_{dc}))^{-1}$	100,028	-0,349	2,810	2,809	100,022	100,033	0,802
	$meff/Var(y)$	100,028	-0,358	2,816	2,815	100,022	100,033	0,806
	$(meff)^{-1}$	100,028	-0,290	2,829	2,829	100,023	100,034	0,809
15	$1/n$	100,031	0,026	2,468	2,467	100,027	100,036	0,609
	$(Var(\hat{\theta}_d))^{-1}$	100,032	0,059	2,469	2,468	100,027	100,037	0,610
	$*(Var(\hat{\theta}_d))^{-1}$	100,031	0,027	2,469	2,468	100,027	100,036	0,610
	$(CV(\hat{\theta}_d))^{-1}$	100,032	0,052	2,467	2,466	100,027	100,037	0,609
	$*(CV(\hat{\theta}_d))^{-1}$	100,032	0,036	2,468	2,467	100,027	100,036	0,609
	$(skewness(s_{dc}))^{-1}$	100,036	0,527	4,337	4,336	100,028	100,045	1,909
	$*(skewness(s_{dc}))^{-1}$	100,036	0,527	4,337	4,336	100,028	100,045	1,909
	$(kurtosis(K_{dc}))^{-1}$	100,031	0,020	2,471	2,470	100,027	100,036	0,610
	$*(kurtosis(K_{dc}))^{-1}$	100,031	0,020	2,471	2,470	100,027	100,036	0,610
	$meff/Var(y)$	100,030	-0,136	2,745	2,745	100,024	100,035	0,756
	$(meff)^{-1}$	100,031	0,017	2,437	2,436	100,027	100,036	0,594
20	$1/n$	100,033	0,137	1,982	1,981	100,029	100,036	0,395
	$(Var(\hat{\theta}_d))^{-1}$	100,033	0,167	1,978	1,977	100,029	100,037	0,394
	$*(Var(\hat{\theta}_d))^{-1}$	100,033	0,139	1,982	1,981	100,029	100,036	0,395
	$(CV(\hat{\theta}_d))^{-1}$	100,033	0,162	1,978	1,977	100,029	100,037	0,394
	$*(CV(\hat{\theta}_d))^{-1}$	100,033	0,148	1,981	1,980	100,029	100,037	0,395
	$(skewness(s_{dc}))^{-1}$	100,034	0,285	4,785	4,783	100,025	100,043	2,298
	$*(skewness(s_{dc}))^{-1}$	100,034	0,285	4,785	4,783	100,025	100,043	2,298
	$(kurtosis(K_{dc}))^{-1}$	100,033	0,133	1,984	1,984	100,029	100,036	0,395
	$*(kurtosis(K_{dc}))^{-1}$	100,033	0,133	1,984	1,984	100,029	100,036	0,395
	$meff/Var(y)$	100,032	0,093	2,049	2,048	100,028	100,036	0,421
	$(meff)^{-1}$	100,033	0,139	2,000	2,000	100,029	100,037	0,402

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Appendix B1

Comparing weighting strategies: Stratified Random Sampling: Skewed normal distribution

Scenario 1: N=1 000 000 aN(100,25) n=1000 D=20, $\theta = 100.056$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		100,052	-4,893	3,479	1,099	100,045	100,058	1,213
2	1/n	100,052	-3,949	2,503	0,791	100,048	100,057	0,628
	$(Var(\hat{\theta}_a))^{-1}$	100,051	-5,485	2,521	0,797	100,046	100,056	0,638
	$*(Var(\hat{\theta}_a))^{-1}$	100,052	-4,736	2,517	0,795	100,047	100,057	0,636
	$(CV(\hat{\theta}_a))^{-1}$	100,052	-4,799	2,518	0,796	100,047	100,057	0,637
	$*(CV(\hat{\theta}_a))^{-1}$	100,052	-4,424	2,517	0,796	100,047	100,057	0,635
	$(skewness(s_{dc}))^{-1}$	100,051	-5,335	2,575	0,814	100,046	100,056	0,666
	$*(skewness(s_{dc}))^{-1}$	100,051	-5,339	2,583	0,816	100,046	100,056	0,670
	$(kurtosis(K_{dc}))^{-1}$	100,052	-4,150	2,518	0,796	100,047	100,057	0,636
	$*(kurtosis(K_{dc}))^{-1}$	100,052	-4,249	2,518	0,796	100,047	100,057	0,636
	$meff/Var(y)$	100,052	-4,735	2,517	0,795	100,047	100,057	0,636
	$(meff)^{-1}$	100,051	-4,977	2,522	0,797	100,047	100,056	0,638
3	1/n	100,053	-3,311	2,037	0,644	100,049	100,057	0,416
	$(Var(\hat{\theta}_a))^{-1}$	100,051	-5,567	2,051	0,648	100,047	100,055	0,424
	$*(Var(\hat{\theta}_a))^{-1}$	100,052	-4,609	2,050	0,648	100,048	100,056	0,423
	$(CV(\hat{\theta}_a))^{-1}$	100,052	-4,682	2,051	0,648	100,048	100,056	0,423
	$*(CV(\hat{\theta}_a))^{-1}$	100,052	-4,203	2,051	0,648	100,048	100,056	0,422
	$(skewness(s_{dc}))^{-1}$	100,052	-4,330	2,113	0,668	100,048	100,056	0,448
	$*(skewness(s_{dc}))^{-1}$	100,052	-4,350	2,136	0,675	100,048	100,056	0,458
	$(kurtosis(K_{dc}))^{-1}$	100,053	-3,880	2,051	0,648	100,049	100,057	0,422
	$*(kurtosis(K_{dc}))^{-1}$	100,052	-3,955	2,051	0,648	100,048	100,056	0,422
	$meff/Var(y)$	100,052	-4,616	2,050	0,648	100,048	100,056	0,423
	$(meff)^{-1}$	100,052	-4,918	2,052	0,649	100,047	100,056	0,424
5	1/n	100,054	-1,992	1,541	0,487	100,051	100,057	0,238
	$(Var(\hat{\theta}_a))^{-1}$	100,053	-3,902	1,537	0,486	100,050	100,056	0,238
	$*(Var(\hat{\theta}_a))^{-1}$	100,054	-2,774	1,540	0,487	100,051	100,057	0,238
	$(CV(\hat{\theta}_a))^{-1}$	100,054	-2,851	1,539	0,486	100,051	100,057	0,238
	$*(CV(\hat{\theta}_a))^{-1}$	100,054	-2,289	1,541	0,487	100,051	100,057	0,238
	$(skewness(s_{dc}))^{-1}$	100,054	-2,399	1,630	0,515	100,051	100,057	0,266
	$*(skewness(s_{dc}))^{-1}$	100,054	-2,500	1,656	0,523	100,051	100,057	0,275
	$(kurtosis(K_{dc}))^{-1}$	100,054	-1,944	1,543	0,488	100,051	100,058	0,238
	$*(kurtosis(K_{dc}))^{-1}$	100,054	-2,043	1,543	0,488	100,051	100,057	0,239
	$meff/Var(y)$	100,054	-2,785	1,540	0,487	100,051	100,057	0,238

	$(meff)^{-1}$	100,053	-3,181	1,540	0,487	100,050	100,056	0,238
10	$1/n$	100,055	-1,312	1,079	0,341	100,053	100,057	0,117
	$(Var(\hat{\theta}_d))^{-1}$	100,053	-3,474	1,080	0,341	100,051	100,055	0,118
	$*(Var(\hat{\theta}_d))^{-1}$	100,054	-2,162	1,080	0,341	100,052	100,056	0,117
	$(CV(\hat{\theta}_d))^{-1}$	100,054	-2,285	1,079	0,341	100,052	100,056	0,117
	$*(CV(\hat{\theta}_d))^{-1}$	100,055	-1,631	1,080	0,341	100,053	100,057	0,117
	$(skewness(s_{dc}))^{-1}$	100,055	-1,198	1,186	0,375	100,053	100,058	0,141
	$*(skewness(s_{dc}))^{-1}$	100,055	-1,540	1,214	0,384	100,053	100,057	0,148
	$(kurtosis(K_{dc}))^{-1}$	100,055	-1,307	1,080	0,341	100,053	100,057	0,117
	$*(kurtosis(K_{dc}))^{-1}$	100,055	-1,413	1,080	0,341	100,053	100,057	0,117
	$meff/Var(y)$	100,054	-2,137	1,081	0,342	100,052	100,056	0,117
	$(meff)^{-1}$	100,054	-2,660	1,079	0,341	100,052	100,056	0,117
15	$1/n$	100,055	-1,722	0,887	0,280	100,053	100,056	0,079
	$(Var(\hat{\theta}_d))^{-1}$	100,052	-3,972	0,886	0,280	100,051	100,054	0,080
	$*(Var(\hat{\theta}_d))^{-1}$	100,054	-2,588	0,887	0,280	100,052	100,056	0,079
	$(CV(\hat{\theta}_d))^{-1}$	100,054	-2,735	0,886	0,280	100,052	100,055	0,079
	$*(CV(\hat{\theta}_d))^{-1}$	100,054	-2,045	0,887	0,280	100,053	100,056	0,079
	$(skewness(s_{dc}))^{-1}$	100,055	-1,384	0,998	0,315	100,053	100,057	0,100
	$*(skewness(s_{dc}))^{-1}$	100,055	-1,811	1,011	0,320	100,053	100,057	0,103
	$(kurtosis(K_{dc}))^{-1}$	100,055	-1,781	0,889	0,281	100,053	100,056	0,079
	$*(kurtosis(K_{dc}))^{-1}$	100,055	-1,903	0,888	0,281	100,053	100,056	0,079
	$meff/Var(y)$	100,054	-2,698	0,884	0,280	100,052	100,055	0,079
	$(meff)^{-1}$	100,053	-3,228	0,885	0,280	100,051	100,055	0,079
20	$1/n$	100,055	-1,731	0,766	0,242	100,053	100,056	0,059
	$(Var(\hat{\theta}_d))^{-1}$	100,052	-4,046	0,765	0,242	100,051	100,054	0,060
	$*(Var(\hat{\theta}_d))^{-1}$	100,054	-2,619	0,766	0,242	100,052	100,055	0,059
	$(CV(\hat{\theta}_d))^{-1}$	100,054	-2,774	0,766	0,242	100,052	100,055	0,059
	$*(CV(\hat{\theta}_d))^{-1}$	100,054	-2,063	0,766	0,242	100,053	100,056	0,059
	$(skewness(s_{dc}))^{-1}$	100,056	-0,683	0,921	0,291	100,054	100,058	0,085
	$*(skewness(s_{dc}))^{-1}$	100,055	-1,019	0,922	0,291	100,054	100,057	0,085
	$(kurtosis(K_{dc}))^{-1}$	100,055	-1,751	0,768	0,243	100,053	100,056	0,059
	$*(kurtosis(K_{dc}))^{-1}$	100,055	-1,875	0,768	0,243	100,053	100,056	0,059
	$meff/Var(y)$	100,054	-2,676	0,771	0,244	100,052	100,055	0,060
	$(meff)^{-1}$	100,053	-3,276	0,782	0,247	100,052	100,055	0,062

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 2: N=1000 000 aN(100,25) n=5000 D=20 $\theta = 100.056$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-4}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		100,057	9,543	1,571	4,966	100,054	100,060	2,470
2	1/n	100,056	-2,360	1,088	3,438	100,054	100,058	1,183
	$(Var(\hat{\theta}_a))^{-1}$	100,057	1,791	1,090	3,446	100,054	100,059	1,189
	$*(Var(\hat{\theta}_a))^{-1}$	100,057	3,159	1,090	3,445	100,055	100,059	1,188
	$(CV(\hat{\theta}_a))^{-1}$	100,057	3,083	1,090	3,445	100,055	100,059	1,188
	$*(CV(\hat{\theta}_a))^{-1}$	100,057	3,765	1,090	3,445	100,055	100,059	1,188
	$(skewness(s_{dc}))^{-1}$	100,057	3,000	1,099	3,472	100,055	100,059	1,207
	$*(skewness(s_{dc}))^{-1}$	100,057	2,703	1,099	3,474	100,055	100,059	1,208
	$(kurtosis(K_{dc}))^{-1}$	100,057	4,092	1,091	3,448	100,055	100,059	1,190
	$*(kurtosis(K_{dc}))^{-1}$	100,057	3,998	1,091	3,447	100,055	100,059	1,190
	$meff/Var(y)$	100,057	3,159	1,090	3,445	100,055	100,059	1,188
	$(meff)^{-1}$	100,057	2,762	1,090	3,446	100,055	100,059	1,189
3	1/n	100,057	1,432	0,884	2,793	100,055	100,058	0,781
	$(Var(\hat{\theta}_a))^{-1}$	100,057	3,312	0,885	2,797	100,055	100,058	0,783
	$*(Var(\hat{\theta}_a))^{-1}$	100,057	5,422	0,884	2,795	100,055	100,059	0,782
	$(CV(\hat{\theta}_a))^{-1}$	100,057	5,188	0,885	2,796	100,055	100,059	0,783
	$*(CV(\hat{\theta}_a))^{-1}$	100,057	6,239	0,884	2,795	100,055	100,059	0,782
	$(skewness(s_{dc}))^{-1}$	100,057	7,177	0,894	2,824	100,055	100,059	0,799
	$*(skewness(s_{dc}))^{-1}$	100,057	6,876	0,893	2,824	100,055	100,059	0,799
	$(kurtosis(K_{dc}))^{-1}$	100,057	6,669	0,886	2,799	100,055	100,059	0,785
	$*(kurtosis(K_{dc}))^{-1}$	100,057	6,478	0,886	2,799	100,055	100,059	0,785
	$meff/Var(y)$	100,057	5,592	0,886	2,800	100,055	100,059	0,785
	$(meff)^{-1}$	100,057	4,624	0,885	2,797	100,055	100,059	0,784
5	1/n	100,057	1,621	0,698	2,205	100,055	100,058	0,487
	$(Var(\hat{\theta}_a))^{-1}$	100,056	-2,415	0,698	2,206	100,055	100,058	0,487
	$*(Var(\hat{\theta}_a))^{-1}$	100,056	0,059	0,698	2,205	100,055	100,058	0,487
	$(CV(\hat{\theta}_a))^{-1}$	100,056	-0,194	0,698	2,205	100,055	100,058	0,487
	$*(CV(\hat{\theta}_a))^{-1}$	100,057	1,038	0,698	2,205	100,055	100,058	0,487
	$(skewness(s_{dc}))^{-1}$	100,057	2,751	0,704	2,226	100,055	100,058	0,496
	$*(skewness(s_{dc}))^{-1}$	100,057	2,839	0,704	2,226	100,055	100,058	0,496
	$(kurtosis(K_{dc}))^{-1}$	100,057	1,735	0,698	2,206	100,055	100,058	0,487
	$*(kurtosis(K_{dc}))^{-1}$	100,057	1,547	0,698	2,206	100,055	100,058	0,487
	$meff/Var(y)$	100,056	-0,124	0,698	2,206	100,055	100,058	0,487
	$(meff)^{-1}$	100,056	-0,937	0,698	2,206	100,055	100,058	0,487
10	1/n	100,056	0,242	0,506	1,598	100,055	100,057	0,256
	$(Var(\hat{\theta}_a))^{-1}$	100,056	-4,181	0,506	1,600	100,055	100,057	0,256
	$*(Var(\hat{\theta}_a))^{-1}$	100,056	-1,420	0,506	1,599	100,055	100,057	0,256
	$(CV(\hat{\theta}_a))^{-1}$	100,056	-1,748	0,506	1,599	100,055	100,057	0,256

	$*(CV(\hat{\theta}_a))^{-1}$	100,056	-0,373	0,506	1,598	100,055	100,057	0,256
	$(\text{skewness}(s_{dc}))^{-1}$	100,057	0,745	0,512	1,617	100,055	100,058	0,262
	$*(\text{skewness}(s_{dc}))^{-1}$	100,056	0,530	0,512	1,619	100,055	100,057	0,262
	$(\text{kurtosis}(K_{dc}))^{-1}$	100,056	0,349	0,506	1,599	100,055	100,057	0,256
	$*(\text{kurtosis}(K_{dc}))^{-1}$	100,056	0,110	0,506	1,599	100,055	100,057	0,256
	$meff/Var(y)$	100,056	-1,084	0,508	1,605	100,055	100,057	0,258
	$(meff)^{-1}$	100,056	-2,316	0,507	1,601	100,055	100,057	0,257
15	$1/n$	100,057	0,928	0,414	1,308	100,056	100,057	0,171
	$(Var(\hat{\theta}_a))^{-1}$	100,056	-3,671	0,414	1,309	100,055	100,057	0,172
	$*(Var(\hat{\theta}_a))^{-1}$	100,056	-0,718	0,414	1,308	100,056	100,057	0,171
	$(CV(\hat{\theta}_a))^{-1}$	100,056	-1,145	0,414	1,308	100,056	100,057	0,171
	$*(CV(\hat{\theta}_a))^{-1}$	100,056	0,325	0,414	1,308	100,056	100,057	0,171
	$(\text{skewness}(s_{dc}))^{-1}$	100,057	1,907	0,416	1,316	100,056	100,057	0,173
	$*(\text{skewness}(s_{dc}))^{-1}$	100,057	1,691	0,417	1,317	100,056	100,057	0,174
	$(\text{kurtosis}(K_{dc}))^{-1}$	100,057	1,046	0,413	1,306	100,056	100,057	0,171
	$*(\text{kurtosis}(K_{dc}))^{-1}$	100,057	0,806	0,413	1,306	100,056	100,057	0,171
	$meff/Var(y)$	100,056	-0,892	0,415	1,311	100,056	100,057	0,172
	$(meff)^{-1}$	100,056	-2,416	0,417	1,318	100,055	100,057	0,174
20	$1/n$	100,057	3,099	0,354	1,118	100,056	100,057	0,125
	$(Var(\hat{\theta}_a))^{-1}$	100,056	-1,580	0,354	1,119	100,056	100,057	0,125
	$*(Var(\hat{\theta}_a))^{-1}$	100,057	1,358	0,354	1,118	100,056	100,057	0,125
	$(CV(\hat{\theta}_a))^{-1}$	100,057	0,989	0,354	1,119	100,056	100,057	0,125
	$*(CV(\hat{\theta}_a))^{-1}$	100,057	2,453	0,354	1,118	100,056	100,057	0,125
	$(\text{skewness}(s_{dc}))^{-1}$	100,057	3,510	0,355	1,122	100,056	100,057	0,126
	$*(\text{skewness}(s_{dc}))^{-1}$	100,057	3,277	0,355	1,123	100,056	100,057	0,126
	$(\text{kurtosis}(K_{dc}))^{-1}$	100,057	3,165	0,354	1,118	100,056	100,057	0,125
	$*(\text{kurtosis}(K_{dc}))^{-1}$	100,057	2,922	0,354	1,118	100,056	100,057	0,125
	$meff/Var(y)$	100,057	1,319	0,354	1,118	100,056	100,057	0,125
	$(meff)^{-1}$	100,056	0,435	0,355	1,121	100,056	100,057	0,126

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 3: N=1000 000 aN(100,25) n=10000 D=20 $\theta = 100.056$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		100,058	1,602	1,086	3,432	100,056	100,060	1,182
2	1/n	100,056	-0,383	0,777	2,457	100,055	100,058	0,604
	$(Var(\hat{\theta}_a))^{-1}$	100,056	-0,406	0,783	2,474	100,054	100,058	0,613
	$*(Var(\hat{\theta}_a))^{-1}$	100,056	-0,331	0,782	2,472	100,055	100,058	0,612
	$(CV(\hat{\theta}_a))^{-1}$	100,056	-0,338	0,783	2,473	100,055	100,058	0,612
	$*(CV(\hat{\theta}_a))^{-1}$	100,056	-0,301	0,782	2,472	100,055	100,058	0,612
	$(skewness(s_{dc}))^{-1}$	100,056	-0,175	0,783	2,474	100,055	100,058	0,613
	$*(skewness(s_{dc}))^{-1}$	100,056	-0,188	0,783	2,474	100,055	100,058	0,613
	$(kurtosis(K_{dc}))^{-1}$	100,056	-0,272	0,782	2,472	100,055	100,058	0,612
	$*(kurtosis(K_{dc}))^{-1}$	100,056	-0,282	0,782	2,472	100,055	100,058	0,612
	$meff/Var(y)$	100,056	-0,331	0,782	2,472	100,055	100,058	0,612
	$(meff)^{-1}$	100,056	-0,358	0,783	2,474	100,055	100,058	0,613
3	1/n	100,056	-0,383	0,635	2,006	100,055	100,057	0,403
	$(Var(\hat{\theta}_a))^{-1}$	100,056	-0,435	0,638	2,017	100,055	100,057	0,407
	$*(Var(\hat{\theta}_a))^{-1}$	100,056	-0,341	0,638	2,016	100,055	100,057	0,407
	$(CV(\hat{\theta}_a))^{-1}$	100,056	-0,349	0,638	2,017	100,055	100,057	0,407
	$*(CV(\hat{\theta}_a))^{-1}$	100,056	-0,302	0,638	2,017	100,055	100,057	0,407
	$(skewness(s_{dc}))^{-1}$	100,056	-0,214	0,640	2,022	100,055	100,057	0,409
	$*(skewness(s_{dc}))^{-1}$	100,056	-0,235	0,640	2,023	100,055	100,057	0,410
	$(kurtosis(K_{dc}))^{-1}$	100,056	-0,272	0,638	2,017	100,055	100,057	0,407
	$*(kurtosis(K_{dc}))^{-1}$	100,056	-0,284	0,638	2,017	100,055	100,057	0,407
	$meff/Var(y)$	100,056	-0,332	0,638	2,017	100,055	100,057	0,407
	$(meff)^{-1}$	100,056	-0,373	0,638	2,018	100,055	100,057	0,408
5	1/n	100,056	-0,452	0,475	1,503	100,055	100,057	0,226
	$(Var(\hat{\theta}_a))^{-1}$	100,056	-0,639	0,475	1,503	100,055	100,057	0,226
	$*(Var(\hat{\theta}_a))^{-1}$	100,056	-0,528	0,475	1,502	100,055	100,057	0,226
	$(CV(\hat{\theta}_a))^{-1}$	100,056	-0,536	0,475	1,503	100,055	100,057	0,226
	$*(CV(\hat{\theta}_a))^{-1}$	100,056	-0,481	0,475	1,502	100,055	100,057	0,226
	$(skewness(s_{dc}))^{-1}$	100,056	-0,396	0,476	1,505	100,055	100,057	0,227
	$*(skewness(s_{dc}))^{-1}$	100,056	-0,427	0,476	1,506	100,055	100,057	0,227
	$(kurtosis(K_{dc}))^{-1}$	100,056	-0,447	0,475	1,503	100,055	100,057	0,226
	$*(kurtosis(K_{dc}))^{-1}$	100,056	-0,461	0,475	1,503	100,055	100,057	0,226
	$meff/Var(y)$	100,056	-0,537	0,475	1,501	100,055	100,057	0,226
	$(meff)^{-1}$	100,056	-0,553	0,476	1,505	100,055	100,057	0,227
10	1/n	100,057	0,301	0,342	1,081	100,056	100,057	0,117
	$(Var(\hat{\theta}_a))^{-1}$	100,057	0,086	0,342	1,081	100,056	100,057	0,117
	$*(Var(\hat{\theta}_a))^{-1}$	100,057	0,220	0,342	1,081	100,056	100,057	0,117
	$(CV(\hat{\theta}_a))^{-1}$	100,057	0,204	0,342	1,081	100,056	100,057	0,117

	$*(CV(\hat{\theta}_a))^{-1}$	100,057	0,271	0,342	1,081	100,056	100,057	0,117
	$(skewness(s_{dc}))^{-1}$	100,057	0,310	0,342	1,081	100,056	100,057	0,117
	$*(skewness(s_{dc}))^{-1}$	100,057	0,301	0,342	1,082	100,056	100,057	0,117
	$(kurtosis(K_{dc}))^{-1}$	100,057	0,300	0,342	1,081	100,056	100,057	0,117
	$*(kurtosis(K_{dc}))^{-1}$	100,057	0,287	0,342	1,081	100,056	100,057	0,117
	$mef/Var(y)$	100,057	0,209	0,342	1,081	100,056	100,057	0,117
	$(mef)^{-1}$	100,057	0,162	0,342	1,081	100,056	100,057	0,117
15	$1/n$	100,057	0,201	0,280	0,885	100,056	100,057	0,078
	$(Var(\hat{\theta}_a))^{-1}$	100,056	-0,024	0,280	0,884	100,056	100,057	0,078
	$*(Var(\hat{\theta}_a))^{-1}$	100,057	0,115	0,280	0,885	100,056	100,057	0,078
	$(CV(\hat{\theta}_a))^{-1}$	100,057	0,100	0,280	0,885	100,056	100,057	0,078
	$*(CV(\hat{\theta}_a))^{-1}$	100,057	0,169	0,280	0,885	100,056	100,057	0,078
	$(skewness(s_{dc}))^{-1}$	100,057	0,211	0,281	0,887	100,056	100,057	0,079
	$*(skewness(s_{dc}))^{-1}$	100,057	0,199	0,281	0,887	100,056	100,057	0,079
	$(kurtosis(K_{dc}))^{-1}$	100,057	0,204	0,280	0,885	100,056	100,057	0,078
	$*(kurtosis(K_{dc}))^{-1}$	100,057	0,191	0,280	0,885	100,056	100,057	0,078
	$mef/Var(y)$	100,057	0,094	0,281	0,888	100,056	100,057	0,079
	$(mef)^{-1}$	100,056	0,043	0,281	0,888	100,056	100,057	0,079
20	$1/n$	100,057	0,329	0,244	0,771	100,056	100,057	0,060
	$(Var(\hat{\theta}_a))^{-1}$	100,057	0,098	0,244	0,771	100,056	100,057	0,059
	$*(Var(\hat{\theta}_a))^{-1}$	100,057	0,243	0,244	0,771	100,056	100,057	0,060
	$(CV(\hat{\theta}_a))^{-1}$	100,057	0,225	0,244	0,771	100,056	100,057	0,060
	$*(CV(\hat{\theta}_a))^{-1}$	100,057	0,297	0,244	0,771	100,056	100,057	0,060
	$(skewness(s_{dc}))^{-1}$	100,057	0,329	0,245	0,774	100,056	100,057	0,060
	$*(skewness(s_{dc}))^{-1}$	100,057	0,317	0,245	0,774	100,056	100,057	0,060
	$(kurtosis(K_{dc}))^{-1}$	100,057	0,330	0,244	0,772	100,056	100,057	0,060
	$*(kurtosis(K_{dc}))^{-1}$	100,057	0,316	0,244	0,771	100,056	100,057	0,060
	$mef/Var(y)$	100,057	0,257	0,245	0,775	100,056	100,057	0,060
	$(mef)^{-1}$	100,057	0,176	0,244	0,770	100,056	100,057	0,059

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Appendix B2

Comparing weighting strategies: Cluster Random Sampling: Skewed normal distribution

Scenario 1: N=1 000 000 aN(100,25) n=1000 D=20, $\theta = 99.914$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	$SE(\hat{\theta}_c)$ $\times 10^{-2}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-1}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-1}$
						LB	UB	
1		99,897	-1,688	3,184	3,187	99,834	99,959	1,017
2	1/n	99,905	-0,834	2,179	2,181	99,863	99,948	0,475
	$(Var(\hat{\theta}_d))^{-1}$	99,922	0,846	2,443	2,445	99,874	99,970	0,597
	$*(Var(\hat{\theta}_d))^{-1}$	99,928	1,476	2,407	2,409	99,881	99,976	0,582
	$(CV(\hat{\theta}_d))^{-1}$	99,928	1,410	2,421	2,422	99,880	99,975	0,588
	$*(CV(\hat{\theta}_d))^{-1}$	99,931	1,730	2,409	2,411	99,884	99,978	0,583
	$(skewness(s_{dc}))^{-1}$	99,935	2,113	2,436	2,437	99,887	99,982	0,598
	$*(skewness(s_{dc}))^{-1}$	99,935	2,113	2,436	2,437	99,887	99,982	0,598
	$(kurtosis(K_{dc}))^{-1}$	99,934	2,025	2,417	2,419	99,886	99,981	0,588
	$*(kurtosis(K_{dc}))^{-1}$	99,934	2,025	2,417	2,419	99,886	99,981	0,588
	$meff/Var(y)$	99,928	1,471	2,406	2,407	99,881	99,975	0,581
	$(meff)^{-1}$	99,926	1,237	2,438	2,440	99,878	99,974	0,596
3	1/n	99,898	-1,526	1,959	1,961	99,860	99,937	0,386
	$(Var(\hat{\theta}_d))^{-1}$	99,919	0,541	1,964	1,965	99,881	99,958	0,386
	$*(Var(\hat{\theta}_d))^{-1}$	99,929	1,553	1,957	1,958	99,891	99,967	0,385
	$(CV(\hat{\theta}_d))^{-1}$	99,928	1,435	1,964	1,965	99,889	99,966	0,388
	$*(CV(\hat{\theta}_d))^{-1}$	99,933	1,961	1,967	1,968	99,895	99,972	0,391
	$(skewness(s_{dc}))^{-1}$	99,937	2,361	1,997	1,998	99,898	99,976	0,404
	$*(skewness(s_{dc}))^{-1}$	99,937	2,361	1,997	1,998	99,898	99,976	0,404
	$(kurtosis(K_{dc}))^{-1}$	99,937	2,305	1,989	1,990	99,898	99,976	0,401
	$*(kurtosis(K_{dc}))^{-1}$	99,937	2,305	1,989	1,990	99,898	99,976	0,401
	$meff/Var(y)$	99,929	1,537	1,955	1,957	99,891	99,967	0,385
	$(meff)^{-1}$	99,925	1,156	1,974	1,975	99,886	99,964	0,391
5	1/n	99,925	1,138	1,418	1,419	99,897	99,953	0,202
	$(Var(\hat{\theta}_d))^{-1}$	99,906	-0,766	1,404	1,405	99,878	99,933	0,198
	$*(Var(\hat{\theta}_d))^{-1}$	99,918	0,424	1,405	1,407	99,890	99,945	0,198
	$(CV(\hat{\theta}_d))^{-1}$	99,916	0,258	1,404	1,405	99,889	99,944	0,197
	$*(CV(\hat{\theta}_d))^{-1}$	99,922	0,874	1,412	1,413	99,895	99,950	0,200
	$(skewness(s_{dc}))^{-1}$	99,927	1,342	1,393	1,394	99,900	99,954	0,196
	$*(skewness(s_{dc}))^{-1}$	99,927	1,342	1,393	1,394	99,900	99,954	0,196
	$(kurtosis(K_{dc}))^{-1}$	99,926	1,219	1,413	1,414	99,898	99,953	0,201
	$*(kurtosis(K_{dc}))^{-1}$	99,926	1,219	1,413	1,414	99,898	99,953	0,201
	$meff/Var(y)$	99,918	0,431	1,407	1,408	99,890	99,945	0,198
	$(meff)^{-1}$	99,914	0,054	1,441	1,442	99,886	99,942	0,208
10	1/n	99,941	2,698	1,031	1,032	99,920	99,961	0,114

	$(Var(\hat{\theta}_a))^{-1}$	99,916	0,193	1,073	1,074	99,895	99,937	0,115
	$*(Var(\hat{\theta}_a))^{-1}$	99,932	1,809	1,040	1,041	99,911	99,952	0,111
	$(CV(\hat{\theta}_a))^{-1}$	99,929	1,525	1,046	1,047	99,908	99,949	0,112
	$*(CV(\hat{\theta}_a))^{-1}$	99,937	2,360	1,035	1,036	99,917	99,957	0,113
	$(skewness(s_{dc}))^{-1}$	99,943	2,943	1,034	1,035	99,923	99,963	0,116
	$*(skewness(s_{dc}))^{-1}$	99,943	2,943	1,034	1,035	99,923	99,963	0,116
	$(kurtosis(K_{dc}))^{-1}$	99,942	2,822	1,033	1,033	99,922	99,962	0,115
	$*(kurtosis(K_{dc}))^{-1}$	99,942	2,822	1,033	1,033	99,922	99,962	0,115
	$meff/Var(y)$	99,935	2,180	1,109	1,109	99,914	99,957	0,128
	$(meff)^{-1}$	99,925	1,189	1,068	1,069	99,905	99,946	0,115
15	$1/n$	99,932	1,806	0,861	0,861	99,915	99,949	0,077
	$(Var(\hat{\theta}_a))^{-1}$	99,906	-0,715	0,883	0,883	99,889	99,924	0,078
	$*(Var(\hat{\theta}_a))^{-1}$	99,923	0,947	0,869	0,869	99,906	99,940	0,076
	$(CV(\hat{\theta}_a))^{-1}$	99,920	0,635	0,866	0,867	99,903	99,937	0,075
	$*(CV(\hat{\theta}_a))^{-1}$	99,929	1,490	0,864	0,865	99,912	99,945	0,077
	$(skewness(s_{dc}))^{-1}$	99,937	2,333	0,857	0,857	99,920	99,954	0,079
	$*(skewness(s_{dc}))^{-1}$	99,937	2,333	0,857	0,857	99,920	99,954	0,079
	$(kurtosis(K_{dc}))^{-1}$	99,934	2,018	0,854	0,854	99,917	99,951	0,077
	$*(kurtosis(K_{dc}))^{-1}$	99,934	2,018	0,854	0,854	99,917	99,951	0,077
	$meff/Var(y)$	99,923	0,964	0,871	0,872	99,906	99,940	0,077
	$(meff)^{-1}$	99,915	0,132	0,873	0,874	99,898	99,932	0,076
20	$1/n$	99,931	1,702	0,720	0,720	99,917	99,945	0,055
	$(Var(\hat{\theta}_a))^{-1}$	99,906	-0,796	0,739	0,740	99,891	99,920	0,055
	$*(Var(\hat{\theta}_a))^{-1}$	99,922	0,855	0,723	0,723	99,908	99,936	0,053
	$(CV(\hat{\theta}_a))^{-1}$	99,919	0,541	0,724	0,724	99,905	99,933	0,053
	$*(CV(\hat{\theta}_a))^{-1}$	99,928	1,390	0,721	0,721	99,913	99,942	0,054
	$(skewness(s_{dc}))^{-1}$	99,937	2,349	0,723	0,724	99,923	99,951	0,058
	$*(skewness(s_{dc}))^{-1}$	99,937	2,349	0,723	0,724	99,923	99,951	0,058
	$(kurtosis(K_{dc}))^{-1}$	99,933	1,963	0,712	0,713	99,919	99,947	0,055
	$*(kurtosis(K_{dc}))^{-1}$	99,933	1,963	0,712	0,713	99,919	99,947	0,055
	$meff/Var(y)$	99,924	1,014	0,729	0,729	99,909	99,938	0,054
	$(meff)^{-1}$	99,917	0,297	0,766	0,766	99,902	99,932	0,059

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 2: N=1000 000 aN(100,25) n=5000 D=20 $\theta = 99.914$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,909	-0,491	1,493	1,495	99,879	99,938	2,232
2	1/n	99,889	-2,481	0,946	0,947	99,870	99,907	0,957
	$(Var(\hat{\theta}_a))^{-1}$	99,894	-1,946	1,024	1,025	99,874	99,914	1,086
	$*(Var(\hat{\theta}_a))^{-1}$	99,895	-1,817	1,020	1,021	99,875	99,915	1,074
	$(CV(\hat{\theta}_a))^{-1}$	99,895	-1,825	1,021	1,022	99,875	99,915	1,075
	$*(CV(\hat{\theta}_a))^{-1}$	99,896	-1,761	1,019	1,020	99,876	99,916	1,070
	$(skewness(s_{dc}))^{-1}$	99,898	-1,610	1,018	1,019	99,878	99,917	1,063
	$*(skewness(s_{dc}))^{-1}$	99,898	-1,610	1,018	1,019	99,878	99,917	1,063
	$(kurtosis(K_{dc}))^{-1}$	99,897	-1,693	1,018	1,019	99,877	99,917	1,065
	$*(kurtosis(K_{dc}))^{-1}$	99,897	-1,693	1,018	1,019	99,877	99,917	1,065
	$mef/Var(y)$	99,895	-1,818	1,020	1,021	99,875	99,915	1,074
	$(mef)^{-1}$	99,895	-1,858	1,022	1,023	99,875	99,915	1,078
3	1/n	99,891	-2,287	0,770	0,771	99,876	99,906	0,646
	$(Var(\hat{\theta}_a))^{-1}$	99,898	-1,599	0,860	0,861	99,881	99,914	0,765
	$*(Var(\hat{\theta}_a))^{-1}$	99,899	-1,511	0,858	0,859	99,882	99,915	0,759
	$(CV(\hat{\theta}_a))^{-1}$	99,899	-1,498	0,859	0,860	99,882	99,915	0,760
	$*(CV(\hat{\theta}_a))^{-1}$	99,899	-1,454	0,858	0,859	99,882	99,916	0,758
	$(skewness(s_{dc}))^{-1}$	99,900	-1,334	0,857	0,858	99,883	99,917	0,752
	$*(skewness(s_{dc}))^{-1}$	99,900	-1,334	0,857	0,858	99,883	99,917	0,752
	$(kurtosis(K_{dc}))^{-1}$	99,900	-1,403	0,857	0,858	99,883	99,916	0,754
	$*(kurtosis(K_{dc}))^{-1}$	99,900	-1,403	0,857	0,858	99,883	99,916	0,754
	$mef/Var(y)$	99,899	-1,482	0,857	0,858	99,882	99,916	0,756
	$(mef)^{-1}$	99,898	-1,513	0,860	0,861	99,882	99,915	0,762
5	1/n	99,907	-0,649	0,612	0,613	99,895	99,919	0,379
	$(Var(\hat{\theta}_a))^{-1}$	99,904	-0,985	0,619	0,620	99,892	99,916	0,393
	$*(Var(\hat{\theta}_a))^{-1}$	99,906	-0,771	0,612	0,613	99,894	99,918	0,381
	$(CV(\hat{\theta}_a))^{-1}$	99,906	-0,800	0,615	0,616	99,894	99,918	0,385
	$*(CV(\hat{\theta}_a))^{-1}$	99,907	-0,693	0,612	0,613	99,895	99,919	0,380
	$(skewness(s_{dc}))^{-1}$	99,907	-0,612	0,612	0,612	99,896	99,919	0,378
	$*(skewness(s_{dc}))^{-1}$	99,907	-0,612	0,612	0,612	99,896	99,919	0,378
	$(kurtosis(K_{dc}))^{-1}$	99,907	-0,639	0,611	0,611	99,895	99,919	0,377
	$*(kurtosis(K_{dc}))^{-1}$	99,907	-0,639	0,611	0,611	99,895	99,919	0,377
	$mef/Var(y)$	99,905	-0,856	0,623	0,623	99,893	99,917	0,395
	$(mef)^{-1}$	99,905	-0,894	0,621	0,621	99,893	99,917	0,393
10	1/n	99,907	-0,663	0,420	0,420	99,899	99,915	0,180
	$(Var(\hat{\theta}_a))^{-1}$	99,903	-1,069	0,420	0,420	99,895	99,911	0,187
	$*(Var(\hat{\theta}_a))^{-1}$	99,906	-0,797	0,419	0,419	99,897	99,914	0,182
	$(CV(\hat{\theta}_a))^{-1}$	99,905	-0,846	0,419	0,420	99,897	99,913	0,183

	$*(CV(\hat{\theta}_a))^{-1}$	99,907	-0,711	0,419	0,420	99,898	99,915	0,181
	$(skewness(s_{dc}))^{-1}$	99,908	-0,587	0,419	0,419	99,900	99,916	0,179
	$*(skewness(s_{dc}))^{-1}$	99,908	-0,587	0,419	0,419	99,900	99,916	0,179
	$(kurtosis(K_{dc}))^{-1}$	99,907	-0,648	0,418	0,419	99,899	99,915	0,179
	$*(kurtosis(K_{dc}))^{-1}$	99,907	-0,648	0,418	0,419	99,899	99,915	0,179
	$meff/Var(y)$	99,905	-0,878	0,426	0,427	99,896	99,913	0,189
	$(meff)^{-1}$	99,903	-1,039	0,432	0,432	99,895	99,912	0,197
15	$1/n$	99,908	-0,543	0,355	0,355	99,901	99,915	0,129
	$(Var(\hat{\theta}_a))^{-1}$	99,904	-1,000	0,354	0,355	99,897	99,911	0,135
	$*(Var(\hat{\theta}_a))^{-1}$	99,907	-0,691	0,354	0,354	99,900	99,914	0,130
	$(CV(\hat{\theta}_a))^{-1}$	99,906	-0,751	0,354	0,355	99,899	99,913	0,131
	$*(CV(\hat{\theta}_a))^{-1}$	99,908	-0,597	0,355	0,355	99,901	99,915	0,129
	$(skewness(s_{dc}))^{-1}$	99,909	-0,455	0,353	0,354	99,902	99,916	0,127
	$*(skewness(s_{dc}))^{-1}$	99,909	-0,455	0,353	0,354	99,902	99,916	0,127
	$(kurtosis(K_{dc}))^{-1}$	99,908	-0,517	0,353	0,354	99,902	99,915	0,127
	$*(kurtosis(K_{dc}))^{-1}$	99,908	-0,517	0,353	0,354	99,902	99,915	0,127
	$meff/Var(y)$	99,908	-0,598	0,372	0,372	99,900	99,915	0,142
	$(meff)^{-1}$	99,906	-0,803	0,361	0,361	99,899	99,913	0,137
20	$1/n$	99,909	-0,479	0,300	0,300	99,903	99,915	0,092
	$(Var(\hat{\theta}_a))^{-1}$	99,904	-0,945	0,300	0,300	99,898	99,910	0,099
	$*(Var(\hat{\theta}_a))^{-1}$	99,907	-0,629	0,299	0,300	99,901	99,913	0,094
	$(CV(\hat{\theta}_a))^{-1}$	99,907	-0,691	0,300	0,300	99,901	99,913	0,094
	$*(CV(\hat{\theta}_a))^{-1}$	99,908	-0,533	0,300	0,300	99,902	99,914	0,093
	$(skewness(s_{dc}))^{-1}$	99,910	-0,360	0,297	0,298	99,904	99,916	0,090
	$*(skewness(s_{dc}))^{-1}$	99,910	-0,360	0,297	0,298	99,904	99,916	0,090
	$(kurtosis(K_{dc}))^{-1}$	99,909	-0,437	0,299	0,299	99,903	99,915	0,091
	$*(kurtosis(K_{dc}))^{-1}$	99,909	-0,437	0,299	0,299	99,903	99,915	0,091
	$meff/Var(y)$	99,908	-0,592	0,302	0,302	99,902	99,914	0,095
	$(meff)^{-1}$	99,906	-0,762	0,302	0,302	99,900	99,912	0,097

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 3: N=1000 000 aN(100,25) n=10000 D=20 $\theta = 99.914$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,912	-1,717	1,011	1,012	99,892	99,932	1,022
2	1/n	99,910	-3,600	0,668	0,668	99,897	99,923	0,447
	$(Var(\hat{\theta}_a))^{-1}$	99,906	-8,014	0,722	0,723	99,891	99,920	0,528
	$*(Var(\hat{\theta}_a))^{-1}$	99,907	-6,928	0,722	0,722	99,893	99,921	0,526
	$(CV(\hat{\theta}_a))^{-1}$	99,906	-7,243	0,722	0,722	99,892	99,921	0,526
	$*(CV(\hat{\theta}_a))^{-1}$	99,907	-6,699	0,721	0,722	99,893	99,921	0,525
	$(skewness(s_{dc}))^{-1}$	99,908	-5,840	0,727	0,728	99,894	99,922	0,532
	$*(skewness(s_{dc}))^{-1}$	99,908	-5,840	0,727	0,728	99,894	99,922	0,532
	$(kurtosis(K_{dc}))^{-1}$	99,907	-6,441	0,723	0,723	99,893	99,921	0,526
	$*(kurtosis(K_{dc}))^{-1}$	99,907	-6,441	0,723	0,723	99,893	99,921	0,526
	$meff/Var(y)$	99,907	-6,931	0,722	0,722	99,893	99,921	0,526
	$(meff)^{-1}$	99,906	-7,661	0,722	0,722	99,892	99,920	0,526
3	1/n	99,914	0,257	0,552	0,552	99,903	99,925	0,305
	$(Var(\hat{\theta}_a))^{-1}$	99,909	-4,652	0,584	0,584	99,898	99,920	0,343
	$*(Var(\hat{\theta}_a))^{-1}$	99,911	-3,067	0,582	0,583	99,899	99,922	0,340
	$(CV(\hat{\theta}_a))^{-1}$	99,910	-3,422	0,583	0,583	99,899	99,922	0,341
	$*(CV(\hat{\theta}_a))^{-1}$	99,911	-2,625	0,582	0,583	99,900	99,922	0,340
	$(skewness(s_{dc}))^{-1}$	99,912	-1,579	0,584	0,584	99,901	99,923	0,341
	$*(skewness(s_{dc}))^{-1}$	99,912	-1,579	0,584	0,584	99,901	99,923	0,341
	$(kurtosis(K_{dc}))^{-1}$	99,912	-2,115	0,583	0,583	99,900	99,923	0,340
	$*(kurtosis(K_{dc}))^{-1}$	99,912	-2,115	0,583	0,583	99,900	99,923	0,340
	$meff/Var(y)$	99,910	-3,194	0,582	0,583	99,899	99,922	0,340
	$(meff)^{-1}$	99,910	-3,931	0,584	0,584	99,898	99,921	0,342
5	1/n	99,914	0,630	0,446	0,446	99,906	99,923	0,199
	$(Var(\hat{\theta}_a))^{-1}$	99,912	-1,857	0,444	0,444	99,903	99,920	0,197
	$*(Var(\hat{\theta}_a))^{-1}$	99,913	-0,166	0,445	0,445	99,905	99,922	0,198
	$(CV(\hat{\theta}_a))^{-1}$	99,913	-0,521	0,445	0,445	99,904	99,922	0,198
	$*(CV(\hat{\theta}_a))^{-1}$	99,914	0,324	0,445	0,446	99,905	99,923	0,198
	$(skewness(s_{dc}))^{-1}$	99,915	1,352	0,451	0,451	99,906	99,924	0,203
	$*(skewness(s_{dc}))^{-1}$	99,915	1,352	0,451	0,451	99,906	99,924	0,203
	$(kurtosis(K_{dc}))^{-1}$	99,914	0,879	0,450	0,450	99,906	99,923	0,202
	$*(kurtosis(K_{dc}))^{-1}$	99,914	0,879	0,450	0,450	99,906	99,923	0,202
	$meff/Var(y)$	99,914	0,326	0,446	0,446	99,905	99,923	0,199
	$(meff)^{-1}$	99,912	-1,367	0,446	0,447	99,904	99,921	0,199
10	1/n	99,911	-2,448	0,306	0,307	99,905	99,917	0,095
	$(Var(\hat{\theta}_a))^{-1}$	99,909	-4,869	0,311	0,311	99,903	99,915	0,099
	$*(Var(\hat{\theta}_a))^{-1}$	99,910	-3,277	0,307	0,308	99,904	99,916	0,096
	$(CV(\hat{\theta}_a))^{-1}$	99,910	-3,555	0,308	0,309	99,904	99,916	0,096

	$*(CV(\hat{\theta}_a))^{-1}$	99,911	-2,758	0,307	0,307	99,905	99,917	0,095
	$(skewness(s_{dc}))^{-1}$	99,912	-1,623	0,307	0,307	99,906	99,918	0,094
	$*(skewness(s_{dc}))^{-1}$	99,912	-1,623	0,307	0,307	99,906	99,918	0,094
	$(kurtosis(K_{dc}))^{-1}$	99,911	-2,180	0,307	0,307	99,905	99,917	0,095
	$*(kurtosis(K_{dc}))^{-1}$	99,911	-2,180	0,307	0,307	99,905	99,917	0,095
	$mef/Var(y)$	99,910	-3,640	0,304	0,304	99,904	99,916	0,094
	$(mef)^{-1}$	99,909	-4,582	0,307	0,307	99,903	99,915	0,096
15	$1/n$	99,911	-2,283	0,236	0,237	99,907	99,916	0,056
	$(Var(\hat{\theta}_a))^{-1}$	99,909	-4,670	0,239	0,239	99,904	99,914	0,059
	$*(Var(\hat{\theta}_a))^{-1}$	99,911	-3,085	0,237	0,237	99,906	99,915	0,057
	$(CV(\hat{\theta}_a))^{-1}$	99,910	-3,374	0,237	0,238	99,906	99,915	0,058
	$*(CV(\hat{\theta}_a))^{-1}$	99,911	-2,581	0,237	0,237	99,906	99,916	0,057
	$(skewness(s_{dc}))^{-1}$	99,912	-1,615	0,238	0,238	99,907	99,917	0,057
	$*(skewness(s_{dc}))^{-1}$	99,912	-1,615	0,238	0,238	99,907	99,917	0,057
	$(kurtosis(K_{dc}))^{-1}$	99,912	-2,065	0,237	0,237	99,907	99,916	0,057
	$*(kurtosis(K_{dc}))^{-1}$	99,912	-2,065	0,237	0,237	99,907	99,916	0,057
	$mef/Var(y)$	99,910	-3,742	0,246	0,246	99,905	99,915	0,062
	$(mef)^{-1}$	99,910	-3,880	0,238	0,238	99,905	99,914	0,058
20	$1/n$	99,911	-2,840	0,217	0,217	99,907	99,915	0,048
	$(Var(\hat{\theta}_a))^{-1}$	99,908	-5,184	0,219	0,219	99,904	99,913	0,051
	$*(Var(\hat{\theta}_a))^{-1}$	99,910	-3,622	0,217	0,217	99,906	99,914	0,048
	$(CV(\hat{\theta}_a))^{-1}$	99,910	-3,910	0,218	0,218	99,905	99,914	0,049
	$*(CV(\hat{\theta}_a))^{-1}$	99,910	-3,130	0,217	0,217	99,906	99,915	0,048
	$(skewness(s_{dc}))^{-1}$	99,911	-2,205	0,218	0,218	99,907	99,916	0,048
	$*(skewness(s_{dc}))^{-1}$	99,911	-2,205	0,218	0,218	99,907	99,916	0,048
	$(kurtosis(K_{dc}))^{-1}$	99,911	-2,643	0,217	0,218	99,907	99,915	0,048
	$*(kurtosis(K_{dc}))^{-1}$	99,911	-2,643	0,217	0,218	99,907	99,915	0,048
	$mef/Var(y)$	99,910	-3,478	0,216	0,216	99,906	99,914	0,048
	$(mef)^{-1}$	99,909	-4,760	0,226	0,226	99,904	99,913	0,053

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Appendix C1

Comparing weighting strategies: Stratified Random Sampling: t-distribution, 3 d.f

Scenario 1: N=1 000 000 t(3)(100,25) n=1000 D=20, $\theta = 99.998$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		100,008	9,871	3,316	1,048	100,002	100,015	1,109
2	1/n	100,003	5,130	2,350	0,743	99,999	100,008	0,555
	$(Var(\hat{\theta}_a))^{-1}$	100,005	6,581	2,346	0,742	100,000	100,009	0,555
	$*(Var(\hat{\theta}_a))^{-1}$	100,004	6,172	2,351	0,743	100,000	100,009	0,556
	$(CV(\hat{\theta}_a))^{-1}$	100,005	6,240	2,339	0,740	100,000	100,009	0,551
	$*(CV(\hat{\theta}_a))^{-1}$	100,004	5,952	2,350	0,743	100,000	100,009	0,556
	$(skewness(s_{dc}))^{-1}$	100,003	4,996	2,640	0,835	99,998	100,008	0,699
	$*(skewness(s_{dc}))^{-1}$	100,003	4,494	2,641	0,835	99,998	100,008	0,700
	$(kurtosis(K_{dc}))^{-1}$	100,004	6,045	2,488	0,787	99,999	100,009	0,623
	$*(kurtosis(K_{dc}))^{-1}$	100,004	6,141	2,493	0,788	100,000	100,009	0,625
	$meff/Var(y)$	100,004	6,163	2,351	0,743	100,000	100,009	0,556
	$(meff)^{-1}$	100,004	6,174	2,335	0,738	100,000	100,009	0,549
3	1/n	100,000	2,155	1,979	0,626	99,997	100,004	0,392
	$(Var(\hat{\theta}_a))^{-1}$	100,002	3,418	1,966	0,622	99,998	100,006	0,388
	$*(Var(\hat{\theta}_a))^{-1}$	100,001	2,943	1,979	0,626	99,997	100,005	0,393
	$(CV(\hat{\theta}_a))^{-1}$	100,001	3,114	1,967	0,622	99,998	100,005	0,388
	$*(CV(\hat{\theta}_a))^{-1}$	100,001	2,796	1,981	0,627	99,997	100,005	0,393
	$(skewness(s_{dc}))^{-1}$	100,001	2,456	2,376	0,751	99,996	100,005	0,565
	$*(skewness(s_{dc}))^{-1}$	100,000	1,887	2,383	0,754	99,996	100,005	0,568
	$(kurtosis(K_{dc}))^{-1}$	100,001	2,918	2,115	0,669	99,997	100,005	0,448
	$*(kurtosis(K_{dc}))^{-1}$	100,001	2,950	2,119	0,670	99,997	100,005	0,450
	$meff/Var(y)$	100,001	2,994	1,980	0,626	99,997	100,005	0,393
	$(meff)^{-1}$	100,001	3,106	1,961	0,620	99,998	100,005	0,385
5	1/n	100,001	2,256	1,537	0,486	99,998	100,004	0,237
	$(Var(\hat{\theta}_a))^{-1}$	100,001	2,547	1,499	0,474	99,998	100,004	0,225
	$*(Var(\hat{\theta}_a))^{-1}$	100,001	2,418	1,511	0,478	99,998	100,004	0,229
	$(CV(\hat{\theta}_a))^{-1}$	100,001	2,568	1,507	0,477	99,998	100,004	0,228
	$*(CV(\hat{\theta}_a))^{-1}$	100,001	2,459	1,521	0,481	99,998	100,004	0,232
	$(skewness(s_{dc}))^{-1}$	100,000	1,188	2,051	0,649	99,995	100,004	0,421
	$*(skewness(s_{dc}))^{-1}$	99,999	0,255	2,069	0,654	99,995	100,003	0,428
	$(kurtosis(K_{dc}))^{-1}$	100,000	1,823	1,575	0,498	99,997	100,003	0,248
	$*(kurtosis(K_{dc}))^{-1}$	100,000	1,891	1,575	0,498	99,997	100,003	0,248
	$meff/Var(y)$	100,001	2,496	1,514	0,479	99,998	100,004	0,230
	$(meff)^{-1}$	100,001	2,551	1,510	0,477	99,998	100,004	0,229
10	1/n	100,000	1,272	1,084	0,343	99,997	100,002	0,118

	$(Var(\hat{\theta}_a))^{-1}$	100,000	1,964	1,050	0,332	99,998	100,002	0,111
	$*(Var(\hat{\theta}_a))^{-1}$	100,000	1,604	1,063	0,336	99,998	100,002	0,113
	$(CV(\hat{\theta}_a))^{-1}$	100,000	1,827	1,058	0,335	99,998	100,002	0,112
	$*(CV(\hat{\theta}_a))^{-1}$	100,000	1,588	1,070	0,338	99,998	100,002	0,115
	$(skewness(s_{dc}))^{-1}$	100,002	4,164	1,699	0,537	99,999	100,006	0,290
	$*(skewness(s_{dc}))^{-1}$	100,001	2,542	1,748	0,553	99,997	100,004	0,306
	$(kurtosis(K_{dc}))^{-1}$	100,000	1,496	1,103	0,349	99,998	100,002	0,122
	$*(kurtosis(K_{dc}))^{-1}$	100,000	1,374	1,101	0,348	99,998	100,002	0,121
	$meff/Var(y)$	100,000	1,603	1,062	0,336	99,998	100,002	0,113
	$(meff)^{-1}$	100,000	1,818	1,061	0,336	99,998	100,002	0,113
15	$1/n$	99,999	1,135	0,887	0,281	99,998	100,001	0,079
	$(Var(\hat{\theta}_a))^{-1}$	100,000	1,806	0,860	0,272	99,998	100,002	0,074
	$*(Var(\hat{\theta}_a))^{-1}$	100,000	1,524	0,867	0,274	99,998	100,002	0,075
	$(CV(\hat{\theta}_a))^{-1}$	100,000	1,697	0,866	0,274	99,998	100,002	0,075
	$*(CV(\hat{\theta}_a))^{-1}$	100,000	1,488	0,874	0,276	99,998	100,002	0,077
	$(skewness(s_{dc}))^{-1}$	100,003	5,127	1,564	0,495	100,000	100,007	0,247
	$*(skewness(s_{dc}))^{-1}$	100,001	2,638	1,621	0,513	99,998	100,004	0,264
	$(kurtosis(K_{dc}))^{-1}$	100,000	1,668	0,914	0,289	99,998	100,002	0,084
	$*(kurtosis(K_{dc}))^{-1}$	100,000	1,558	0,910	0,288	99,998	100,002	0,083
	$meff/Var(y)$	100,000	1,522	0,867	0,274	99,998	100,002	0,075
	$(meff)^{-1}$	100,000	1,626	0,870	0,275	99,998	100,002	0,076
20	$1/n$	100,000	1,221	0,772	0,244	99,998	100,001	0,060
	$(Var(\hat{\theta}_a))^{-1}$	100,000	1,811	0,749	0,237	99,999	100,002	0,056
	$*(Var(\hat{\theta}_a))^{-1}$	100,000	1,613	0,754	0,239	99,998	100,001	0,057
	$(CV(\hat{\theta}_a))^{-1}$	100,000	1,758	0,754	0,238	99,999	100,002	0,057
	$*(CV(\hat{\theta}_a))^{-1}$	100,000	1,589	0,760	0,240	99,998	100,001	0,058
	$(skewness(s_{dc}))^{-1}$	100,002	3,237	1,509	0,477	99,999	100,005	0,229
	$*(skewness(s_{dc}))^{-1}$	100,001	2,965	1,532	0,484	99,998	100,004	0,235
	$(kurtosis(K_{dc}))^{-1}$	100,000	1,677	0,788	0,249	99,998	100,002	0,062
	$*(kurtosis(K_{dc}))^{-1}$	100,000	1,633	0,787	0,249	99,998	100,001	0,062
	$meff/Var(y)$	100,000	1,693	0,759	0,240	99,999	100,002	0,058
	$(meff)^{-1}$	100,000	1,665	0,759	0,240	99,998	100,001	0,058

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 2: N=1000 000 t(3)(100,25) n=5000 D=20 $\theta = 99.998$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		99,997	-1,094	1,604	5,074	99,994	100,000	2,575
2	1/n	99,997	-0,851	1,120	3,543	99,995	100,000	1,256
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,208	1,123	3,551	99,996	100,001	1,261
	$*(Var(\hat{\theta}_a))^{-1}$	99,998	-0,064	1,126	3,559	99,996	100,000	1,267
	$(CV(\hat{\theta}_a))^{-1}$	99,998	-0,005	1,123	3,553	99,996	100,001	1,262
	$*(CV(\hat{\theta}_a))^{-1}$	99,998	-0,150	1,126	3,562	99,996	100,000	1,269
	$(skewness(s_{dc}))^{-1}$	100,000	1,372	1,273	4,025	99,997	100,002	1,622
	$*(skewness(s_{dc}))^{-1}$	99,999	0,760	1,290	4,079	99,997	100,002	1,664
	$(kurtosis(K_{dc}))^{-1}$	99,999	0,979	1,213	3,837	99,997	100,002	1,473
	$*(kurtosis(K_{dc}))^{-1}$	99,999	0,859	1,218	3,850	99,997	100,002	1,483
	$meff/Var(y)$	99,998	-0,064	1,126	3,559	99,996	100,000	1,267
	$(meff)^{-1}$	99,998	0,027	1,123	3,550	99,996	100,001	1,261
3	1/n	99,996	-2,045	0,919	2,907	99,994	99,998	0,849
	$(Var(\hat{\theta}_a))^{-1}$	99,997	-0,985	0,928	2,934	99,996	99,999	0,862
	$*(Var(\hat{\theta}_a))^{-1}$	99,997	-1,281	0,927	2,932	99,995	99,999	0,861
	$(CV(\hat{\theta}_a))^{-1}$	99,997	-1,216	0,926	2,930	99,995	99,999	0,860
	$*(CV(\hat{\theta}_a))^{-1}$	99,997	-1,384	0,928	2,934	99,995	99,999	0,862
	$(skewness(s_{dc}))^{-1}$	99,998	-0,630	1,134	3,587	99,995	100,000	1,287
	$*(skewness(s_{dc}))^{-1}$	99,997	-1,087	1,143	3,616	99,995	99,999	1,308
	$(kurtosis(K_{dc}))^{-1}$	99,998	-0,231	1,016	3,214	99,996	100,000	1,033
	$*(kurtosis(K_{dc}))^{-1}$	99,998	-0,349	1,017	3,217	99,996	100,000	1,035
	$meff/Var(y)$	99,997	-1,291	0,927	2,931	99,995	99,999	0,861
	$(meff)^{-1}$	99,997	-1,193	0,927	2,932	99,995	99,999	0,861
5	1/n	99,997	-1,054	0,709	2,241	99,996	99,999	0,503
	$(Var(\hat{\theta}_a))^{-1}$	99,998	-0,532	0,706	2,233	99,996	99,999	0,499
	$*(Var(\hat{\theta}_a))^{-1}$	99,998	-0,817	0,706	2,233	99,996	99,999	0,499
	1	99,998	-0,738	0,705	2,231	99,996	99,999	0,498
	$(CV(\hat{\theta}_a))^{-1}$	99,997	-0,909	0,707	2,235	99,996	99,999	0,501
	$*(CV(\hat{\theta}_a))^{-1}$	99,999	0,241	0,980	3,100	99,997	100,000	0,961
	$(skewness(s_{dc}))^{-1}$	99,999	0,222	0,987	3,123	99,997	100,000	0,975
	$*(skewness(s_{dc}))^{-1}$	99,998	0,168	0,783	2,476	99,997	100,000	0,613
	$(kurtosis(K_{dc}))^{-1}$	99,998	0,054	0,782	2,474	99,997	100,000	0,612
	$*(kurtosis(K_{dc}))^{-1}$	99,998	-0,782	0,708	2,238	99,996	99,999	0,501
	$meff/Var(y)$	99,998	-0,703	0,706	2,233	99,996	99,999	0,499
	$(meff)^{-1}$	99,998	-0,045	0,499	1,577	99,997	99,999	0,249
10	1/n	99,999	0,386	0,493	1,560	99,998	100,000	0,243
	$(Var(\hat{\theta}_a))^{-1}$	99,998	0,156	0,495	1,565	99,998	99,999	0,245
	$*(Var(\hat{\theta}_a))^{-1}$	99,999	0,229	0,494	1,563	99,998	100,000	0,244

	$(CV(\hat{\theta}_d))^{-1}$	99,998	0,086	0,496	1,570	99,997	99,999	0,246
	$*(CV(\hat{\theta}_d))^{-1}$	99,999	0,677	0,792	2,506	99,997	100,001	0,628
	$(\text{skewness}(s_{dc}))^{-1}$	99,999	0,321	0,821	2,595	99,997	100,000	0,674
	$*(\text{skewness}(s_{dc}))^{-1}$	99,999	0,667	0,542	1,713	99,998	100,000	0,294
	$(\text{kurtosis}(K_{dc}))^{-1}$	99,999	0,608	0,542	1,713	99,998	100,000	0,294
	$*(\text{kurtosis}(K_{dc}))^{-1}$	99,998	0,172	0,495	1,564	99,998	99,999	0,245
	$mef f / Var(y)$	99,999	0,269	0,494	1,563	99,998	100,000	0,244
	$(mef f)^{-1}$	99,998	0,033	0,407	1,287	99,998	99,999	0,166
15	$1/n$	99,999	0,438	0,402	1,270	99,998	100,000	0,161
	$(Var(\hat{\theta}_d))^{-1}$	99,999	0,240	0,404	1,276	99,998	99,999	0,163
	$*(Var(\hat{\theta}_d))^{-1}$	99,999	0,300	0,403	1,274	99,998	99,999	0,162
	$(CV(\hat{\theta}_d))^{-1}$	99,998	0,169	0,405	1,281	99,998	99,999	0,164
	$*(CV(\hat{\theta}_d))^{-1}$	99,999	0,721	0,744	2,352	99,998	100,000	0,554
	$(\text{skewness}(s_{dc}))^{-1}$	99,999	0,368	0,759	2,399	99,997	100,000	0,576
	$*(\text{skewness}(s_{dc}))^{-1}$	99,999	0,591	0,433	1,369	99,998	100,000	0,188
	$(\text{kurtosis}(K_{dc}))^{-1}$	99,999	0,547	0,434	1,371	99,998	100,000	0,188
	$*(\text{kurtosis}(K_{dc}))^{-1}$	99,999	0,242	0,404	1,277	99,998	99,999	0,163
	$mef f / Var(y)$	99,999	0,295	0,403	1,275	99,998	99,999	0,163
	$(mef f)^{-1}$	99,998	-0,395	0,356	1,127	99,997	99,999	0,127
20	$1/n$	99,998	0,018	0,352	1,113	99,998	99,999	0,124
	$(Var(\hat{\theta}_d))^{-1}$	99,998	-0,180	0,353	1,118	99,997	99,999	0,125
	$*(Var(\hat{\theta}_d))^{-1}$	99,998	-0,122	0,353	1,116	99,998	99,999	0,125
	$(CV(\hat{\theta}_d))^{-1}$	99,998	-0,254	0,355	1,121	99,997	99,999	0,126
	$*(CV(\hat{\theta}_d))^{-1}$	99,999	0,954	0,697	2,205	99,998	100,001	0,487
	$(\text{skewness}(s_{dc}))^{-1}$	99,999	0,262	0,711	2,248	99,997	100,000	0,505
	$*(\text{skewness}(s_{dc}))^{-1}$	99,998	0,160	0,380	1,202	99,998	99,999	0,145
	$(\text{kurtosis}(K_{dc}))^{-1}$	99,998	0,125	0,380	1,201	99,998	99,999	0,144
	$*(\text{kurtosis}(K_{dc}))^{-1}$	99,998	-0,178	0,353	1,118	99,997	99,999	0,125
	$mef f / Var(y)$	99,998	-0,131	0,353	1,117	99,997	99,999	0,125
	$(mef f)^{-1}$	99,997	-1,094	1,604	5,074	99,994	100,000	2,575

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 3: N=1000 000 t(3)(100,25) n=10000 D=20 $\theta = 99.998$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		100,000	1,207	1,093	3,456	99,997	100,002	1,196
2	1/n	99,999	0,622	0,783	2,477	99,997	100,000	0,614
	$(Var(\hat{\theta}_a))^{-1}$	100,000	1,354	0,789	2,496	99,998	100,001	0,625
	$*(Var(\hat{\theta}_a))^{-1}$	99,999	1,173	0,787	2,489	99,998	100,001	0,621
	$(CV(\hat{\theta}_a))^{-1}$	100,000	1,207	0,788	2,490	99,998	100,001	0,622
	$*(CV(\hat{\theta}_a))^{-1}$	99,999	1,113	0,787	2,489	99,998	100,001	0,621
	$(skewness(s_{dc}))^{-1}$	100,000	2,095	0,887	2,804	99,999	100,002	0,791
	$*(skewness(s_{dc}))^{-1}$	100,001	2,191	0,890	2,814	99,999	100,002	0,797
	$(kurtosis(K_{dc}))^{-1}$	100,000	2,179	0,860	2,720	99,999	100,002	0,744
	$*(kurtosis(K_{dc}))^{-1}$	100,000	2,114	0,863	2,729	99,999	100,002	0,749
	$meff/Var(y)$	99,999	1,173	0,787	2,489	99,998	100,001	0,621
	$(meff)^{-1}$	100,000	1,228	0,788	2,493	99,998	100,001	0,623
3	1/n	99,999	1,149	0,635	2,009	99,998	100,001	0,405
	$(Var(\hat{\theta}_a))^{-1}$	100,000	1,723	0,635	2,009	99,999	100,001	0,406
	$*(Var(\hat{\theta}_a))^{-1}$	100,000	1,578	0,634	2,006	99,999	100,001	0,405
	$(CV(\hat{\theta}_a))^{-1}$	100,000	1,608	0,635	2,007	99,999	100,001	0,405
	$*(CV(\hat{\theta}_a))^{-1}$	100,000	1,529	0,635	2,007	99,999	100,001	0,405
	$(skewness(s_{dc}))^{-1}$	100,001	2,296	0,776	2,453	99,999	100,002	0,607
	$*(skewness(s_{dc}))^{-1}$	100,001	2,204	0,781	2,469	99,999	100,002	0,614
	$(kurtosis(K_{dc}))^{-1}$	100,001	2,317	0,703	2,222	99,999	100,002	0,499
	$*(kurtosis(K_{dc}))^{-1}$	100,001	2,270	0,706	2,233	99,999	100,002	0,504
	$meff/Var(y)$	100,000	1,600	0,635	2,007	99,999	100,001	0,405
	$(meff)^{-1}$	100,000	1,616	0,635	2,009	99,999	100,001	0,406
5	1/n	99,999	0,683	0,479	1,516	99,998	100,000	0,230
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,996	0,476	1,505	99,998	100,000	0,228
	$*(Var(\hat{\theta}_a))^{-1}$	99,999	0,834	0,478	1,511	99,998	100,000	0,229
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,862	0,477	1,508	99,998	100,000	0,228
	$*(CV(\hat{\theta}_a))^{-1}$	99,999	0,770	0,478	1,513	99,998	100,000	0,229
	$(skewness(s_{dc}))^{-1}$	100,000	2,084	0,658	2,081	99,999	100,002	0,437
	$*(skewness(s_{dc}))^{-1}$	100,000	1,773	0,671	2,123	99,999	100,001	0,454
	$(kurtosis(K_{dc}))^{-1}$	100,000	1,607	0,534	1,690	99,999	100,001	0,288
	$*(kurtosis(K_{dc}))^{-1}$	100,000	1,549	0,537	1,699	99,999	100,001	0,291
	$meff/Var(y)$	99,999	0,859	0,479	1,513	99,998	100,000	0,230
	$(meff)^{-1}$	99,999	0,866	0,476	1,507	99,998	100,000	0,228
10	1/n	99,999	0,210	0,346	1,093	99,998	99,999	0,119
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,533	0,344	1,088	99,998	100,000	0,119
	$*(Var(\hat{\theta}_a))^{-1}$	99,999	0,347	0,345	1,090	99,998	99,999	0,119
	$(CV(\hat{\theta}_a))^{-1}$	99,999	0,402	0,344	1,089	99,998	99,999	0,119

	$*(CV(\hat{\theta}_a))^{-1}$	99,999	0,292	0,345	1,091	99,998	99,999	0,119
	$(skewness(s_{dc}))^{-1}$	100,000	1,764	0,544	1,719	99,999	100,001	0,299
	$*(skewness(s_{dc}))^{-1}$	100,000	1,878	0,557	1,762	99,999	100,001	0,314
	$(kurtosis(K_{dc}))^{-1}$	99,999	1,024	0,376	1,191	99,999	100,000	0,143
	$*(kurtosis(K_{dc}))^{-1}$	99,999	0,973	0,377	1,193	99,999	100,000	0,143
	$mef/Var(y)$	99,999	0,336	0,345	1,090	99,998	99,999	0,119
	$(mef)^{-1}$	99,999	0,407	0,345	1,091	99,998	99,999	0,119
15	$1/n$	99,998	-0,010	0,282	0,892	99,998	99,999	0,079
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,273	0,281	0,887	99,998	99,999	0,079
	$*(Var(\hat{\theta}_a))^{-1}$	99,998	0,114	0,281	0,888	99,998	99,999	0,079
	$(CV(\hat{\theta}_a))^{-1}$	99,998	0,161	0,281	0,888	99,998	99,999	0,079
	$*(CV(\hat{\theta}_a))^{-1}$	99,998	0,066	0,281	0,890	99,998	99,999	0,079
	$(skewness(s_{dc}))^{-1}$	100,000	1,297	0,520	1,643	99,999	100,001	0,272
	$*(skewness(s_{dc}))^{-1}$	99,999	1,073	0,520	1,643	99,998	100,000	0,271
	$(kurtosis(K_{dc}))^{-1}$	99,999	0,654	0,311	0,984	99,998	100,000	0,097
	$*(kurtosis(K_{dc}))^{-1}$	99,999	0,614	0,312	0,986	99,998	100,000	0,098
	$mef/Var(y)$	99,998	0,130	0,281	0,888	99,998	99,999	0,079
	$(mef)^{-1}$	99,999	0,190	0,281	0,888	99,998	99,999	0,079
20	$1/n$	99,998	0,006	0,240	0,759	99,998	99,999	0,058
	$(Var(\hat{\theta}_a))^{-1}$	99,999	0,275	0,238	0,753	99,998	99,999	0,057
	$*(Var(\hat{\theta}_a))^{-1}$	99,998	0,123	0,239	0,756	99,998	99,999	0,057
	$(CV(\hat{\theta}_a))^{-1}$	99,998	0,171	0,239	0,755	99,998	99,999	0,057
	$*(CV(\hat{\theta}_a))^{-1}$	99,998	0,079	0,240	0,757	99,998	99,999	0,057
	$(skewness(s_{dc}))^{-1}$	99,999	0,703	0,506	1,599	99,998	100,000	0,256
	$*(skewness(s_{dc}))^{-1}$	99,999	0,578	0,514	1,626	99,998	100,000	0,265
	$(kurtosis(K_{dc}))^{-1}$	99,999	0,538	0,266	0,841	99,998	99,999	0,071
	$*(kurtosis(K_{dc}))^{-1}$	99,999	0,510	0,267	0,845	99,998	99,999	0,072
	$mef/Var(y)$	99,998	0,131	0,239	0,757	99,998	99,999	0,057
	$(mef)^{-1}$	99,999	0,183	0,238	0,754	99,998	99,999	0,057

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Appendix C2

Comparing weighting strategies: Cluster Random Sampling: t-distribution, 3 d.f

Scenario 1: N=1 000 000 t(3)(100,25) n=1000 D=20, $\theta = 100.010$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,981	-2,931	3,051	3,051	99,921	100,041	9,393
2	1/n	100,015	0,461	1,881	1,880	99,978	100,052	3,539
	$(Var(\hat{\theta}_a))^{-1}$	100,006	-0,389	2,199	2,199	99,963	100,049	4,839
	$*(Var(\hat{\theta}_a))^{-1}$	100,003	-0,704	2,177	2,177	99,960	100,046	4,744
	$(CV(\hat{\theta}_a))^{-1}$	100,005	-0,480	2,176	2,176	99,963	100,048	4,739
	$*(CV(\hat{\theta}_a))^{-1}$	100,004	-0,653	2,174	2,174	99,961	100,046	4,732
	$(skewness(s_{dc}))^{-1}$	100,002	-0,791	2,538	2,538	99,952	100,052	6,448
	$*(skewness(s_{dc}))^{-1}$	100,002	-0,791	2,538	2,538	99,952	100,052	6,448
	$(kurtosis(K_{dc}))^{-1}$	99,999	-1,145	2,367	2,367	99,952	100,045	5,614
	$*(kurtosis(K_{dc}))^{-1}$	99,999	-1,145	2,367	2,367	99,952	100,045	5,614
	$meff/Var(y)$	100,003	-0,727	2,175	2,175	99,960	100,046	4,735
	$(meff)^{-1}$	100,007	-0,358	2,200	2,199	99,963	100,050	4,839
3	1/n	100,019	0,908	1,731	1,731	99,985	100,053	3,005
	$(Var(\hat{\theta}_a))^{-1}$	100,009	-0,069	1,968	1,968	99,971	100,048	3,875
	$*(Var(\hat{\theta}_a))^{-1}$	100,007	-0,266	1,958	1,958	99,969	100,046	3,835
	$(CV(\hat{\theta}_a))^{-1}$	100,009	-0,081	1,958	1,958	99,971	100,048	3,834
	$*(CV(\hat{\theta}_a))^{-1}$	100,008	-0,167	1,969	1,968	99,970	100,047	3,876
	$(skewness(s_{dc}))^{-1}$	100,010	-0,058	2,394	2,394	99,963	100,056	5,731
	$*(skewness(s_{dc}))^{-1}$	100,010	-0,058	2,394	2,394	99,963	100,056	5,731
	$(kurtosis(K_{dc}))^{-1}$	100,001	-0,923	2,201	2,201	99,958	100,044	4,852
	$*(kurtosis(K_{dc}))^{-1}$	100,001	-0,923	2,201	2,201	99,958	100,044	4,852
	$meff/Var(y)$	100,005	-0,531	2,005	2,005	99,966	100,044	4,023
	$(meff)^{-1}$	100,009	-0,153	1,995	1,995	99,970	100,048	3,981
5	1/n	100,007	-0,338	1,469	1,469	99,978	100,036	2,159
	$(Var(\hat{\theta}_a))^{-1}$	100,005	-0,480	1,477	1,477	99,976	100,034	2,185
	$*(Var(\hat{\theta}_a))^{-1}$	100,006	-0,465	1,441	1,441	99,977	100,034	2,078
	$(CV(\hat{\theta}_a))^{-1}$	100,006	-0,376	1,459	1,458	99,978	100,035	2,129
	$*(CV(\hat{\theta}_a))^{-1}$	100,007	-0,315	1,450	1,450	99,979	100,035	2,104
	$(skewness(s_{dc}))^{-1}$	99,987	-2,302	2,000	2,000	99,948	100,026	4,052
	$*(skewness(s_{dc}))^{-1}$	99,987	-2,302	2,000	2,000	99,948	100,026	4,052
	$(kurtosis(K_{dc}))^{-1}$	100,000	-0,968	1,576	1,576	99,970	100,031	2,492
	$*(kurtosis(K_{dc}))^{-1}$	100,000	-0,968	1,576	1,576	99,970	100,031	2,492
	$meff/Var(y)$	100,005	-0,528	1,443	1,443	99,977	100,033	2,085
	$(meff)^{-1}$	100,005	-0,541	1,516	1,516	99,975	100,034	2,301
10	1/n	100,003	-0,750	0,903	0,903	99,985	100,020	0,821

	$(Var(\hat{\theta}_a))^{-1}$	99,994	-1,590	0,930	0,930	99,976	100,012	0,891
	$*(Var(\hat{\theta}_a))^{-1}$	100,001	-0,919	0,894	0,894	99,983	100,018	0,807
	$(CV(\hat{\theta}_a))^{-1}$	99,999	-1,143	0,901	0,901	99,981	100,016	0,825
	$*(CV(\hat{\theta}_a))^{-1}$	100,003	-0,725	0,894	0,894	99,985	100,020	0,805
	$(skewness(s_{dc}))^{-1}$	100,004	-0,619	1,764	1,764	99,969	100,039	3,116
	$*(skewness(s_{dc}))^{-1}$	100,004	-0,619	1,764	1,764	99,969	100,039	3,116
	$(kurtosis(K_{dc}))^{-1}$	100,003	-0,683	1,033	1,033	99,983	100,024	1,073
	$*(kurtosis(K_{dc}))^{-1}$	100,003	-0,683	1,033	1,033	99,983	100,024	1,073
	$meff/Var(y)$	100,002	-0,813	0,917	0,917	99,984	100,020	0,848
	$(meff)^{-1}$	99,995	-1,548	0,976	0,976	99,976	100,014	0,977
15	$1/n$	100,010	0,002	0,749	0,749	99,995	100,025	0,561
	$(Var(\hat{\theta}_a))^{-1}$	99,996	-1,426	0,729	0,729	99,982	100,010	0,552
	$*(Var(\hat{\theta}_a))^{-1}$	100,006	-0,407	0,732	0,732	99,992	100,020	0,537
	$(CV(\hat{\theta}_a))^{-1}$	100,003	-0,707	0,728	0,728	99,989	100,017	0,535
	$*(CV(\hat{\theta}_a))^{-1}$	100,009	-0,104	0,737	0,737	99,995	100,024	0,543
	$(skewness(s_{dc}))^{-1}$	100,005	-0,466	1,786	1,786	99,970	100,040	3,191
	$*(skewness(s_{dc}))^{-1}$	100,005	-0,466	1,786	1,786	99,970	100,040	3,191
	$(kurtosis(K_{dc}))^{-1}$	100,012	0,137	0,884	0,884	99,994	100,029	0,782
	$*(kurtosis(K_{dc}))^{-1}$	100,012	0,137	0,884	0,884	99,994	100,029	0,782
	$meff/Var(y)$	100,004	-0,593	0,758	0,758	99,989	100,019	0,578
	$(meff)^{-1}$	99,996	-1,430	0,773	0,774	99,981	100,011	0,619
20	$1/n$	100,014	0,362	0,746	0,746	99,999	100,028	0,557
	$(Var(\hat{\theta}_a))^{-1}$	100,000	-1,027	0,733	0,733	99,986	100,014	0,548
	$*(Var(\hat{\theta}_a))^{-1}$	100,009	-0,075	0,735	0,735	99,995	100,024	0,540
	$(CV(\hat{\theta}_a))^{-1}$	100,007	-0,319	0,732	0,731	99,993	100,021	0,536
	$*(CV(\hat{\theta}_a))^{-1}$	100,013	0,243	0,737	0,737	99,998	100,027	0,544
	$(skewness(s_{dc}))^{-1}$	100,002	-0,787	1,695	1,695	99,969	100,036	2,878
	$*(skewness(s_{dc}))^{-1}$	100,002	-0,787	1,695	1,695	99,969	100,036	2,878
	$(kurtosis(K_{dc}))^{-1}$	100,013	0,321	0,843	0,843	99,997	100,030	0,711
	$*(kurtosis(K_{dc}))^{-1}$	100,013	0,321	0,843	0,843	99,997	100,030	0,711
	$meff/Var(y)$	100,007	-0,351	0,789	0,789	99,991	100,022	0,624
	$(meff)^{-1}$	99,998	-1,213	0,864	0,864	99,981	100,015	0,762

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 2: N=1000 000 t(3)(100,25) n=5000 D=20 $\theta = 100.010$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,994	-1,627	1,527	1,527	99,964	100,024	2,357
2	1/n	100,025	1,532	0,945	0,944	100,007	100,044	0,916
	$(Var(\hat{\theta}_d))^{-1}$	100,010	0,010	1,062	1,062	99,989	100,031	1,128
	$*(Var(\hat{\theta}_d))^{-1}$	100,012	0,195	1,069	1,069	99,991	100,033	1,143
	$(CV(\hat{\theta}_d))^{-1}$	100,012	0,161	1,060	1,059	99,991	100,033	1,123
	$*(CV(\hat{\theta}_d))^{-1}$	100,013	0,254	1,064	1,064	99,992	100,034	1,133
	$(skewness(s_{dc}))^{-1}$	100,007	-0,346	1,354	1,354	99,980	100,033	1,834
	$*(skewness(s_{dc}))^{-1}$	100,007	-0,346	1,354	1,354	99,980	100,033	1,834
	$(kurtosis(K_{dc}))^{-1}$	100,014	0,421	1,250	1,250	99,990	100,039	1,565
	$*(kurtosis(K_{dc}))^{-1}$	100,014	0,421	1,250	1,250	99,990	100,039	1,565
	$meff/Var(y)$	100,012	0,195	1,069	1,069	99,991	100,033	1,143
	$(meff)^{-1}$	100,011	0,110	1,057	1,056	99,991	100,032	1,117
3	1/n	100,021	1,090	0,715	0,714	100,007	100,035	0,522
	$(Var(\hat{\theta}_d))^{-1}$	100,008	-0,265	0,801	0,801	99,992	100,023	0,642
	$*(Var(\hat{\theta}_d))^{-1}$	100,010	-0,040	0,811	0,811	99,994	100,026	0,658
	$(CV(\hat{\theta}_d))^{-1}$	100,009	-0,069	0,797	0,797	99,994	100,025	0,636
	$*(CV(\hat{\theta}_d))^{-1}$	100,011	0,037	0,804	0,804	99,995	100,026	0,646
	$(skewness(s_{dc}))^{-1}$	100,002	-0,854	1,129	1,129	99,979	100,024	1,281
	$*(skewness(s_{dc}))^{-1}$	100,002	-0,854	1,129	1,129	99,979	100,024	1,281
	$(kurtosis(K_{dc}))^{-1}$	100,009	-0,138	0,977	0,977	99,990	100,028	0,955
	$*(kurtosis(K_{dc}))^{-1}$	100,009	-0,138	0,977	0,977	99,990	100,028	0,955
	$meff/Var(y)$	100,011	0,042	0,814	0,814	99,995	100,027	0,663
	$(meff)^{-1}$	100,009	-0,086	0,791	0,791	99,994	100,025	0,625
5	1/n	100,007	-0,364	0,670	0,670	99,993	100,020	0,450
	$(Var(\hat{\theta}_d))^{-1}$	100,004	-0,644	0,669	0,669	99,991	100,017	0,452
	$*(Var(\hat{\theta}_d))^{-1}$	100,006	-0,441	0,673	0,673	99,993	100,019	0,454
	1	100,005	-0,489	0,668	0,668	99,992	100,018	0,449
	$(CV(\hat{\theta}_d))^{-1}$	100,006	-0,389	0,671	0,671	99,993	100,019	0,452
	$*(CV(\hat{\theta}_d))^{-1}$	99,998	-1,254	0,990	0,990	99,978	100,017	0,997
	$(skewness(s_{dc}))^{-1}$	99,998	-1,254	0,990	0,990	99,978	100,017	0,997
	$*(skewness(s_{dc}))^{-1}$	100,005	-0,554	0,770	0,770	99,990	100,020	0,596
	$(kurtosis(K_{dc}))^{-1}$	100,005	-0,554	0,770	0,770	99,990	100,020	0,596
	$*(kurtosis(K_{dc}))^{-1}$	100,005	-0,525	0,680	0,680	99,992	100,018	0,465
	$meff/Var(y)$	100,005	-0,559	0,669	0,669	99,991	100,018	0,451
	$(meff)^{-1}$	100,006	-0,398	0,457	0,457	99,997	100,015	0,211
10	1/n	100,003	-0,743	0,436	0,436	99,994	100,011	0,196
	$(Var(\hat{\theta}_d))^{-1}$	100,006	-0,415	0,449	0,449	99,997	100,015	0,203
	$*(Var(\hat{\theta}_d))^{-1}$	100,005	-0,552	0,445	0,445	99,996	100,013	0,201

	$(CV(\hat{\theta}_d))^{-1}$	100,006	-0,387	0,452	0,452	99,997	100,015	0,206
	$*(CV(\hat{\theta}_d))^{-1}$	99,997	-1,280	0,907	0,907	99,980	100,015	0,840
	$(\text{skewness}(s_{dc}))^{-1}$	99,997	-1,280	0,907	0,907	99,980	100,015	0,840
	$*(\text{skewness}(s_{dc}))^{-1}$	100,008	-0,241	0,536	0,536	99,997	100,018	0,287
	$(\text{kurtosis}(K_{dc}))^{-1}$	100,008	-0,241	0,536	0,536	99,997	100,018	0,287
	$*(\text{kurtosis}(K_{dc}))^{-1}$	100,005	-0,482	0,459	0,459	99,996	100,014	0,213
	$meff/Var(y)$	100,003	-0,742	0,452	0,452	99,994	100,012	0,210
	$(meff)^{-1}$	100,006	-0,371	0,370	0,370	99,999	100,014	0,139
15	$1/n$	100,002	-0,790	0,353	0,353	99,995	100,009	0,131
	$(Var(\hat{\theta}_d))^{-1}$	100,006	-0,431	0,363	0,363	99,999	100,013	0,134
	$*(Var(\hat{\theta}_d))^{-1}$	100,004	-0,566	0,360	0,360	99,997	100,012	0,133
	$(CV(\hat{\theta}_d))^{-1}$	100,006	-0,381	0,366	0,366	99,999	100,014	0,136
	$*(CV(\hat{\theta}_d))^{-1}$	99,999	-1,080	0,760	0,760	99,984	100,014	0,590
	$(\text{skewness}(s_{dc}))^{-1}$	99,999	-1,080	0,760	0,760	99,984	100,014	0,590
	$*(\text{skewness}(s_{dc}))^{-1}$	100,007	-0,338	0,404	0,404	99,999	100,015	0,165
	$(\text{kurtosis}(K_{dc}))^{-1}$	100,007	-0,338	0,404	0,404	99,999	100,015	0,165
	$*(\text{kurtosis}(K_{dc}))^{-1}$	100,006	-0,461	0,365	0,365	99,998	100,013	0,135
	$meff/Var(y)$	100,003	-0,723	0,362	0,362	99,996	100,010	0,136
	$(meff)^{-1}$	100,008	-0,199	0,327	0,327	100,002	100,015	0,107
20	$1/n$	100,004	-0,635	0,319	0,319	99,998	100,010	0,106
	$(Var(\hat{\theta}_d))^{-1}$	100,007	-0,269	0,321	0,321	100,001	100,014	0,104
	$*(Var(\hat{\theta}_d))^{-1}$	100,006	-0,404	0,322	0,322	100,000	100,012	0,105
	$(CV(\hat{\theta}_d))^{-1}$	100,008	-0,213	0,324	0,324	100,002	100,014	0,105
	$*(CV(\hat{\theta}_d))^{-1}$	100,004	-0,574	0,752	0,752	99,990	100,019	0,568
	$(\text{skewness}(s_{dc}))^{-1}$	100,004	-0,574	0,752	0,752	99,990	100,019	0,568
	$*(\text{skewness}(s_{dc}))^{-1}$	100,007	-0,276	0,368	0,367	100,000	100,015	0,136
	$(\text{kurtosis}(K_{dc}))^{-1}$	100,007	-0,276	0,368	0,367	100,000	100,015	0,136
	$*(\text{kurtosis}(K_{dc}))^{-1}$	100,007	-0,343	0,332	0,332	100,000	100,013	0,111
	$meff/Var(y)$	100,003	-0,672	0,340	0,340	99,997	100,010	0,120
	$(meff)^{-1}$	99,994	-1,627	1,527	1,527	99,964	100,024	2,357

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_d$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 3: N=1000 000 t(3)(100,25) n=10000 D=20 $\theta = 99.998$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		99,994	-1,574	9,449	9,449	99,976	100,013	9,176
2	1/n	99,999	-1,082	6,461	6,461	99,987	100,012	4,291
	$(Var(\hat{\theta}_a))^{-1}$	100,003	-0,690	6,410	6,410	99,991	100,016	4,156
	$*(Var(\hat{\theta}_a))^{-1}$	100,005	-0,539	6,488	6,488	99,992	100,017	4,238
	$(CV(\hat{\theta}_a))^{-1}$	100,004	-0,617	6,410	6,410	99,991	100,017	4,147
	$*(CV(\hat{\theta}_a))^{-1}$	100,005	-0,540	6,454	6,453	99,992	100,017	4,194
	$(skewness(s_{dc}))^{-1}$	100,008	-0,208	8,589	8,588	99,991	100,025	7,382
	$*(skewness(s_{dc}))^{-1}$	100,008	-0,208	8,589	8,588	99,991	100,025	7,382
	$(kurtosis(K_{dc}))^{-1}$	100,006	-0,438	7,682	7,682	99,991	100,021	5,921
	$*(kurtosis(K_{dc}))^{-1}$	100,006	-0,438	7,682	7,682	99,991	100,021	5,921
	$meff/Var(y)$	100,005	-0,543	6,491	6,491	99,992	100,017	4,243
	$(meff)^{-1}$	100,003	-0,704	6,379	6,379	99,991	100,016	4,119
3	1/n	99,999	-1,090	5,228	5,228	99,989	100,010	2,852
	$(Var(\hat{\theta}_a))^{-1}$	100,003	-0,694	5,267	5,266	99,993	100,014	2,822
	$*(Var(\hat{\theta}_a))^{-1}$	100,005	-0,540	5,272	5,272	99,994	100,015	2,809
	$(CV(\hat{\theta}_a))^{-1}$	100,004	-0,628	5,255	5,255	99,994	100,014	2,801
	$*(CV(\hat{\theta}_a))^{-1}$	100,005	-0,549	5,261	5,260	99,994	100,015	2,798
	$(skewness(s_{dc}))^{-1}$	100,008	-0,252	7,140	7,139	99,994	100,022	5,104
	$*(skewness(s_{dc}))^{-1}$	100,008	-0,252	7,140	7,139	99,994	100,022	5,104
	$(kurtosis(K_{dc}))^{-1}$	100,007	-0,323	6,070	6,070	99,995	100,019	3,695
	$*(kurtosis(K_{dc}))^{-1}$	100,007	-0,323	6,070	6,070	99,995	100,019	3,695
	$meff/Var(y)$	100,005	-0,561	5,263	5,263	99,994	100,015	2,802
	$(meff)^{-1}$	100,003	-0,734	5,274	5,274	99,992	100,013	2,836
5	1/n	100,002	-0,850	4,491	4,491	99,993	100,010	2,089
	$(Var(\hat{\theta}_a))^{-1}$	100,001	-0,964	4,489	4,489	99,992	100,009	2,108
	$*(Var(\hat{\theta}_a))^{-1}$	100,002	-0,852	4,533	4,533	99,993	100,011	2,128
	$(CV(\hat{\theta}_a))^{-1}$	100,001	-0,899	4,487	4,487	99,992	100,010	2,094
	$*(CV(\hat{\theta}_a))^{-1}$	100,002	-0,844	4,511	4,510	99,993	100,011	2,106
	$(skewness(s_{dc}))^{-1}$	100,010	0,021	6,777	6,777	99,997	100,024	4,593
	$*(skewness(s_{dc}))^{-1}$	100,010	0,021	6,777	6,777	99,997	100,024	4,593
	$(kurtosis(K_{dc}))^{-1}$	100,004	-0,626	5,214	5,214	99,994	100,014	2,758
	$*(kurtosis(K_{dc}))^{-1}$	100,004	-0,626	5,214	5,214	99,994	100,014	2,758
	$meff/Var(y)$	100,002	-0,840	4,545	4,545	99,993	100,011	2,137
	$(meff)^{-1}$	100,001	-0,922	4,512	4,512	99,992	100,010	2,121
10	1/n	100,004	-0,642	3,162	3,162	99,998	100,010	1,041
	$(Var(\hat{\theta}_a))^{-1}$	100,003	-0,756	3,205	3,205	99,996	100,009	1,084
	$*(Var(\hat{\theta}_a))^{-1}$	100,004	-0,646	3,218	3,218	99,997	100,010	1,077
	$(CV(\hat{\theta}_a))^{-1}$	100,003	-0,691	3,179	3,179	99,997	100,009	1,058

	$*(CV(\hat{\theta}_a))^{-1}$	100,004	-0,635	3,189	3,189	99,998	100,010	1,057
	$(skewness(s_{dc}))^{-1}$	100,006	-0,394	5,886	5,886	99,995	100,018	3,480
	$*(skewness(s_{dc}))^{-1}$	100,006	-0,394	5,886	5,886	99,995	100,018	3,480
	$(kurtosis(K_{dc}))^{-1}$	100,006	-0,406	3,931	3,931	99,998	100,014	1,562
	$*(kurtosis(K_{dc}))^{-1}$	100,006	-0,406	3,931	3,931	99,998	100,014	1,562
	$mef/Var(y)$	100,004	-0,604	3,219	3,219	99,998	100,010	1,073
	$(mef)^{-1}$	100,003	-0,735	3,140	3,140	99,997	100,009	1,040
15	$1/n$	100,007	-0,279	2,501	2,501	100,002	100,012	0,633
	$(Var(\hat{\theta}_a))^{-1}$	100,006	-0,420	2,471	2,471	100,001	100,011	0,628
	$*(Var(\hat{\theta}_a))^{-1}$	100,007	-0,284	2,505	2,505	100,002	100,012	0,636
	$(CV(\hat{\theta}_a))^{-1}$	100,007	-0,279	2,501	2,501	100,002	100,012	0,633
	$*(CV(\hat{\theta}_a))^{-1}$	100,007	-0,341	2,481	2,481	100,002	100,012	0,627
	$(skewness(s_{dc}))^{-1}$	100,007	-0,271	2,501	2,501	100,003	100,012	0,633
	$*(skewness(s_{dc}))^{-1}$	100,005	-0,534	5,437	5,436	99,994	100,015	2,984
	$(kurtosis(K_{dc}))^{-1}$	100,005	-0,534	5,437	5,436	99,994	100,015	2,984
	$*(kurtosis(K_{dc}))^{-1}$	100,009	-0,108	3,058	3,058	100,003	100,015	0,936
	$mef/Var(y)$	100,009	-0,108	3,058	3,058	100,003	100,015	0,936
	$(mef)^{-1}$	100,007	-0,324	2,552	2,552	100,002	100,012	0,662
20	$1/n$	100,005	-0,480	2,566	2,566	100,000	100,010	0,681
	$(Var(\hat{\theta}_a))^{-1}$	100,007	-0,313	2,291	2,291	100,003	100,012	0,535
	$*(Var(\hat{\theta}_a))^{-1}$	100,005	-0,470	2,259	2,259	100,001	100,010	0,532
	$(CV(\hat{\theta}_a))^{-1}$	100,007	-0,325	2,285	2,284	100,002	100,011	0,533
	$*(CV(\hat{\theta}_a))^{-1}$	100,006	-0,382	2,271	2,270	100,002	100,011	0,530
	$(skewness(s_{dc}))^{-1}$	100,007	-0,309	2,286	2,286	100,003	100,012	0,532
	$*(skewness(s_{dc}))^{-1}$	100,003	-0,711	4,919	4,919	99,993	100,013	2,470
	$(kurtosis(K_{dc}))^{-1}$	100,003	-0,711	4,919	4,919	99,993	100,013	2,470
	$*(kurtosis(K_{dc}))^{-1}$	100,009	-0,094	2,594	2,594	100,004	100,014	0,674
	$mef/Var(y)$	100,009	-0,094	2,594	2,594	100,004	100,014	0,674
	$(mef)^{-1}$	100,007	-0,327	2,284	2,284	100,002	100,011	0,532

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Appendix D1

Comparing weighting strategies: Stratified Random Sampling: Skewed t-distribution, 3 d.f

Scenario 1: N=1 000 000 at(3)(100,25) n=1000 D=20, $\theta = 99.970$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	$SE(\hat{\theta}_c)$ $\times 10^{-3}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-1}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-2}$
						LB	UB	
1		99,974	0,396	3,509	1,110	99,967	99,981	1,233
2	1/n	99,975	0,490	2,469	0,781	99,970	99,980	0,612
	$(Var(\hat{\theta}_a))^{-1}$	99,965	-0,563	2,474	0,783	99,960	99,969	0,615
	$*(Var(\hat{\theta}_a))^{-1}$	99,970	-0,044	2,455	0,776	99,965	99,975	0,602
	$(CV(\hat{\theta}_a))^{-1}$	99,969	-0,098	2,459	0,778	99,964	99,974	0,605
	$*(CV(\hat{\theta}_a))^{-1}$	99,972	0,182	2,465	0,780	99,967	99,977	0,608
	$(skewness(s_{dc}))^{-1}$	99,960	-1,016	2,626	0,831	99,955	99,965	0,700
	$*(skewness(s_{dc}))^{-1}$	99,960	-0,989	2,630	0,832	99,955	99,966	0,701
	$(kurtosis(K_{dc}))^{-1}$	99,960	-1,004	2,681	0,848	99,955	99,965	0,729
	$*(kurtosis(K_{dc}))^{-1}$	99,961	-0,967	2,675	0,846	99,955	99,966	0,725
	$meff/Var(y)$	99,970	-0,041	2,455	0,776	99,965	99,975	0,603
	$(meff)^{-1}$	99,969	-0,157	2,473	0,782	99,964	99,974	0,612
3	1/n	99,974	0,373	1,980	0,626	99,970	99,978	0,394
	$(Var(\hat{\theta}_a))^{-1}$	99,961	-0,912	1,980	0,626	99,957	99,965	0,400
	$*(Var(\hat{\theta}_a))^{-1}$	99,968	-0,250	1,959	0,620	99,964	99,972	0,384
	$(CV(\hat{\theta}_a))^{-1}$	99,967	-0,331	1,964	0,621	99,963	99,971	0,387
	$*(CV(\hat{\theta}_a))^{-1}$	99,970	0,024	1,966	0,622	99,967	99,974	0,387
	$(skewness(s_{dc}))^{-1}$	99,955	-1,514	2,134	0,675	99,951	99,959	0,478
	$*(skewness(s_{dc}))^{-1}$	99,955	-1,483	2,146	0,679	99,951	99,960	0,483
	$(kurtosis(K_{dc}))^{-1}$	99,956	-1,428	2,182	0,690	99,952	99,960	0,497
	$*(kurtosis(K_{dc}))^{-1}$	99,957	-1,373	2,180	0,690	99,952	99,961	0,494
	$meff/Var(y)$	99,968	-0,253	1,959	0,620	99,964	99,972	0,384
	$(meff)^{-1}$	99,966	-0,411	1,981	0,627	99,962	99,970	0,394
5	1/n	99,971	0,033	1,543	0,488	99,968	99,974	0,238
	$(Var(\hat{\theta}_a))^{-1}$	99,956	-1,448	1,528	0,484	99,953	99,959	0,255
	$*(Var(\hat{\theta}_a))^{-1}$	99,964	-0,639	1,508	0,477	99,961	99,967	0,231
	$(CV(\hat{\theta}_a))^{-1}$	99,963	-0,745	1,517	0,480	99,960	99,966	0,236
	$*(CV(\hat{\theta}_a))^{-1}$	99,967	-0,307	1,518	0,480	99,964	99,970	0,231
	$(skewness(s_{dc}))^{-1}$	99,948	-2,242	1,657	0,524	99,945	99,951	0,325
	$*(skewness(s_{dc}))^{-1}$	99,948	-2,225	1,671	0,529	99,945	99,951	0,329
	$(kurtosis(K_{dc}))^{-1}$	99,949	-2,082	1,665	0,527	99,946	99,953	0,320
	$*(kurtosis(K_{dc}))^{-1}$	99,950	-2,029	1,662	0,526	99,947	99,953	0,317

	$meff/Var(y)$	99,964	-0,639	1,508	0,477	99,961	99,967	0,231
	$(meff)^{-1}$	99,962	-0,848	1,536	0,486	99,959	99,965	0,243
10	$1/n$	99,971	0,075	1,077	0,341	99,969	99,973	0,116
	$(Var(\hat{\theta}_a))^{-1}$	99,955	-1,514	1,072	0,339	99,953	99,957	0,138
	$*(Var(\hat{\theta}_a))^{-1}$	99,964	-0,643	1,062	0,336	99,962	99,966	0,117
	$(CV(\hat{\theta}_a))^{-1}$	99,963	-0,753	1,062	0,336	99,961	99,965	0,118
	$*(CV(\hat{\theta}_a))^{-1}$	99,967	-0,288	1,065	0,337	99,965	99,969	0,114
	$(skewness(s_{dc}))^{-1}$	99,946	-2,438	1,190	0,377	99,944	99,948	0,201
	$*(skewness(s_{dc}))^{-1}$	99,946	-2,415	1,207	0,382	99,944	99,948	0,204
	$(kurtosis(K_{dc}))^{-1}$	99,948	-2,221	1,189	0,376	99,946	99,950	0,191
	$*(kurtosis(K_{dc}))^{-1}$	99,949	-2,157	1,190	0,377	99,946	99,951	0,188
	$meff/Var(y)$	99,964	-0,643	1,062	0,336	99,962	99,966	0,117
	$(meff)^{-1}$	99,962	-0,860	1,071	0,339	99,960	99,964	0,122
15	$1/n$	99,971	0,109	0,879	0,278	99,970	99,973	0,077
	$(Var(\hat{\theta}_a))^{-1}$	99,955	-1,522	0,876	0,277	99,953	99,957	0,100
	$*(Var(\hat{\theta}_a))^{-1}$	99,964	-0,617	0,866	0,274	99,962	99,966	0,079
	$(CV(\hat{\theta}_a))^{-1}$	99,963	-0,740	0,868	0,274	99,961	99,965	0,081
	$*(CV(\hat{\theta}_a))^{-1}$	99,968	-0,258	0,869	0,275	99,966	99,969	0,076
	$(skewness(s_{dc}))^{-1}$	99,946	-2,455	0,980	0,310	99,944	99,948	0,156
	$*(skewness(s_{dc}))^{-1}$	99,946	-2,432	0,998	0,316	99,944	99,948	0,159
	$(kurtosis(K_{dc}))^{-1}$	99,948	-2,210	0,971	0,307	99,946	99,950	0,143
	$*(kurtosis(K_{dc}))^{-1}$	99,949	-2,142	0,972	0,308	99,947	99,951	0,140
	$meff/Var(y)$	99,964	-0,610	0,868	0,275	99,962	99,966	0,079
	$(meff)^{-1}$	99,962	-0,858	0,876	0,277	99,960	99,963	0,084
20	$1/n$	99,970	0,025	0,773	0,245	99,969	99,972	0,060
	$(Var(\hat{\theta}_a))^{-1}$	99,954	-1,637	0,767	0,243	99,952	99,955	0,086
	$*(Var(\hat{\theta}_a))^{-1}$	99,963	-0,710	0,764	0,242	99,962	99,965	0,063
	$(CV(\hat{\theta}_a))^{-1}$	99,962	-0,840	0,761	0,241	99,960	99,963	0,065
	$*(CV(\hat{\theta}_a))^{-1}$	99,967	-0,347	0,765	0,242	99,965	99,968	0,060
	$(skewness(s_{dc}))^{-1}$	99,944	-2,607	0,864	0,273	99,942	99,946	0,142
	$*(skewness(s_{dc}))^{-1}$	99,944	-2,580	0,882	0,279	99,943	99,946	0,144
	$(kurtosis(K_{dc}))^{-1}$	99,947	-2,346	0,854	0,270	99,945	99,948	0,128
	$*(kurtosis(K_{dc}))^{-1}$	99,948	-2,274	0,856	0,271	99,946	99,949	0,125
	$meff/Var(y)$	99,963	-0,702	0,763	0,241	99,962	99,965	0,063
	$(meff)^{-1}$	99,961	-0,964	0,764	0,242	99,959	99,962	0,068

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 2: N=1000 000 at(3)(100,25) n=5000 D=20 $\theta = 99.970$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		99,971	0,607	1,533	4,848	99,968	99,974	2,350
2	1/n	99,972	1,321	1,079	3,413	99,969	99,974	1,166
	$(Var(\hat{\theta}_a))^{-1}$	99,969	-0,763	1,082	3,423	99,967	99,972	1,171
	$*(Var(\hat{\theta}_a))^{-1}$	99,971	0,427	1,081	3,418	99,969	99,973	1,168
	$(CV(\hat{\theta}_a))^{-1}$	99,971	0,295	1,082	3,422	99,968	99,973	1,170
	$*(CV(\hat{\theta}_a))^{-1}$	99,971	0,899	1,083	3,425	99,969	99,973	1,173
	$(skewness(s_{dc}))^{-1}$	99,967	-2,793	1,130	3,574	99,965	99,970	1,285
	$*(skewness(s_{dc}))^{-1}$	99,968	-2,669	1,134	3,588	99,965	99,970	1,294
	$(kurtosis(K_{dc}))^{-1}$	99,967	-3,599	1,186	3,751	99,964	99,969	1,419
	$*(kurtosis(K_{dc}))^{-1}$	99,967	-3,387	1,188	3,757	99,965	99,969	1,422
	$meff/Var(y)$	99,971	0,429	1,081	3,418	99,969	99,973	1,168
	$(meff)^{-1}$	99,970	0,143	1,084	3,430	99,968	99,973	1,176
3	1/n	99,971	0,276	0,859	2,716	99,969	99,972	0,738
	$(Var(\hat{\theta}_a))^{-1}$	99,968	-2,510	0,864	2,733	99,966	99,969	0,753
	$*(Var(\hat{\theta}_a))^{-1}$	99,969	-0,850	0,860	2,722	99,968	99,971	0,741
	$(CV(\hat{\theta}_a))^{-1}$	99,969	-1,090	0,862	2,727	99,967	99,971	0,744
	$*(CV(\hat{\theta}_a))^{-1}$	99,970	-0,242	0,862	2,726	99,968	99,972	0,743
	$(skewness(s_{dc}))^{-1}$	99,965	-5,156	0,917	2,902	99,963	99,967	0,868
	$*(skewness(s_{dc}))^{-1}$	99,965	-4,988	0,921	2,915	99,963	99,967	0,874
	$(kurtosis(K_{dc}))^{-1}$	99,964	-6,340	0,976	3,088	99,962	99,966	0,993
	$*(kurtosis(K_{dc}))^{-1}$	99,964	-6,064	0,978	3,095	99,962	99,966	0,994
	$meff/Var(y)$	99,969	-0,846	0,860	2,722	99,968	99,971	0,741
	$(meff)^{-1}$	99,969	-1,348	0,865	2,736	99,967	99,971	0,750
5	1/n	99,971	0,298	0,676	2,137	99,969	99,972	0,456
	$(Var(\hat{\theta}_a))^{-1}$	99,967	-3,077	0,674	2,133	99,966	99,968	0,464
	$*(Var(\hat{\theta}_a))^{-1}$	99,969	-1,091	0,672	2,125	99,968	99,970	0,453
	$(CV(\hat{\theta}_a))^{-1}$	99,969	-1,404	0,673	2,128	99,968	99,970	0,454
	$*(CV(\hat{\theta}_a))^{-1}$	99,970	-0,385	0,673	2,129	99,969	99,971	0,453
	$(skewness(s_{dc}))^{-1}$	99,964	-6,190	0,716	2,266	99,963	99,965	0,552
	$*(skewness(s_{dc}))^{-1}$	99,964	-6,038	0,718	2,273	99,963	99,966	0,553
	$(kurtosis(K_{dc}))^{-1}$	99,963	-7,573	0,763	2,413	99,961	99,964	0,639
	$*(kurtosis(K_{dc}))^{-1}$	99,963	-7,329	0,762	2,412	99,961	99,964	0,635
	$meff/Var(y)$	99,969	-1,040	0,674	2,133	99,968	99,971	0,456
	$(meff)^{-1}$	99,969	-1,685	0,677	2,140	99,967	99,970	0,461
10	1/n	99,971	0,582	0,488	1,542	99,970	99,972	0,238
	$(Var(\hat{\theta}_a))^{-1}$	99,967	-3,225	0,480	1,520	99,966	99,968	0,241
	$*(Var(\hat{\theta}_a))^{-1}$	99,969	-1,008	0,483	1,528	99,968	99,970	0,234
	$(CV(\hat{\theta}_a))^{-1}$	99,969	-1,342	0,482	1,525	99,968	99,970	0,234

	$*(CV(\hat{\theta}_a))^{-1}$	99,970	-0,201	0,485	1,534	99,969	99,971	0,235
	$(skewness(s_{dc}))^{-1}$	99,963	-6,963	0,511	1,615	99,962	99,964	0,309
	$*(skewness(s_{dc}))^{-1}$	99,963	-6,841	0,513	1,623	99,962	99,964	0,310
	$(kurtosis(K_{dc}))^{-1}$	99,962	-8,589	0,544	1,722	99,961	99,963	0,370
	$*(kurtosis(K_{dc}))^{-1}$	99,962	-8,342	0,544	1,721	99,961	99,963	0,366
	$mef/Var(y)$	99,969	-1,004	0,483	1,528	99,968	99,970	0,234
	$(mef)^{-1}$	99,969	-1,696	0,483	1,527	99,968	99,969	0,236
15	$1/n$	99,971	0,396	0,406	1,284	99,970	99,971	0,165
	$(Var(\hat{\theta}_a))^{-1}$	99,967	-3,550	0,402	1,271	99,966	99,967	0,174
	$*(Var(\hat{\theta}_a))^{-1}$	99,969	-1,272	0,403	1,273	99,968	99,970	0,164
	$(CV(\hat{\theta}_a))^{-1}$	99,969	-1,602	0,402	1,273	99,968	99,969	0,164
	$*(CV(\hat{\theta}_a))^{-1}$	99,970	-0,427	0,404	1,278	99,969	99,971	0,163
	$(skewness(s_{dc}))^{-1}$	99,963	-7,491	0,427	1,352	99,962	99,964	0,239
	$*(skewness(s_{dc}))^{-1}$	99,963	-7,384	0,430	1,360	99,962	99,964	0,239
	$(kurtosis(K_{dc}))^{-1}$	99,961	-9,198	0,456	1,443	99,960	99,962	0,293
	$*(kurtosis(K_{dc}))^{-1}$	99,961	-8,950	0,457	1,444	99,960	99,962	0,288
	$mef/Var(y)$	99,969	-1,316	0,403	1,276	99,968	99,970	0,164
	$(mef)^{-1}$	99,968	-1,970	0,403	1,275	99,967	99,969	0,166
20	$1/n$	99,970	0,236	0,342	1,083	99,970	99,971	0,117
	$(Var(\hat{\theta}_a))^{-1}$	99,966	-3,774	0,338	1,070	99,966	99,967	0,129
	$*(Var(\hat{\theta}_a))^{-1}$	99,969	-1,443	0,339	1,073	99,968	99,969	0,117
	$(CV(\hat{\theta}_a))^{-1}$	99,968	-1,794	0,339	1,072	99,968	99,969	0,118
	$*(CV(\hat{\theta}_a))^{-1}$	99,970	-0,592	0,341	1,077	99,969	99,970	0,116
	$(skewness(s_{dc}))^{-1}$	99,963	-7,716	0,361	1,141	99,962	99,963	0,190
	$*(skewness(s_{dc}))^{-1}$	99,963	-7,587	0,362	1,146	99,962	99,963	0,189
	$(kurtosis(K_{dc}))^{-1}$	99,961	-9,421	0,387	1,223	99,960	99,962	0,238
	$*(kurtosis(K_{dc}))^{-1}$	99,961	-9,153	0,386	1,220	99,960	99,962	0,232
	$mef/Var(y)$	99,969	-1,411	0,340	1,076	99,968	99,970	0,118
	$(mef)^{-1}$	99,968	-2,152	0,340	1,074	99,967	99,969	0,120

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 3: N=1000 000 at(3)(100,25) n=10000 D=20 $\theta = 99.970$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-3}$	SE($\hat{\theta}_c$) $\times 10^{-3}$	CV($\hat{\theta}_c$) (%) $\times 10^{-2}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-3}$
						LB	UB	
1		99,971	1,173	1,083	3,425	99,969	99,974	1,174
2	1/n	99,971	0,885	0,757	2,395	99,970	99,973	0,574
	$(Var(\hat{\theta}_a))^{-1}$	99,970	-0,312	0,763	2,413	99,968	99,971	0,582
	$*(Var(\hat{\theta}_a))^{-1}$	99,971	0,319	0,762	2,411	99,969	99,972	0,581
	$(CV(\hat{\theta}_a))^{-1}$	99,970	0,229	0,761	2,409	99,969	99,972	0,580
	$*(CV(\hat{\theta}_a))^{-1}$	99,971	0,548	0,762	2,410	99,969	99,972	0,581
	$(skewness(s_{dc}))^{-1}$	99,969	-1,508	0,796	2,516	99,967	99,970	0,635
	$*(skewness(s_{dc}))^{-1}$	99,969	-1,445	0,799	2,526	99,967	99,970	0,640
	$(kurtosis(K_{dc}))^{-1}$	99,968	-2,202	0,835	2,643	99,966	99,970	0,703
	$*(kurtosis(K_{dc}))^{-1}$	99,968	-2,087	0,837	2,648	99,967	99,970	0,705
	$meff/Var(y)$	99,971	0,320	0,762	2,411	99,969	99,972	0,581
	$(meff)^{-1}$	99,970	0,129	0,761	2,408	99,969	99,972	0,579
3	1/n	99,971	0,761	0,623	1,969	99,970	99,972	0,388
	$(Var(\hat{\theta}_a))^{-1}$	99,969	-0,833	0,625	1,978	99,968	99,971	0,392
	$*(Var(\hat{\theta}_a))^{-1}$	99,970	0,018	0,627	1,982	99,969	99,971	0,393
	$(CV(\hat{\theta}_a))^{-1}$	99,970	-0,092	0,625	1,977	99,969	99,971	0,391
	$*(CV(\hat{\theta}_a))^{-1}$	99,971	0,339	0,626	1,982	99,969	99,972	0,393
	$(skewness(s_{dc}))^{-1}$	99,968	-2,501	0,667	2,109	99,966	99,969	0,451
	$*(skewness(s_{dc}))^{-1}$	99,968	-2,455	0,673	2,127	99,966	99,969	0,458
	$(kurtosis(K_{dc}))^{-1}$	99,967	-3,406	0,722	2,284	99,965	99,968	0,533
	$*(kurtosis(K_{dc}))^{-1}$	99,967	-3,306	0,726	2,297	99,966	99,968	0,538
	$meff/Var(y)$	99,970	0,011	0,627	1,982	99,969	99,971	0,393
	$(meff)^{-1}$	99,970	-0,215	0,624	1,974	99,969	99,971	0,390
5	1/n	99,970	-0,027	0,494	1,564	99,969	99,971	0,245
	$(Var(\hat{\theta}_a))^{-1}$	99,969	-1,744	0,493	1,561	99,968	99,969	0,247
	$*(Var(\hat{\theta}_a))^{-1}$	99,969	-0,746	0,493	1,560	99,969	99,970	0,244
	$(CV(\hat{\theta}_a))^{-1}$	99,969	-0,885	0,493	1,560	99,968	99,970	0,244
	$*(CV(\hat{\theta}_a))^{-1}$	99,970	-0,379	0,494	1,561	99,969	99,971	0,244
	$(skewness(s_{dc}))^{-1}$	99,967	-3,717	0,521	1,648	99,966	99,968	0,285
	$*(skewness(s_{dc}))^{-1}$	99,967	-3,659	0,524	1,657	99,966	99,968	0,288
	$(kurtosis(K_{dc}))^{-1}$	99,965	-4,806	0,565	1,788	99,964	99,967	0,343
	$*(kurtosis(K_{dc}))^{-1}$	99,966	-4,699	0,568	1,796	99,964	99,967	0,344
	$meff/Var(y)$	99,969	-0,775	0,492	1,555	99,969	99,970	0,242
	$(meff)^{-1}$	99,969	-1,034	0,494	1,562	99,968	99,970	0,245
10	1/n	99,971	0,439	0,353	1,116	99,970	99,971	0,125
	$(Var(\hat{\theta}_a))^{-1}$	99,969	-1,495	0,351	1,111	99,968	99,969	0,126
	$*(Var(\hat{\theta}_a))^{-1}$	99,970	-0,361	0,352	1,112	99,969	99,971	0,124
	$(CV(\hat{\theta}_a))^{-1}$	99,970	-0,530	0,351	1,111	99,969	99,970	0,124
	$*(CV(\hat{\theta}_a))^{-1}$	99,970	0,047	0,352	1,114	99,970	99,971	0,124
	$(skewness(s_{dc}))^{-1}$	99,967	-3,705	0,371	1,175	99,966	99,967	0,152

	$*(skewness(s_{dc}))^{-1}$	99,967	-3,621	0,373	1,181	99,966	99,967	0,152
	$(kurtosis(K_{dc}))^{-1}$	99,965	-5,003	0,404	1,278	99,964	99,966	0,188
	$*(kurtosis(K_{dc}))^{-1}$	99,965	-4,850	0,405	1,281	99,965	99,966	0,188
	$meff/Var(y)$	99,970	-0,354	0,352	1,113	99,969	99,971	0,124
	$(meff)^{-1}$	99,970	-0,724	0,351	1,111	99,969	99,970	0,124
15	$1/n$	99,971	0,412	0,277	0,876	99,970	99,971	0,077
	$(Var(\hat{\theta}_a))^{-1}$	99,969	-1,603	0,275	0,868	99,968	99,969	0,078
	$*(Var(\hat{\theta}_a))^{-1}$	99,970	-0,407	0,275	0,871	99,969	99,970	0,076
	$(CV(\hat{\theta}_a))^{-1}$	99,970	-0,599	0,275	0,870	99,969	99,970	0,076
	$*(CV(\hat{\theta}_a))^{-1}$	99,970	0,010	0,276	0,873	99,970	99,971	0,076
	$(skewness(s_{dc}))^{-1}$	99,966	-3,858	0,293	0,927	99,966	99,967	0,101
	$*(skewness(s_{dc}))^{-1}$	99,967	-3,746	0,295	0,932	99,966	99,967	0,101
	$(kurtosis(K_{dc}))^{-1}$	99,965	-5,191	0,323	1,021	99,964	99,966	0,131
	$*(kurtosis(K_{dc}))^{-1}$	99,965	-5,016	0,323	1,022	99,965	99,966	0,129
	$meff/Var(y)$	99,970	-0,385	0,278	0,878	99,969	99,970	0,077
	$(meff)^{-1}$	99,969	-0,810	0,275	0,869	99,969	99,970	0,076
20	$1/n$	99,971	0,467	0,242	0,767	99,970	99,971	0,059
	$(Var(\hat{\theta}_a))^{-1}$	99,969	-1,598	0,241	0,761	99,968	99,969	0,060
	$*(Var(\hat{\theta}_a))^{-1}$	99,970	-0,377	0,241	0,763	99,969	99,970	0,058
	$(CV(\hat{\theta}_a))^{-1}$	99,970	-0,569	0,241	0,762	99,969	99,970	0,058
	$*(CV(\hat{\theta}_a))^{-1}$	99,970	0,053	0,242	0,765	99,970	99,971	0,058
	$(skewness(s_{dc}))^{-1}$	99,966	-3,948	0,257	0,814	99,966	99,967	0,082
	$*(skewness(s_{dc}))^{-1}$	99,966	-3,829	0,259	0,818	99,966	99,967	0,082
	$(kurtosis(K_{dc}))^{-1}$	99,965	-5,336	0,283	0,894	99,964	99,965	0,108
	$*(kurtosis(K_{dc}))^{-1}$	99,965	-5,141	0,283	0,894	99,965	99,966	0,106
	$meff/Var(y)$	99,970	-0,390	0,242	0,766	99,969	99,970	0,059
	$(meff)^{-1}$	99,969	-0,772	0,241	0,762	99,969	99,970	0,059

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, $CI_{95\%}(\hat{\theta}_c)$ = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Appendix D2

Comparing weighting strategies: Cluster Random Sampling: Skewed t-distribution, 3 d.f

Scenario 1: N=1 000 000 at(3)(100,25) n=1000 D=20, $\theta = 99.940$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	$SE(\hat{\theta}_c)$ $\times 10^{-2}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-1}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-2}$
						LB	UB	
1		99,936	-0,595	2,697	2,698	99,883	99,989	7,276
2	1/n	99,918	-2,374	1,835	1,836	99,882	99,954	3,423
	$(Var(\hat{\theta}_d))^{-1}$	99,899	-4,257	1,978	1,980	99,861	99,938	4,095
	$*(Var(\hat{\theta}_d))^{-1}$	99,916	-2,623	1,973	1,974	99,877	99,954	3,960
	$(CV(\hat{\theta}_d))^{-1}$	99,921	-2,088	1,947	1,948	99,883	99,959	3,834
	$*(CV(\hat{\theta}_d))^{-1}$	99,930	-1,225	1,958	1,959	99,891	99,968	3,848
	$(skewness(s_{dc}))^{-1}$	99,924	-1,815	2,010	2,012	99,884	99,963	4,073
	$*(skewness(s_{dc}))^{-1}$	99,924	-1,815	2,010	2,012	99,884	99,963	4,073
	$(kurtosis(K_{dc}))^{-1}$	99,919	-2,329	2,080	2,082	99,878	99,959	4,382
	$*(kurtosis(K_{dc}))^{-1}$	99,919	-2,329	2,080	2,082	99,878	99,959	4,382
	$meff/Var(y)$	99,916	-2,627	1,973	1,974	99,877	99,954	3,961
	$(meff)^{-1}$	99,927	-1,510	1,990	1,992	99,888	99,966	3,983
3	1/n	99,927	-1,529	1,523	1,524	99,897	99,956	2,343
	$(Var(\hat{\theta}_d))^{-1}$	99,888	-5,384	1,567	1,568	99,857	99,919	2,743
	$*(Var(\hat{\theta}_d))^{-1}$	99,913	-2,865	1,593	1,594	99,882	99,944	2,620
	$(CV(\hat{\theta}_d))^{-1}$	99,918	-2,419	1,571	1,572	99,887	99,948	2,525
	$*(CV(\hat{\theta}_d))^{-1}$	99,931	-1,084	1,604	1,605	99,900	99,962	2,586
	$(skewness(s_{dc}))^{-1}$	99,921	-2,106	1,692	1,694	99,888	99,954	2,909
	$*(skewness(s_{dc}))^{-1}$	99,921	-2,106	1,692	1,694	99,888	99,954	2,909
	$(kurtosis(K_{dc}))^{-1}$	99,913	-2,904	1,809	1,811	99,877	99,948	3,357
	$*(kurtosis(K_{dc}))^{-1}$	99,913	-2,904	1,809	1,811	99,877	99,948	3,357
	$meff/Var(y)$	99,914	-2,743	1,595	1,596	99,883	99,946	2,618
	$(meff)^{-1}$	99,924	-1,804	1,631	1,632	99,892	99,956	2,693
5	1/n	99,935	-0,688	1,316	1,317	99,909	99,961	1,736
	$(Var(\hat{\theta}_d))^{-1}$	99,869	-7,298	1,295	1,297	99,844	99,894	2,208
	$*(Var(\hat{\theta}_d))^{-1}$	99,898	-4,378	1,273	1,274	99,873	99,923	1,811
	$(CV(\hat{\theta}_d))^{-1}$	99,901	-4,059	1,264	1,265	99,877	99,926	1,762
	$*(CV(\hat{\theta}_d))^{-1}$	99,917	-2,520	1,280	1,282	99,892	99,942	1,703
	$(skewness(s_{dc}))^{-1}$	99,908	-3,339	1,339	1,340	99,882	99,935	1,903
	$*(skewness(s_{dc}))^{-1}$	99,908	-3,339	1,339	1,340	99,882	99,935	1,903
	$(kurtosis(K_{dc}))^{-1}$	99,904	-3,781	1,400	1,402	99,877	99,932	2,104
	$*(kurtosis(K_{dc}))^{-1}$	99,904	-3,781	1,400	1,402	99,877	99,932	2,104
	$meff/Var(y)$	99,901	-4,038	1,313	1,314	99,876	99,927	1,886
	$(meff)^{-1}$	99,908	-3,404	1,329	1,330	99,882	99,934	1,881
10	1/n	99,927	-1,439	0,922	0,922	99,909	99,946	0,870

	$(Var(\hat{\theta}_a))^{-1}$	99,862	-7,945	0,989	0,991	99,843	99,882	1,608
	$*(Var(\hat{\theta}_a))^{-1}$	99,894	-4,830	0,924	0,925	99,875	99,912	1,086
	$(CV(\hat{\theta}_a))^{-1}$	99,895	-4,727	0,920	0,921	99,877	99,913	1,070
	$*(CV(\hat{\theta}_a))^{-1}$	99,911	-3,108	0,915	0,915	99,893	99,929	0,933
	$(skewness(s_{dc}))^{-1}$	99,908	-3,339	0,982	0,983	99,889	99,928	1,076
	$*(skewness(s_{dc}))^{-1}$	99,908	-3,339	0,982	0,983	99,889	99,928	1,076
	$(kurtosis(K_{dc}))^{-1}$	99,908	-3,395	1,062	1,063	99,887	99,929	1,244
	$*(kurtosis(K_{dc}))^{-1}$	99,908	-3,395	1,062	1,063	99,887	99,929	1,244
	$meff/Var(y)$	99,894	-4,816	0,925	0,926	99,876	99,912	1,086
	$(meff)^{-1}$	99,896	-4,559	0,965	0,966	99,877	99,915	1,138
15	$1/n$	99,932	-1,023	0,781	0,782	99,916	99,947	0,621
	$(Var(\hat{\theta}_a))^{-1}$	99,865	-7,727	0,828	0,829	99,848	99,881	1,281
	$*(Var(\hat{\theta}_a))^{-1}$	99,897	-4,522	0,774	0,775	99,881	99,912	0,803
	$(CV(\hat{\theta}_a))^{-1}$	99,898	-4,416	0,777	0,778	99,882	99,913	0,799
	$*(CV(\hat{\theta}_a))^{-1}$	99,915	-2,734	0,770	0,771	99,899	99,930	0,668
	$(skewness(s_{dc}))^{-1}$	99,910	-3,208	0,787	0,788	99,894	99,925	0,722
	$*(skewness(s_{dc}))^{-1}$	99,910	-3,208	0,787	0,788	99,894	99,925	0,722
	$(kurtosis(K_{dc}))^{-1}$	99,910	-3,229	0,842	0,843	99,893	99,926	0,813
	$*(kurtosis(K_{dc}))^{-1}$	99,910	-3,229	0,842	0,843	99,893	99,926	0,813
	$meff/Var(y)$	99,899	-4,305	0,801	0,802	99,883	99,915	0,827
	$(meff)^{-1}$	99,901	-4,098	0,817	0,818	99,885	99,917	0,835
20	$1/n$	99,933	-0,910	0,681	0,682	99,919	99,946	0,473
	$(Var(\hat{\theta}_a))^{-1}$	99,867	-7,459	0,684	0,685	99,854	99,881	1,023
	$*(Var(\hat{\theta}_a))^{-1}$	99,898	-4,338	0,664	0,665	99,885	99,912	0,629
	$(CV(\hat{\theta}_a))^{-1}$	99,899	-4,249	0,664	0,665	99,886	99,912	0,621
	$*(CV(\hat{\theta}_a))^{-1}$	99,916	-2,596	0,666	0,666	99,903	99,929	0,511
	$(skewness(s_{dc}))^{-1}$	99,912	-3,019	0,694	0,695	99,898	99,925	0,573
	$*(skewness(s_{dc}))^{-1}$	99,912	-3,019	0,694	0,695	99,898	99,925	0,573
	$(kurtosis(K_{dc}))^{-1}$	99,912	-2,974	0,748	0,748	99,897	99,927	0,648
	$*(kurtosis(K_{dc}))^{-1}$	99,912	-2,974	0,748	0,748	99,897	99,927	0,648
	$meff/Var(y)$	99,899	-4,329	0,666	0,666	99,886	99,912	0,630
	$(meff)^{-1}$	99,905	-3,710	0,762	0,762	99,890	99,920	0,717

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 2: N=1000 000 at(3)(100,25) n=5000 D=20 $\theta = 99.940$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	$SE(\hat{\theta}_c)$ $\times 10^{-2}$	$CV(\hat{\theta}_c)$ (%) $\times 10^{-1}$	$CI_{95\%}(\hat{\theta}_c)$		MSE $\times 10^{-2}$
						LB	UB	
1		99,9219	-1,777	1,494	1,495	99,89263	99,95118	2,262
2	1/n	99,93113	-0,853	0,992	0,992	99,91117	99,95057	0,991
	$(Var(\hat{\theta}_a))^{-1}$	99,91804	-2,164	1,041	1,042	99,89763	99,93845	1,131
	$*(Var(\hat{\theta}_a))^{-1}$	99,92681	-1,286	1,037	1,038	99,90647	99,94714	1,093
	$(CV(\hat{\theta}_a))^{-1}$	99,92706	-1,261	1,027	1,028	99,90692	99,94719	1,071
	$*(CV(\hat{\theta}_a))^{-1}$	99,93211	-0,755	1,038	1,038	99,91178	99,95245	1,083
	$(skewness(s_{dc}))^{-1}$	99,92618	-1,349	1,069	1,070	99,90522	99,94714	1,162
	$*(skewness(s_{dc}))^{-1}$	99,92618	-1,349	1,069	1,070	99,90522	99,94714	1,162
	$(kurtosis(K_{dc}))^{-1}$	99,92294	-1,673	1,138	1,139	99,90064	99,94525	1,323
	$*(kurtosis(K_{dc}))^{-1}$	99,92294	-1,673	1,138	1,139	99,90064	99,94525	1,323
	$meff/Var(y)$	99,92732	-1,235	1,032	1,033	99,90708	99,94755	1,081
	$(meff)^{-1}$	99,92719	-1,248	1,028	1,028	99,90705	99,94733	1,072
3	1/n	99,9304	-0,926	0,824	0,825	99,91425	99,94656	0,688
	$(Var(\hat{\theta}_a))^{-1}$	99,92427	-1,540	0,824	0,824	99,90813	99,94042	0,702
	$*(Var(\hat{\theta}_a))^{-1}$	99,93219	-0,747	0,838	0,838	99,91577	99,94861	0,707
	$(CV(\hat{\theta}_a))^{-1}$	99,93212	-0,755	0,827	0,827	99,91591	99,94832	0,689
	$*(CV(\hat{\theta}_a))^{-1}$	99,93675	-0,291	0,842	0,842	99,92025	99,95325	0,710
	$(skewness(s_{dc}))^{-1}$	99,93281	-0,686	0,886	0,886	99,91545	99,95017	0,789
	$*(skewness(s_{dc}))^{-1}$	99,93281	-0,686	0,886	0,886	99,91545	99,95017	0,789
	$(kurtosis(K_{dc}))^{-1}$	99,9309	-0,877	0,984	0,985	99,91161	99,95019	0,977
	$*(kurtosis(K_{dc}))^{-1}$	99,9309	-0,877	0,984	0,985	99,91161	99,95019	0,977
	$meff/Var(y)$	99,9321	-0,756	0,837	0,838	99,91569	99,94851	0,707
	$(meff)^{-1}$	99,93265	-0,701	0,831	0,831	99,91637	99,94894	0,695
5	1/n	99,93075	-0,892	0,716	0,717	99,91671	99,94479	0,521
	$(Var(\hat{\theta}_a))^{-1}$	99,90996	-2,972	0,648	0,649	99,89726	99,92267	0,508
	$*(Var(\hat{\theta}_a))^{-1}$	99,92045	-1,923	0,695	0,696	99,90682	99,93407	0,520
	$(CV(\hat{\theta}_a))^{-1}$	99,91957	-2,011	0,671	0,672	99,90642	99,93272	0,491
	$*(CV(\hat{\theta}_a))^{-1}$	99,9256	-1,407	0,703	0,704	99,91182	99,93938	0,514
	$(skewness(s_{dc}))^{-1}$	99,92138	-1,830	0,754	0,755	99,90659	99,93617	0,603
	$*(skewness(s_{dc}))^{-1}$	99,92138	-1,830	0,754	0,755	99,90659	99,93617	0,603
	$(kurtosis(K_{dc}))^{-1}$	99,91838	-2,130	0,859	0,860	99,90154	99,93521	0,783
	$*(kurtosis(K_{dc}))^{-1}$	99,91838	-2,130	0,859	0,860	99,90154	99,93521	0,783
	$meff/Var(y)$	99,92011	-1,957	0,700	0,700	99,90639	99,93382	0,528
	$(meff)^{-1}$	99,91847	-2,121	0,658	0,658	99,90558	99,93136	0,477
10	1/n	99,93584	-0,382	0,474	0,474	99,92655	99,94514	0,226
	$(Var(\hat{\theta}_a))^{-1}$	99,90998	-2,971	0,448	0,448	99,9012	99,91876	0,289
	$*(Var(\hat{\theta}_a))^{-1}$	99,92362	-1,605	0,460	0,460	99,9146	99,93264	0,237
	$(CV(\hat{\theta}_a))^{-1}$	99,92149	-1,819	0,451	0,451	99,91264	99,93033	0,237

	$*(CV(\hat{\theta}_a))^{-1}$	99,92965	-1,002	0,464	0,465	99,92055	99,93875	0,226
	$(skewness(s_{dc}))^{-1}$	99,92695	-1,272	0,484	0,485	99,91745	99,93644	0,251
	$*(skewness(s_{dc}))^{-1}$	99,92695	-1,272	0,484	0,485	99,91745	99,93644	0,251
	$(kurtosis(K_{dc}))^{-1}$	99,92514	-1,453	0,542	0,542	99,91453	99,93576	0,314
	$*(kurtosis(K_{dc}))^{-1}$	99,92514	-1,453	0,542	0,542	99,91453	99,93576	0,314
	$mef/Var(y)$	99,92274	-1,694	0,444	0,444	99,91404	99,93143	0,225
	$(mef)^{-1}$	99,91812	-2,156	0,439	0,440	99,90951	99,92673	0,239
15	$1/n$	99,93693	-0,274	0,373	0,374	99,92961	99,94424	0,140
	$(Var(\hat{\theta}_a))^{-1}$	99,91126	-2,843	0,348	0,348	99,90444	99,91807	0,201
	$*(Var(\hat{\theta}_a))^{-1}$	99,92459	-1,508	0,361	0,361	99,91751	99,93167	0,153
	$(CV(\hat{\theta}_a))^{-1}$	99,92269	-1,699	0,350	0,351	99,91582	99,92955	0,151
	$*(CV(\hat{\theta}_a))^{-1}$	99,93065	-0,902	0,364	0,365	99,92351	99,93779	0,141
	$(skewness(s_{dc}))^{-1}$	99,92792	-1,175	0,394	0,394	99,92021	99,93563	0,169
	$*(skewness(s_{dc}))^{-1}$	99,92792	-1,175	0,394	0,394	99,92021	99,93563	0,169
	$(kurtosis(K_{dc}))^{-1}$	99,92646	-1,321	0,452	0,453	99,91759	99,93532	0,222
	$*(kurtosis(K_{dc}))^{-1}$	99,92646	-1,321	0,452	0,453	99,91759	99,93532	0,222
	$mef/Var(y)$	99,9262	-1,348	0,392	0,393	99,91851	99,93388	0,172
	$(mef)^{-1}$	99,92175	-1,793	0,355	0,355	99,91479	99,9287	0,158
20	$1/n$	99,93611	-0,355	0,303	0,304	99,93017	99,94206	0,093
	$(Var(\hat{\theta}_a))^{-1}$	99,90937	-3,031	0,305	0,306	99,90339	99,91536	0,185
	$*(Var(\hat{\theta}_a))^{-1}$	99,92302	-1,665	0,308	0,308	99,91699	99,92906	0,122
	$(CV(\hat{\theta}_a))^{-1}$	99,92113	-1,855	0,299	0,299	99,91526	99,92699	0,124
	$*(CV(\hat{\theta}_a))^{-1}$	99,92943	-1,024	0,304	0,304	99,92347	99,93538	0,103
	$(skewness(s_{dc}))^{-1}$	99,92572	-1,395	0,345	0,345	99,91896	99,93247	0,138
	$*(skewness(s_{dc}))^{-1}$	99,92572	-1,395	0,345	0,345	99,91896	99,93247	0,138
	$(kurtosis(K_{dc}))^{-1}$	99,9234	-1,627	0,399	0,400	99,91557	99,93123	0,186
	$*(kurtosis(K_{dc}))^{-1}$	99,9234	-1,627	0,399	0,400	99,91557	99,93123	0,186
	$mef/Var(y)$	99,92325	-1,643	0,307	0,308	99,91722	99,92927	0,121
	$(mef)^{-1}$	99,91988	-1,980	0,297	0,298	99,91405	99,9257	0,128

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, CI_{95%}($\hat{\theta}_c$)= 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Scenario 3: N=1000 000 at(3)(100,25) n=10000 D=20 $\theta = 99.940$

D	Method	$E(\hat{\theta}_c)$	Rel. Bias (%) $\times 10^{-2}$	SE($\hat{\theta}_c$) $\times 10^{-2}$	CV($\hat{\theta}_c$) (%) $\times 10^{-1}$	CI _{95%} ($\hat{\theta}_c$)		MSE $\times 10^{-2}$
						LB	UB	
1		99,920	-1,985	1,228	1,229	99,896	99,944	1,547
2	1/n	99,934	-0,542	0,728	0,729	99,920	99,949	0,533
	$(Var(\hat{\theta}_a))^{-1}$	99,921	-1,823	0,755	0,756	99,907	99,936	0,604
	$*(Var(\hat{\theta}_a))^{-1}$	99,928	-1,202	0,755	0,755	99,913	99,942	0,584
	$(CV(\hat{\theta}_a))^{-1}$	99,926	-1,343	0,756	0,756	99,911	99,941	0,589
	$*(CV(\hat{\theta}_a))^{-1}$	99,930	-1,015	0,761	0,761	99,915	99,944	0,589
	$(skewness(s_{dc}))^{-1}$	99,930	-1,009	0,770	0,771	99,914	99,945	0,603
	$*(skewness(s_{dc}))^{-1}$	99,930	-1,009	0,770	0,771	99,914	99,945	0,603
	$(kurtosis(K_{dc}))^{-1}$	99,931	-0,908	0,814	0,815	99,915	99,947	0,671
	$*(kurtosis(K_{dc}))^{-1}$	99,931	-0,908	0,814	0,815	99,915	99,947	0,671
	$meff/Var(y)$	99,928	-1,200	0,755	0,755	99,913	99,942	0,584
	$(meff)^{-1}$	99,925	-1,488	0,761	0,761	99,910	99,940	0,601
3	1/n	99,937	-0,282	0,601	0,602	99,925	99,949	0,362
	$(Var(\hat{\theta}_a))^{-1}$	99,925	-1,490	0,602	0,603	99,913	99,937	0,385
	$*(Var(\hat{\theta}_a))^{-1}$	99,933	-0,694	0,610	0,611	99,921	99,945	0,377
	$(CV(\hat{\theta}_a))^{-1}$	99,931	-0,837	0,607	0,607	99,919	99,943	0,375
	$*(CV(\hat{\theta}_a))^{-1}$	99,936	-0,415	0,617	0,617	99,923	99,948	0,382
	$(skewness(s_{dc}))^{-1}$	99,934	-0,573	0,650	0,650	99,921	99,947	0,426
	$*(skewness(s_{dc}))^{-1}$	99,934	-0,573	0,650	0,650	99,921	99,947	0,426
	$(kurtosis(K_{dc}))^{-1}$	99,934	-0,610	0,725	0,726	99,919	99,948	0,530
	$*(kurtosis(K_{dc}))^{-1}$	99,934	-0,610	0,725	0,726	99,919	99,948	0,530
	$meff/Var(y)$	99,933	-0,693	0,610	0,611	99,921	99,945	0,377
	$(meff)^{-1}$	99,930	-0,984	0,611	0,611	99,918	99,942	0,383
5	1/n	99,940	0,035	0,496	0,497	99,930	99,950	0,246
	$(Var(\hat{\theta}_a))^{-1}$	99,925	-1,440	0,477	0,477	99,916	99,935	0,248
	$*(Var(\hat{\theta}_a))^{-1}$	99,934	-0,571	0,485	0,485	99,924	99,943	0,239
	$(CV(\hat{\theta}_a))^{-1}$	99,932	-0,734	0,481	0,481	99,923	99,942	0,236
	$*(CV(\hat{\theta}_a))^{-1}$	99,937	-0,262	0,490	0,490	99,927	99,947	0,241
	$(skewness(s_{dc}))^{-1}$	99,935	-0,477	0,510	0,510	99,925	99,945	0,263
	$*(skewness(s_{dc}))^{-1}$	99,935	-0,477	0,510	0,510	99,925	99,945	0,263
	$(kurtosis(K_{dc}))^{-1}$	99,933	-0,661	0,566	0,566	99,922	99,944	0,325
	$*(kurtosis(K_{dc}))^{-1}$	99,933	-0,661	0,566	0,566	99,922	99,944	0,325
	$meff/Var(y)$	99,934	-0,598	0,487	0,487	99,924	99,943	0,241
	$(meff)^{-1}$	99,930	-0,949	0,485	0,485	99,921	99,940	0,244
10	1/n	99,942	0,195	0,335	0,335	99,935	99,948	0,113
	$(Var(\hat{\theta}_a))^{-1}$	99,926	-1,325	0,323	0,323	99,920	99,933	0,122
	$*(Var(\hat{\theta}_a))^{-1}$	99,936	-0,400	0,330	0,330	99,929	99,942	0,111
	$(CV(\hat{\theta}_a))^{-1}$	99,934	-0,606	0,325	0,325	99,927	99,940	0,109
	$*(CV(\hat{\theta}_a))^{-1}$	99,939	-0,098	0,332	0,332	99,932	99,945	0,110
	$(skewness(s_{dc}))^{-1}$	99,937	-0,259	0,351	0,351	99,930	99,944	0,124

	$*(skewness(s_{dc}))^{-1}$	99,937	-0,259	0,351	0,351	99,930	99,944	0,124
	$(kurtosis(K_{dc}))^{-1}$	99,936	-0,351	0,396	0,396	99,928	99,944	0,158
	$*(kurtosis(K_{dc}))^{-1}$	99,936	-0,351	0,396	0,396	99,928	99,944	0,158
	$meff/Var(y)$	99,935	-0,426	0,329	0,330	99,929	99,942	0,110
	$(meff)^{-1}$	99,932	-0,807	0,325	0,325	99,925	99,938	0,112
15	$1/n$	99,940	0,032	0,260	0,260	99,935	99,945	0,068
	$(Var(\hat{\theta}_a))^{-1}$	99,924	-1,567	0,253	0,253	99,919	99,929	0,089
	$*(Var(\hat{\theta}_a))^{-1}$	99,933	-0,623	0,255	0,255	99,928	99,938	0,069
	$(CV(\hat{\theta}_a))^{-1}$	99,931	-0,819	0,253	0,253	99,927	99,936	0,071
	$*(CV(\hat{\theta}_a))^{-1}$	99,937	-0,291	0,257	0,257	99,932	99,942	0,067
	$(skewness(s_{dc}))^{-1}$	99,934	-0,523	0,268	0,268	99,929	99,940	0,074
	$*(skewness(s_{dc}))^{-1}$	99,934	-0,523	0,268	0,268	99,929	99,940	0,074
	$(kurtosis(K_{dc}))^{-1}$	99,933	-0,689	0,308	0,308	99,927	99,939	0,100
	$*(kurtosis(K_{dc}))^{-1}$	99,933	-0,689	0,308	0,308	99,927	99,939	0,100
	$meff/Var(y)$	99,935	-0,468	0,288	0,288	99,929	99,941	0,085
	$(meff)^{-1}$	99,931	-0,913	0,264	0,264	99,925	99,936	0,078
20	$1/n$	99,940	0,007	0,238	0,238	99,935	99,944	0,057
	$(Var(\hat{\theta}_a))^{-1}$	99,923	-1,634	0,229	0,229	99,919	99,928	0,079
	$*(Var(\hat{\theta}_a))^{-1}$	99,933	-0,666	0,234	0,234	99,928	99,938	0,059
	$(CV(\hat{\theta}_a))^{-1}$	99,931	-0,866	0,230	0,230	99,927	99,936	0,060
	$*(CV(\hat{\theta}_a))^{-1}$	99,936	-0,326	0,235	0,235	99,932	99,941	0,056
	$(skewness(s_{dc}))^{-1}$	99,934	-0,576	0,249	0,249	99,929	99,939	0,065
	$*(skewness(s_{dc}))^{-1}$	99,934	-0,576	0,249	0,249	99,929	99,939	0,065
	$(kurtosis(K_{dc}))^{-1}$	99,932	-0,794	0,281	0,282	99,926	99,937	0,086
	$*(kurtosis(K_{dc}))^{-1}$	99,932	-0,794	0,281	0,282	99,926	99,937	0,086
	$meff/Var(y)$	99,932	-0,746	0,252	0,252	99,927	99,937	0,069
	$(meff)^{-1}$	99,929	-1,085	0,231	0,231	99,924	99,933	0,065

D=number of samples combined, $\hat{\theta}_c$ = combined estimate, $\hat{\theta}_a$ = estimate calculated from each individual survey, Rel. Bias (%) = relative bias, SE ($\hat{\theta}_c$)=standard error of the combined estimate, CV ($\hat{\theta}_c$) (%) = coefficient of variation, $CI_{95\%}(\hat{\theta}_c)$ = 95% confidence interval, LB=lower 95% confidence limit, UB=upper 95% confidence limit, MSE=mean square error, method=weighting strategy.

Appendix E

R code used in simulations for simple random sampling, point estimators

E1 Normal Distribution

```
#####  
#####  
# Variance Estimation: Normal distribution  
# author: "Loveness Dzikiti"  
# date: "20 April 2017"  
#  
#####  
#Scenario 1: N=100000 N(100,25) n=1000 K=1000  
#####  
  
dagostino.pearson.test <- function(x) {  
  # from Zar (1999), implemented by Doug Scofield, scofield at bio.indiana.edu  
  DNAME <- deparse(substitute(x))  
  n <- length(x)  
  x2 <- x * x  
  x3 <- x * x2  
  x4 <- x * x3  
  # compute Z_g1  
  k3 <- ((n*sum(x3)) - (3*sum(x)*sum(x2)) + (2*(sum(x)^3)/n))/((n-1)*(n-2))  
  g1 <- k3 / sqrt(var(x)^3)  
  sqrtb1 <- ((n - 2)*g1) / sqrt(n*(n - 1))  
  A <- sqrtb1 * sqrt(((n + 1)*(n + 3)) / (6*(n - 2)))  
  B <- (3*(n*n + 27*n - 70)*(n+1)*(n+3)) / ((n-2)*(n+5)*(n+7)*(n+9))  
  C <- sqrt(2*(B - 1)) - 1  
  D <- sqrt(C)  
  E <- 1 / sqrt(log(D))  
  F <- A / sqrt(2/(C - 1))  
  Zg1 <- E * log(F + sqrt(F*F + 1))  
  # compute Z_g2  
  G <- (24*n*(n-2)*(n-3)) / (((n+1)^2)*(n+3)*(n+5))  
  k4 <- (((n*n*n + n*n)*sum(x4)) - (4*(n*n + n)*sum(x3)*sum(x)) -  
    (3*(n*n - n)*sum(x2)^2) + (12*n*sum(x2)*sum(x)^2) - (6*sum(x)^4)) /  
    (n*(n-1)*(n-2)*(n-3))  
  g2 <- k4 / var(x)^2  
  H <- ((n-2)*(n-3)*abs(g2)) / ((n+1)*(n-1)*sqrt(G))  
  J <- ((6*(n*n - 5*n + 2)) / ((n+7)*(n+9))) * sqrt((6*(n+3)*(n+5)) /  
    (n*(n-2)*(n-3)))  
  K <- 6 + (8/J)*(2/J + sqrt(1 + 4/(J*J)))  
  L <- (1 - 2/K) / (1 + H*sqrt(2/(K-4)))  
  Zg2 <- (1 - 2/(9*K) - (L^(1/3))) / (sqrt(2/(9*K)))  
  K2 <- Zg1*Zg1 + Zg2*Zg2  
  pk2 <- pchisq(K2, 2, lower.tail=FALSE)  
  RVAL <- list(statistic = c(K2 = K2), p.value = pk2, method =  
    "D'Agostino-Pearson normality test\n\nK2 is distributed as Chi-squared  
with df=2", alternative = "distribution is not normal", data.name =  
    DNAME)  
  class(RVAL) <- "htest"  
  #return(RVAL)  
  return(K2) # adapting the test to return K2 instead of the usual results stored in RVAL  
}  
#####  
  
setwd("C:\\Users\\User\\Dropbox\\Review\\simulations_2019\\srs\\normal\\scenario1\\nodo")  
  
library(moments)  
  
N=100000 #finite population size  
n=1000 #sample size  
nsamp=20 #number of samples
```



```

combmeans16 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
combmeans26 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
combmeans36 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans46 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans56 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5
combmeans66 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 6
combmeans76 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 7

#####

for (k in 1:nvar){

for (j in 1:nsamp){

w1<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1) #store weights
w2<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w3<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w4<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w5<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w6<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w7<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)

}

# Sampling from finite population
for (i in 1:nsamp) {

s[,i]<-sample(x,n)

means[i]<-mean(s[,i]) #sample mean for ith sample
variance[i]<-sum((s[,i]-means[i])^2)/(n-1) #sample variance
designvar[i]<- (1-(n/N))*(variance[i])/n #design consistent variance
cv[i]<-abs(sqrt(designvar[i])/means[i]) #coefficient of variation
skewness[i]<-skewness(s[,i])
kurtosis[i]<-kurtosis(s[,i])
DP[i]<-dagostino.pearson.test(s[,i]) #dagostino pearson statistic

w1[i]<-n #inverse of the sample size
w2[i]<-1/(designvar[i]) #inverse of the variance of the estimator of the mean
w3[i]=1 #simple average weights
w4[i]=1/cv[i] #inverse of cv
w5[i]=abs(1/skewness[i]) #inverse skewness
w6[i]=1/kurtosis[i] #inverse kurtosis
w7[i]=1/DP[i] #inverse dagostino pearson k_squared

combmeans1[i]<-round(sum(means[1:i]*w1[1:i])/sum(w1[1:i]),digits=9) #combine means using inverse sample size
combmeans2[i]<-round(sum(means[1:i]*w2[1:i])/sum(w2[1:i]),digits=9) # combine using inverse variance
combmeans3[i]<-round(sum(means[1:i]*w3[1:i])/sum(w3[1:i]),digits=9) # combine using a weight of 1
combmeans4[i]<-round(sum(means[1:i]*w4[1:i])/sum(w4[1:i]),digits=9) # combine using a weight of n
combmeans5[i]<-round(sum(means[1:i]*w5[1:i])/sum(w5[1:i]),digits=9) # combine using a weight of inverse CV
combmeans6[i]<-round(sum(means[1:i]*w6[1:i])/sum(w6[1:i]),digits=9) # combine using a weight of inverse skewness
combmeans7[i]<-round(sum(means[1:i]*w7[1:i])/sum(w7[1:i]),digits=9) # combine using a weight of inverse kurtosis

}

sampmean[k]=means[1] # capturing the combined mean, where nsamp reps are combined, all weighting strategies

combmeans11[k]=combmeans1[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans21[k]=combmeans2[2]
combmeans31[k]=combmeans3[2]
combmeans41[k]=combmeans4[2]
combmeans51[k]=combmeans5[2]
combmeans61[k]=combmeans6[2]
combmeans71[k]=combmeans7[2]

combmeans12[k]=combmeans1[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans22[k]=combmeans2[3]
combmeans32[k]=combmeans3[3]

```

```

combmeans42[k]=combmeans4[3]
combmeans52[k]=combmeans5[3]
combmeans62[k]=combmeans6[3]
combmeans72[k]=combmeans7[3]

combmeans13[k]=combmeans1[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans23[k]=combmeans2[5]
combmeans33[k]=combmeans3[5]
combmeans43[k]=combmeans4[5]
combmeans53[k]=combmeans5[5]
combmeans63[k]=combmeans6[5]
combmeans73[k]=combmeans7[5]

combmeans14[k]=combmeans1[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans24[k]=combmeans2[10]
combmeans34[k]=combmeans3[10]
combmeans44[k]=combmeans4[10]
combmeans54[k]=combmeans5[10]
combmeans64[k]=combmeans6[10]
combmeans74[k]=combmeans7[10]

combmeans15[k]=combmeans1[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans25[k]=combmeans2[15]
combmeans35[k]=combmeans3[15]
combmeans45[k]=combmeans4[15]
combmeans55[k]=combmeans5[15]
combmeans65[k]=combmeans6[15]
combmeans75[k]=combmeans7[15]

combmeans16[k]=combmeans1[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans26[k]=combmeans2[20]
combmeans36[k]=combmeans3[20]
combmeans46[k]=combmeans4[20]
combmeans56[k]=combmeans5[20]
combmeans66[k]=combmeans6[20]
combmeans76[k]=combmeans7[20]

}

expected1=mean(sampmean)      #expected mean for one sample, not combining
expected=matrix(0,7,6)

expected[1,1]=mean(combmeans11) #strategy 1 nsamp=2
expected[2,1]=mean(combmeans21)
expected[3,1]=mean(combmeans31)
expected[4,1]=mean(combmeans41)
expected[5,1]=mean(combmeans51)
expected[6,1]=mean(combmeans61)
expected[7,1]=mean(combmeans71)

expected[1,2]=mean(combmeans12) #strategy 1 nsamp=3
expected[2,2]=mean(combmeans22)
expected[3,2]=mean(combmeans32)
expected[4,2]=mean(combmeans42)
expected[5,2]=mean(combmeans52)
expected[6,2]=mean(combmeans62)
expected[7,2]=mean(combmeans72)

expected[1,3]=mean(combmeans13) #strategy 1 nsamp=5
expected[2,3]=mean(combmeans23)
expected[3,3]=mean(combmeans33)
expected[4,3]=mean(combmeans43)
expected[5,3]=mean(combmeans53)
expected[6,3]=mean(combmeans63)
expected[7,3]=mean(combmeans73)

expected[1,4]=mean(combmeans14) #strategy 1 nsamp=10
expected[2,4]=mean(combmeans24)
expected[3,4]=mean(combmeans34)
expected[4,4]=mean(combmeans44)
expected[5,4]=mean(combmeans54)

```

```
expected[6,4]=mean(combmeans64)
expected[7,4]=mean(combmeans74)
```

```
expected[1,5]=mean(combmeans15) #strategy 1 nsamp=2
expected[2,5]=mean(combmeans25)
expected[3,5]=mean(combmeans35)
expected[4,5]=mean(combmeans45)
expected[5,5]=mean(combmeans55)
expected[6,5]=mean(combmeans65)
expected[7,5]=mean(combmeans75)
```

```
expected[1,6]=mean(combmeans16) #strategy 1 nsamp=20
expected[2,6]=mean(combmeans26)
expected[3,6]=mean(combmeans36)
expected[4,6]=mean(combmeans46)
expected[5,6]=mean(combmeans56)
expected[6,6]=mean(combmeans66)
expected[7,6]=mean(combmeans76)
```

```
expectedcol=c(expected1,expected[,1],expected[,2],expected[,3],expected[,4],expected[,5],expected[,6])
```

```
means=cbind(combmeans11,combmeans21, combmeans31,combmeans41,combmeans51,combmeans61,combmeans71,
combmeans12,combmeans22, combmeans32,combmeans42,combmeans52,combmeans62,combmeans72,
combmeans13,combmeans23, combmeans33,combmeans43,combmeans53,combmeans63,combmeans73,
combmeans14,combmeans24, combmeans34,combmeans44,combmeans54,combmeans64,combmeans74,
combmeans15,combmeans25, combmeans35,combmeans45,combmeans55,combmeans65,combmeans75,
combmeans16,combmeans26, combmeans36,combmeans46,combmeans56,combmeans66,combmeans76)
```

```
variance1=(1/(nvar))*(1/1)*sum(sampmean[,1]-expected1)**2
```

```
variance11=(1/(nvar))*(1/(2-1))*sum(means[,1]-expected[1,1])**2
variance21=(1/(nvar))*(1/(2-1))*sum(means[,2]-expected[2,1])**2
variance31=(1/(nvar))*(1/(2-1))*sum(means[,3]-expected[3,1])**2
variance41=(1/(nvar))*(1/(2-1))*sum(means[,4]-expected[4,1])**2
variance51=(1/(nvar))*(1/(2-1))*sum(means[,5]-expected[5,1])**2
variance61=(1/(nvar))*(1/(2-1))*sum(means[,6]-expected[6,1])**2
variance71=(1/(nvar))*(1/(2-1))*sum(means[,7]-expected[7,1])**2
```

```
variance12=(1/(nvar))*(1/(3-1))*sum(means[,8]-expected[1,2])**2
variance22=(1/(nvar))*(1/(3-1))*sum(means[,9]-expected[2,2])**2
variance32=(1/(nvar))*(1/(3-1))*sum(means[,10]-expected[3,2])**2
variance42=(1/(nvar))*(1/(3-1))*sum(means[,11]-expected[4,2])**2
variance52=(1/(nvar))*(1/(3-1))*sum(means[,12]-expected[5,2])**2
variance62=(1/(nvar))*(1/(3-1))*sum(means[,13]-expected[6,2])**2
variance72=(1/(nvar))*(1/(3-1))*sum(means[,14]-expected[7,2])**2
```

```
variance13=(1/(nvar))*(1/(5-1))*sum(means[,15]-expected[1,3])**2
variance23=(1/(nvar))*(1/(5-1))*sum(means[,16]-expected[2,3])**2
variance33=(1/(nvar))*(1/(5-1))*sum(means[,17]-expected[3,3])**2
variance43=(1/(nvar))*(1/(5-1))*sum(means[,18]-expected[4,3])**2
variance53=(1/(nvar))*(1/(5-1))*sum(means[,19]-expected[5,3])**2
variance63=(1/(nvar))*(1/(5-1))*sum(means[,20]-expected[6,3])**2
variance73=(1/(nvar))*(1/(5-1))*sum(means[,21]-expected[7,3])**2
```

```
variance14=(1/(nvar))*(1/(10-1))*sum(means[,22]-expected[1,4])**2
variance24=(1/(nvar))*(1/(10-1))*sum(means[,23]-expected[2,4])**2
variance34=(1/(nvar))*(1/(10-1))*sum(means[,24]-expected[3,4])**2
variance44=(1/(nvar))*(1/(10-1))*sum(means[,25]-expected[4,4])**2
variance54=(1/(nvar))*(1/(10-1))*sum(means[,26]-expected[5,4])**2
variance64=(1/(nvar))*(1/(10-1))*sum(means[,27]-expected[6,4])**2
variance74=(1/(nvar))*(1/(10-1))*sum(means[,28]-expected[7,4])**2
```

```
variance15=(1/(nvar))*(1/(15-1))*sum(means[,29]-expected[1,5])**2
variance25=(1/(nvar))*(1/(15-1))*sum(means[,30]-expected[2,5])**2
variance35=(1/(nvar))*(1/(15-1))*sum(means[,31]-expected[3,5])**2
variance45=(1/(nvar))*(1/(15-1))*sum(means[,32]-expected[4,5])**2
variance55=(1/(nvar))*(1/(15-1))*sum(means[,33]-expected[5,5])**2
variance65=(1/(nvar))*(1/(15-1))*sum(means[,34]-expected[6,5])**2
variance75=(1/(nvar))*(1/(15-1))*sum(means[,35]-expected[7,5])**2
```

```
variance16=(1/(nvar))*(1/(20-1))*sum(means[,36]-expected[1,6])**2
```

```

variance26=(1/(nvar))*(1/(20-1))*sum(means[,37]-expected[2,6])**2
variance36=(1/(nvar))*(1/(20-1))*sum(means[,38]-expected[3,6])**2
variance46=(1/(nvar))*(1/(20-1))*sum(means[,39]-expected[4,6])**2
variance56=(1/(nvar))*(1/(20-1))*sum(means[,40]-expected[5,6])**2
variance66=(1/(nvar))*(1/(20-1))*sum(means[,41]-expected[6,6])**2
variance76=(1/(nvar))*(1/(20-1))*sum(means[,42]-expected[7,6])**2

```

```

variance=c(variance1,variance11,variance21,variance31,variance41,variance51, variance61,variance71,
variance12,variance22,variance32,variance42,variance52, variance62,variance72,
variance13,variance23,variance33,variance43,variance53, variance63,variance73,
variance14,variance24,variance34,variance44,variance54, variance64,variance74,
variance15,variance25,variance35,variance45,variance55, variance65,variance75,
variance16,variance26,variance36,variance46,variance56, variance66,variance76)

```

```

bias= expectedcol-popmean

```

```

relbias=(bias/expectedcol)*100
mse= variance + bias**2
cv= (sqrt(variance)/expectedcol)*100
s.e= (sqrt(variance)/sqrt(nsamp))
LB= expectedcol - 1.96*s.e
UB= expectedcol + 1.96*s.e

```

```

result=cbind(popmean,expectedcol,relbias,s.e,cv,LB,UB,mse)
write.table(result,"result_normal_scenario1.txt",sep="\t")

```

E2 T Distribution, 3 degrees of freedom

```
#####
# Simulations for the T- distribution: Convergence and other performance measures of a
# Sequence of Simple Random Samples from a single Finite population using "7 DIFFERENT
# WEIGHTING STRATEGIES" with varying population size, sample size and variance.
# author: "Loveness Dzikiti"
# date: "23 October 2016"
#####
#
# Scenario 1: N=1000000 t(3)(100,25) n=1000 D=nsamp=20
#
#####
# Function 1.

dagostino.pearson.test <- function(x) {
  # from Zar (1999), implemented by Doug Scofield, scofield at bio.indiana.edu
  DNAME <- deparse(substitute(x))
  n <- length(x)
  x2 <- x * x
  x3 <- x * x2
  x4 <- x * x3
  # compute Z_g1
  k3 <- ((n*sum(x3)) - (3*sum(x)*sum(x2)) + (2*(sum(x)^3)/n))/((n-1)*(n-2))
  g1 <- k3 / sqrt(var(x)^3)
  sqrtb1 <- ((n - 2)*g1) / sqrt(n*(n - 1))
  A <- sqrtb1 * sqrt(((n + 1)*(n + 3)) / (6*(n - 2)))
  B <- (3*(n*n + 27*n - 70)*(n+1)*(n+3)) / ((n-2)*(n+5)*(n+7)*(n+9))
  C <- sqrt(2*(B - 1)) - 1
  D <- sqrt(C)
  E <- 1 / sqrt(log(D))
  F <- A / sqrt(2/(C - 1))
  Zg1 <- E * log(F + sqrt(F*F + 1))
  # compute Z_g2
  G <- (24*n*(n-2)*(n-3)) / (((n+1)^2)*(n+3)*(n+5))
  k4 <- (((n*n*n + n*n)*sum(x4)) - (4*(n*n + n)*sum(x3)*sum(x)) -
    (3*(n*n - n)*sum(x2)^2) + (12*n*sum(x2)*sum(x)^2) - (6*sum(x)^4)) /
    (n*(n-1)*(n-2)*(n-3))
  g2 <- k4 / var(x)^2
  H <- ((n-2)*(n-3)*abs(g2)) / ((n+1)*(n-1)*sqrt(G))
  J <- ((6*(n*n - 5*n + 2)) / ((n+7)*(n+9))) * sqrt((6*(n+3)*(n+5)) /
    (n*(n-2)*(n-3)))
  K <- 6 + (8/J)*(2/J + sqrt(1 + 4/(J*J)))
  L <- (1 - 2/K) / (1 + H*sqrt(2/(K-4)))
  Zg2 <- (1 - 2/(9*K) - (L^(1/3))) / (sqrt(2/(9*K)))
  K2 <- Zg1*Zg1 + Zg2*Zg2
  pk2 <- pchisq(K2, 2, lower.tail=FALSE)
  RVAL <- list(statistic = c(K2 = K2), p.value = pk2, method =
    "D'Agostino-Pearson normality test\n\nK2 is distributed as Chi-squared
    with df=2", alternative = "distribution is not normal", data.name =
    DNAME)
  class(RVAL) <- "htest"
  #return(RVAL)
  return(K2) # adapting the test to return K2 instead of the usual results stored in RVAL
}

#####
# Function 2
# the rt function in R does not allow one to specify variance. to compare, with the normal distribution, we have to transform
# the data given by the rt function.
# For r=2, the following expression gives the variance of the t distribution, according to # Johnson, Kotz and Balakrishnan, 1995
# See also Vieira (2005, Phd Thesis, pages 115-117
```

```
#####
#r=2
#t_Var<-v^(r/2)*((gamma(1/2*(r+1))*gamma(1/2*(v-r)))/(gamma(0.5)*gamma(1/2*v))
#
# Generates a t distribution with population size N, v degrees of freedom, mean=0 and unspecified standard deviation
#N=50000000
#v=5
#t=rt(N,v)
#
# Standardizing the standard deviation to be equal to 1
#t<-t/sqrt(t_Var)

# Standardizing the standard deviation to be sigma and mean to be mu
#sigma=5 # then variance=25
#mu=100
#t=(t*sigma)+mu
#####
##

setwd("C:\\Users\\User\\Dropbox\\Review\\simulations_2019\\srs\\t3")

library(moments)

N=1000000      #finite population size
n=1000         #sample size
nsamp=20       #number of samples
k=nvar=1000

set.seed(123456)

#####
# T distribution with 3 degrees of freedom

v=3
r=2
mu=100
sigma=5

Var<-v^(r/2)*((gamma(1/2*(r+1))*gamma(1/2*(v-r)))/(gamma(0.5)*gamma(1/2*v)) # variance of t-distribution
x<-rt(N,v)      #simulate data from a t distribution random mean and variance
x<-x/sqrt(Var)
x=(x*sigma)+mu

#####
#####

popmean=mean(x) #finite population mean

s <- matrix(c(rep.int(NA,n)), nrow = n, ncol = nsamp) #store samples
means <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample means
variance <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample variance
designvar <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store design consistent variance
cv <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample coefficient of variation
skewness <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample skewness
kurtosis <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample skewness
DP <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample dagostino k_squared

combmeans1 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 1
combmeans2 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 2
combmeans3 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 3
combmeans4 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 4
combmeans5 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 5
combmeans6 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 6
combmeans7 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 7

samplemean <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1

combmeans11 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
combmeans21 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
```



```

designvar[i]<- (1-(n/N))*(variance[i])/n      #design consistent variance
cv[i]<-abs(sqrt(designvar[i])/means[i])      #coefficient of variation
skewness[i]<-skewness(s[,i])
kurtosis[i]<-kurtosis(s[,i])
DP[i]<-dagostino.pearson.test(s[,i])         #dagostino pearson statistic

w1[i]<-1/n      #inverse of the sample size
w2[i]<-1/(designvar[i]) #inverse of the variance of the estimator of the mean
w3[i]=1        #simple average weights
w4[i]=1/cv[i]  #inverse of cv
w5[i]=abs(1/skewness[i]) #inverse skewness
w6[i]=1/kurtosis[i] #inverse kurtosis
w7[i]=1/DP[i]  #inverse dagostino pearson k_squared

combmeans1[i]<-round(sum(means[1:i]*w1[1:i])/sum(w1[1:i]),digits=9) #combine means using inverse sample size
combmeans2[i]<-round(sum(means[1:i]*w2[1:i])/sum(w2[1:i]),digits=9) # combine using inverse variance
combmeans3[i]<-round(sum(means[1:i]*w3[1:i])/sum(w3[1:i]),digits=9) # combine using a weight of 1
combmeans4[i]<-round(sum(means[1:i]*w4[1:i])/sum(w4[1:i]),digits=9) # combine using a weight of n
combmeans5[i]<-round(sum(means[1:i]*w5[1:i])/sum(w5[1:i]),digits=9) # combine using a weight of inverse CV
combmeans6[i]<-round(sum(means[1:i]*w6[1:i])/sum(w6[1:i]),digits=9) # combine using a weight of inverse skewness
combmeans7[i]<-round(sum(means[1:i]*w7[1:i])/sum(w7[1:i]),digits=9) # combine using a weight of inverse kurtosis

}

popmean=mean(x) #finite population mean

#####Measures of performance of the different estimators

sampmean[k]=means[1] # capturing the mean of 1 sample, where nsamp=1 reps are combined

combmeans11[k]=combmeans1[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans21[k]=combmeans2[2]
combmeans31[k]=combmeans3[2]
combmeans41[k]=combmeans4[2]
combmeans51[k]=combmeans5[2]
combmeans61[k]=combmeans6[2]
combmeans71[k]=combmeans7[2]

combmeans12[k]=combmeans1[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans22[k]=combmeans2[3]
combmeans32[k]=combmeans3[3]
combmeans42[k]=combmeans4[3]
combmeans52[k]=combmeans5[3]
combmeans62[k]=combmeans6[3]
combmeans72[k]=combmeans7[3]

combmeans13[k]=combmeans1[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans23[k]=combmeans2[5]
combmeans33[k]=combmeans3[5]
combmeans43[k]=combmeans4[5]
combmeans53[k]=combmeans5[5]
combmeans63[k]=combmeans6[5]
combmeans73[k]=combmeans7[5]

combmeans14[k]=combmeans1[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans24[k]=combmeans2[10]
combmeans34[k]=combmeans3[10]
combmeans44[k]=combmeans4[10]
combmeans54[k]=combmeans5[10]
combmeans64[k]=combmeans6[10]
combmeans74[k]=combmeans7[10]

combmeans15[k]=combmeans1[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans25[k]=combmeans2[15]
combmeans35[k]=combmeans3[15]
combmeans45[k]=combmeans4[15]
combmeans55[k]=combmeans5[15]

```

```

combmeans65[k]=combmeans6[15]
combmeans75[k]=combmeans7[15]

combmeans16[k]=combmeans1[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans26[k]=combmeans2[20]
combmeans36[k]=combmeans3[20]
combmeans46[k]=combmeans4[20]
combmeans56[k]=combmeans5[20]
combmeans66[k]=combmeans6[20]
combmeans76[k]=combmeans7[20]

}

expected1=mean(sampmean)      #expected mean for one sample, not combining
expected=matrix(0,7,6)

expected[1,1]=mean(combmeans11) #strategy 1 nsamp=2
expected[2,1]=mean(combmeans21)
expected[3,1]=mean(combmeans31)
expected[4,1]=mean(combmeans41)
expected[5,1]=mean(combmeans51)
expected[6,1]=mean(combmeans61)
expected[7,1]=mean(combmeans71)

expected[1,2]=mean(combmeans12) #strategy 1 nsamp=3
expected[2,2]=mean(combmeans22)
expected[3,2]=mean(combmeans32)
expected[4,2]=mean(combmeans42)
expected[5,2]=mean(combmeans52)
expected[6,2]=mean(combmeans62)
expected[7,2]=mean(combmeans72)

expected[1,3]=mean(combmeans13) #strategy 1 nsamp=5
expected[2,3]=mean(combmeans23)
expected[3,3]=mean(combmeans33)
expected[4,3]=mean(combmeans43)
expected[5,3]=mean(combmeans53)
expected[6,3]=mean(combmeans63)
expected[7,3]=mean(combmeans73)

expected[1,4]=mean(combmeans14) #strategy 1 nsamp=10
expected[2,4]=mean(combmeans24)
expected[3,4]=mean(combmeans34)
expected[4,4]=mean(combmeans44)
expected[5,4]=mean(combmeans54)
expected[6,4]=mean(combmeans64)
expected[7,4]=mean(combmeans74)

expected[1,5]=mean(combmeans15) #strategy 1 nsamp=2
expected[2,5]=mean(combmeans25)
expected[3,5]=mean(combmeans35)
expected[4,5]=mean(combmeans45)
expected[5,5]=mean(combmeans55)
expected[6,5]=mean(combmeans65)
expected[7,5]=mean(combmeans75)

expected[1,6]=mean(combmeans16) #strategy 1 nsamp=20
expected[2,6]=mean(combmeans26)
expected[3,6]=mean(combmeans36)
expected[4,6]=mean(combmeans46)
expected[5,6]=mean(combmeans56)
expected[6,6]=mean(combmeans66)
expected[7,6]=mean(combmeans76)

expectedcol=c(expected1,expected[,1],expected[,2],expected[,3],expected[,4],expected[,5],expected[,6])

means=cbind(combmeans11,combmeans21, combmeans31,combmeans41,combmeans51,combmeans61,combmeans71,
combmeans12,combmeans22, combmeans32,combmeans42,combmeans52,combmeans62,combmeans72,
combmeans13,combmeans23, combmeans33,combmeans43,combmeans53,combmeans63,combmeans73,

```

combmeans14,combmeans24, combmeans34,combmeans44,combmeans54,combmeans64,combmeans74,
 combmeans15,combmeans25, combmeans35,combmeans45,combmeans55,combmeans65,combmeans75,
 combmeans16,combmeans26, combmeans36,combmeans46,combmeans56,combmeans66,combmeans76)

$$\text{variance1}=(1/(\text{nvar}))*\text{sum}((\text{sampmean}[,1]-\text{expected1})**\text{2})$$

$$\begin{aligned} \text{variance11}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,1]-\text{expected}[1,1])**\text{2}) \\ \text{variance21}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,2]-\text{expected}[2,1])**\text{2}) \\ \text{variance31}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,3]-\text{expected}[3,1])**\text{2}) \\ \text{variance41}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,4]-\text{expected}[4,1])**\text{2}) \\ \text{variance51}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,5]-\text{expected}[5,1])**\text{2}) \\ \text{variance61}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,6]-\text{expected}[6,1])**\text{2}) \\ \text{variance71}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,7]-\text{expected}[7,1])**\text{2}) \end{aligned}$$

$$\begin{aligned} \text{variance12}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,8]-\text{expected}[1,2])**\text{2}) \\ \text{variance22}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,9]-\text{expected}[2,2])**\text{2}) \\ \text{variance32}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,10]-\text{expected}[3,2])**\text{2}) \\ \text{variance42}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,11]-\text{expected}[4,2])**\text{2}) \\ \text{variance52}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,12]-\text{expected}[5,2])**\text{2}) \\ \text{variance62}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,13]-\text{expected}[6,2])**\text{2}) \\ \text{variance72}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,14]-\text{expected}[7,2])**\text{2}) \end{aligned}$$

$$\begin{aligned} \text{variance13}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,15]-\text{expected}[1,3])**\text{2}) \\ \text{variance23}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,16]-\text{expected}[2,3])**\text{2}) \\ \text{variance33}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,17]-\text{expected}[3,3])**\text{2}) \\ \text{variance43}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,18]-\text{expected}[4,3])**\text{2}) \\ \text{variance53}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,19]-\text{expected}[5,3])**\text{2}) \\ \text{variance63}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,20]-\text{expected}[6,3])**\text{2}) \\ \text{variance73}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,21]-\text{expected}[7,3])**\text{2}) \end{aligned}$$

$$\begin{aligned} \text{variance14}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,22]-\text{expected}[1,4])**\text{2}) \\ \text{variance24}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,23]-\text{expected}[2,4])**\text{2}) \\ \text{variance34}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,24]-\text{expected}[3,4])**\text{2}) \\ \text{variance44}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,25]-\text{expected}[4,4])**\text{2}) \\ \text{variance54}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,26]-\text{expected}[5,4])**\text{2}) \\ \text{variance64}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,27]-\text{expected}[6,4])**\text{2}) \\ \text{variance74}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,28]-\text{expected}[7,4])**\text{2}) \end{aligned}$$

$$\begin{aligned} \text{variance15}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,29]-\text{expected}[1,5])**\text{2}) \\ \text{variance25}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,30]-\text{expected}[2,5])**\text{2}) \\ \text{variance35}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,31]-\text{expected}[3,5])**\text{2}) \\ \text{variance45}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,32]-\text{expected}[4,5])**\text{2}) \\ \text{variance55}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,33]-\text{expected}[5,5])**\text{2}) \\ \text{variance65}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,34]-\text{expected}[6,5])**\text{2}) \\ \text{variance75}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,35]-\text{expected}[7,5])**\text{2}) \end{aligned}$$

$$\begin{aligned} \text{variance16}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,36]-\text{expected}[1,6])**\text{2}) \\ \text{variance26}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,37]-\text{expected}[2,6])**\text{2}) \\ \text{variance36}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,38]-\text{expected}[3,6])**\text{2}) \\ \text{variance46}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,39]-\text{expected}[4,6])**\text{2}) \\ \text{variance56}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,40]-\text{expected}[5,6])**\text{2}) \\ \text{variance66}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,41]-\text{expected}[6,6])**\text{2}) \\ \text{variance76}&=(1/(\text{nvar}-1))*\text{sum}((\text{means}[,42]-\text{expected}[7,6])**\text{2}) \end{aligned}$$

variance=c(variance1,variance11,variance21,variance31,variance41,variance51, variance61,variance71,
 variance12,variance22,variance32,variance42,variance52, variance62,variance72,
 variance13,variance23,variance33,variance43,variance53, variance63,variance73,
 variance14,variance24,variance34,variance44,variance54, variance64,variance74,
 variance15,variance25,variance35,variance45,variance55, variance65,variance75,
 variance16,variance26,variance36,variance46,variance56, variance66,variance76)

bias= expectedcol-popmean

$$\text{relbias}=(\text{bias}/\text{expectedcol})*\text{100}$$

$$\text{mse}=\text{variance}+\text{bias}**\text{2}$$

$$\text{cv}=(\text{sqrt}(\text{variance})/\text{expectedcol})*\text{100}$$

$$\text{s.e.}=(\text{sqrt}(\text{variance})/\text{sqrt}(\text{nvar}))$$

```
LB= expectedcol - 1.96*s.e
UB= expectedcol + 1.96*s.e

result=cbind(popmean,expectedcol,relbias,s.e,cv,LB,UB,mse)

#write.table(result,"result_t3_scenario1_1000.txt",sep="\t")
write.table(result,"million_t3_scenario1_1000.txt",sep="\t")

rm(list = ls())
```

E3 Skewed Normal Distribution

```
#####
# Simulations for the Skewed Normal distribution: Convergence and other performance measures of
# a Sequence of Simple Random Samples from a single Finite population using "7 DIFFERENT
# WEIGHTING STRATEGIES" with varying population size, sample size and variance.
# author: "Loveness Dzikiti"
# date: "27 September 2017"
#####

#Scenario 1: N=1000000 SN(100,25) n=1000 D=nsamp=20
#####

dagostino.pearson.test <- function(x) {
  # from Zar (1999), implemented by Doug Scofield, scofield at bio.indiana.edu
  DNAME <- deparse(substitute(x))
  n <- length(x)
  x2 <- x * x
  x3 <- x * x2
  x4 <- x * x3
  # compute Z_g1
  k3 <- ((n*sum(x3)) - (3*sum(x)*sum(x2)) + (2*(sum(x)^3)/n))/((n-1)*(n-2))
  g1 <- k3 / sqrt(var(x)^3)
  sqrtb1 <- ((n - 2)*g1) / sqrt(n*(n - 1))
  A <- sqrtb1 * sqrt(((n + 1)*(n + 3)) / (6*(n - 2)))
  B <- (3*(n*n + 27*n - 70)*(n+1)*(n+3)) / ((n-2)*(n+5)*(n+7)*(n+9))
  C <- sqrt(2*(B - 1)) - 1
  D <- sqrt(C)
  E <- 1 / sqrt(log(D))
  F <- A / sqrt(2/(C - 1))
  Zg1 <- E * log(F + sqrt(F*F + 1))
  # compute Z_g2
  G <- (24*n*(n-2)*(n-3)) / (((n+1)^2)*(n+3)*(n+5))
  k4 <- (((n*n*n + n*n)*sum(x4)) - (4*(n*n + n)*sum(x3)*sum(x)) -
    (3*(n*n - n)*sum(x2)^2) + (12*n*sum(x2)*sum(x)^2) - (6*sum(x)^4)) /
    (n*(n-1)*(n-2)*(n-3))
  g2 <- k4 / var(x)^2
  H <- ((n-2)*(n-3)*abs(g2)) / (((n+1)*(n-1)*sqrt(G))
  J <- ((6*(n*n - 5*n + 2)) / ((n+7)*(n+9))) * sqrt((6*(n+3)*(n+5)) /
    (n*(n-2)*(n-3)))
  K <- 6 + (8/J)*(2/J + sqrt(1 + 4/(J*J)))
  L <- (1 - 2/K) / (1 + H*sqrt(2/(K-4)))
  Zg2 <- (1 - 2/(9*K) - (L^(1/3))) / (sqrt(2/(9*K)))
  K2 <- Zg1*Zg1 + Zg2*Zg2
  pk2 <- pchisq(K2, 2, lower.tail=FALSE)
  RVAL <- list(statistic = c(K2 = K2), p.value = pk2, method =
    "D'Agostino-Pearson normality test\n\nK2 is distributed as Chi-squared
    with df=2", alternative = "distribution is not normal", data.name =
    DNAME)
  class(RVAL) <- "htest"
  #return(RVAL)
  return(K2) # adapting the test to return K2 instead of the usual results stored in RVAL
}
#####

setwd("C:\\Users\\User\\Dropbox\\Review\\simulations_2019\\srs\\anormal")

library(moments)
library(sn)

N=1000000 #finite population size
n=1000 #sample size
nsamp=20 #number of samples
k=nvar=1000

set.seed(123456)

#####
# Asymmetric Normal Distribution #
```

```

xi=-1.15          #
omega=1.541      #
alpha=3          # asymmetry parameter for strong asymmetry

x <- rsn(N, xi, omega, alpha)

sigma=5
mu=100

x=(x*sigma)+mu

#####
#####

popmean=mean(x) #finite population mean

s <- matrix(c(rep.int(NA,n)), nrow = n, ncol = nsamp) #store samples
means <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample means
variance <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample variance
designvar <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store design consistent variance
cv <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample coefficient of variation
skewness <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample skewness
kurtosis <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample kurtosis
DP <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample dagostino k_squared

combmeans1 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 1
combmeans2 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 2
combmeans3 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 3
combmeans4 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 4
combmeans5 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 5
combmeans6 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 6
combmeans7 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 7

sampmean <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1

combmeans11 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
combmeans21 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
combmeans31 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans41 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans51 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5
combmeans61 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 6
combmeans71 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 7

combmeans12 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
combmeans22 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
combmeans32 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans42 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans52 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5
combmeans62 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 6
combmeans72 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 7

combmeans13 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
combmeans23 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
combmeans33 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans43 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans53 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5
combmeans63 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 6
combmeans73 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 7

combmeans14 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
combmeans24 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
combmeans34 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans44 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans54 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5
combmeans64 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 6
combmeans74 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 7

```

```

combmeans15 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
combmeans25 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
combmeans35 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans45 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans55 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5
combmeans65 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 6
combmeans75 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 7

```

```

combmeans16 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
combmeans26 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
combmeans36 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans46 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans56 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5
combmeans66 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 6
combmeans76 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 7

```

```
#####
```

```
for (k in 1:nvar){
```

```
  for (j in 1:nsamp){
```

```

    w1<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1) #store weights
    w2<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
    w3<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
    w4<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
    w5<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
    w6<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
    w7<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)

```

```
  }
```

```
##### Sampling from finite population#####
```

```
for (i in 1:nsamp) {
```

```
  s[,i]<-sample(x,n)
```

```

  means[i]<-mean(s[,i])           #sample mean for ith sample
  variance[i]<-sum((s[,i]-means[i])^2)/(n-1)   #sample variance
  designvar[i]<- (1-(n/N))*(variance[i])/n      #design consistent variance
  cv[i]<-abs(sqrt(designvar[i])/means[i])       #coefficient of variation
  skewness[i]<-skewness(s[,i])
  kurtosis[i]<-kurtosis(s[,i])
  DP[i]<-dagostino.pearson.test(s[,i])         #dagostino pearson statistic

```

```

  w1[i]<-1/n           #inverse of the sample size
  w2[i]<-1/(designvar[i]) #inverse of the variance of the estimator of the mean
  w3[i]=1             #simple average weights
  w4[i]=1/cv[i]       #inverse of cv
  w5[i]=abs(1/skewness[i]) #inverse skewness
  w6[i]=1/kurtosis[i] #inverse kurtosis
  w7[i]=1/DP[i]       #inverse dagostino pearson k_squared

```

```

  combmeans1[i]<-round(sum(means[1:i]*w1[1:i])/sum(w1[1:i]),digits=9) #combine means using inverse sample size
  combmeans2[i]<-round(sum(means[1:i]*w2[1:i])/sum(w2[1:i]),digits=9) # combine using inverse variance
  combmeans3[i]<-round(sum(means[1:i]*w3[1:i])/sum(w3[1:i]),digits=9) # combine using a weight of 1
  combmeans4[i]<-round(sum(means[1:i]*w4[1:i])/sum(w4[1:i]),digits=9) # combine using a weight of n
  combmeans5[i]<-round(sum(means[1:i]*w5[1:i])/sum(w5[1:i]),digits=9) # combine using a weight of inverse CV
  combmeans6[i]<-round(sum(means[1:i]*w6[1:i])/sum(w6[1:i]),digits=9) # combine using a weight of inverse skewness
  combmeans7[i]<-round(sum(means[1:i]*w7[1:i])/sum(w7[1:i]),digits=9) # combine using a weight of inverse kurtosis

```

```
}
```

```
#####Measures of performance of the different estimators#####
```

```
sampmean[k]=means[1] # capturing the mean of 1 sample, where nsamp=1 reps are combined
```

```
combmeans11[k]=combmeans1[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies  
combmeans21[k]=combmeans2[2]  
combmeans31[k]=combmeans3[2]  
combmeans41[k]=combmeans4[2]  
combmeans51[k]=combmeans5[2]  
combmeans61[k]=combmeans6[2]  
combmeans71[k]=combmeans7[2]
```

```
combmeans12[k]=combmeans1[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies  
combmeans22[k]=combmeans2[3]  
combmeans32[k]=combmeans3[3]  
combmeans42[k]=combmeans4[3]  
combmeans52[k]=combmeans5[3]  
combmeans62[k]=combmeans6[3]  
combmeans72[k]=combmeans7[3]
```

```
combmeans13[k]=combmeans1[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies  
combmeans23[k]=combmeans2[5]  
combmeans33[k]=combmeans3[5]  
combmeans43[k]=combmeans4[5]  
combmeans53[k]=combmeans5[5]  
combmeans63[k]=combmeans6[5]  
combmeans73[k]=combmeans7[5]
```

```
combmeans14[k]=combmeans1[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies  
combmeans24[k]=combmeans2[10]  
combmeans34[k]=combmeans3[10]  
combmeans44[k]=combmeans4[10]  
combmeans54[k]=combmeans5[10]  
combmeans64[k]=combmeans6[10]  
combmeans74[k]=combmeans7[10]
```

```
combmeans15[k]=combmeans1[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies  
combmeans25[k]=combmeans2[15]  
combmeans35[k]=combmeans3[15]  
combmeans45[k]=combmeans4[15]  
combmeans55[k]=combmeans5[15]  
combmeans65[k]=combmeans6[15]  
combmeans75[k]=combmeans7[15]
```

```
combmeans16[k]=combmeans1[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies  
combmeans26[k]=combmeans2[20]  
combmeans36[k]=combmeans3[20]  
combmeans46[k]=combmeans4[20]  
combmeans56[k]=combmeans5[20]  
combmeans66[k]=combmeans6[20]  
combmeans76[k]=combmeans7[20]
```

```
}
```

```
expected1=mean(sampmean) #expected mean for one sample, not combining  
expected=matrix(0,7,6)
```

```
expected[1,1]=mean(combmeans11) #strategy 1 nsamp=2  
expected[2,1]=mean(combmeans21)  
expected[3,1]=mean(combmeans31)  
expected[4,1]=mean(combmeans41)  
expected[5,1]=mean(combmeans51)  
expected[6,1]=mean(combmeans61)  
expected[7,1]=mean(combmeans71)
```

```
expected[1,2]=mean(combmeans12) #strategy 1 nsamp=3  
expected[2,2]=mean(combmeans22)  
expected[3,2]=mean(combmeans32)
```

```
expected[4,2]=mean(combmeans42)
expected[5,2]=mean(combmeans52)
expected[6,2]=mean(combmeans62)
expected[7,2]=mean(combmeans72)
```

```
expected[1,3]=mean(combmeans13) #strategy 1 nsamp=5
expected[2,3]=mean(combmeans23)
expected[3,3]=mean(combmeans33)
expected[4,3]=mean(combmeans43)
expected[5,3]=mean(combmeans53)
expected[6,3]=mean(combmeans63)
expected[7,3]=mean(combmeans73)
```

```
expected[1,4]=mean(combmeans14) #strategy 1 nsamp=10
expected[2,4]=mean(combmeans24)
expected[3,4]=mean(combmeans34)
expected[4,4]=mean(combmeans44)
expected[5,4]=mean(combmeans54)
expected[6,4]=mean(combmeans64)
expected[7,4]=mean(combmeans74)
```

```
expected[1,5]=mean(combmeans15) #strategy 1 nsamp=2
expected[2,5]=mean(combmeans25)
expected[3,5]=mean(combmeans35)
expected[4,5]=mean(combmeans45)
expected[5,5]=mean(combmeans55)
expected[6,5]=mean(combmeans65)
expected[7,5]=mean(combmeans75)
```

```
expected[1,6]=mean(combmeans16) #strategy 1 nsamp=20
expected[2,6]=mean(combmeans26)
expected[3,6]=mean(combmeans36)
expected[4,6]=mean(combmeans46)
expected[5,6]=mean(combmeans56)
expected[6,6]=mean(combmeans66)
expected[7,6]=mean(combmeans76)
```

```
expectedcol=c(expected1,expected[,1],expected[,2],expected[,3],expected[,4],expected[,5],expected[,6])
```

```
means=cbind(combmeans11,combmeans21, combmeans31,combmeans41,combmeans51,combmeans61,combmeans71,
combmeans12,combmeans22, combmeans32,combmeans42,combmeans52,combmeans62,combmeans72,
combmeans13,combmeans23, combmeans33,combmeans43,combmeans53,combmeans63,combmeans73,
combmeans14,combmeans24, combmeans34,combmeans44,combmeans54,combmeans64,combmeans74,
combmeans15,combmeans25, combmeans35,combmeans45,combmeans55,combmeans65,combmeans75,
combmeans16,combmeans26, combmeans36,combmeans46,combmeans56,combmeans66,combmeans76)
```

```
variance1=(1/(nvar))*sum((sampmean[,1]-expected1)**2)
```

```
variance11=(1/(nvar-1))*sum((means[,1]-expected[1,1])**2)
variance21=(1/(nvar-1))*sum((means[,2]-expected[2,1])**2)
variance31=(1/(nvar-1))*sum((means[,3]-expected[3,1])**2)
variance41=(1/(nvar-1))*sum((means[,4]-expected[4,1])**2)
variance51=(1/(nvar-1))*sum((means[,5]-expected[5,1])**2)
variance61=(1/(nvar-1))*sum((means[,6]-expected[6,1])**2)
variance71=(1/(nvar-1))*sum((means[,7]-expected[7,1])**2)
```

```
variance12=(1/(nvar-1))*sum((means[,8]-expected[1,2])**2)
variance22=(1/(nvar-1))*sum((means[,9]-expected[2,2])**2)
variance32=(1/(nvar-1))*sum((means[,10]-expected[3,2])**2)
variance42=(1/(nvar-1))*sum((means[,11]-expected[4,2])**2)
variance52=(1/(nvar-1))*sum((means[,12]-expected[5,2])**2)
variance62=(1/(nvar-1))*sum((means[,13]-expected[6,2])**2)
variance72=(1/(nvar-1))*sum((means[,14]-expected[7,2])**2)
```

```
variance13=(1/(nvar-1))*sum((means[,15]-expected[1,3])**2)
variance23=(1/(nvar-1))*sum((means[,16]-expected[2,3])**2)
variance33=(1/(nvar-1))*sum((means[,17]-expected[3,3])**2)
variance43=(1/(nvar-1))*sum((means[,18]-expected[4,3])**2)
variance53=(1/(nvar-1))*sum((means[,19]-expected[5,3])**2)
variance63=(1/(nvar-1))*sum((means[,20]-expected[6,3])**2)
variance73=(1/(nvar-1))*sum((means[,21]-expected[7,3])**2)
```

```

variance14=(1/(nvar-1))*sum((means[,22]-expected[1,4])**2)
variance24=(1/(nvar-1))*sum((means[,23]-expected[2,4])**2)
variance34=(1/(nvar-1))*sum((means[,24]-expected[3,4])**2)
variance44=(1/(nvar-1))*sum((means[,25]-expected[4,4])**2)
variance54=(1/(nvar-1))*sum((means[,26]-expected[5,4])**2)
variance64=(1/(nvar-1))*sum((means[,27]-expected[6,4])**2)
variance74=(1/(nvar-1))*sum((means[,28]-expected[7,4])**2)

variance15=(1/(nvar-1))*sum((means[,29]-expected[1,5])**2)
variance25=(1/(nvar-1))*sum((means[,30]-expected[2,5])**2)
variance35=(1/(nvar-1))*sum((means[,31]-expected[3,5])**2)
variance45=(1/(nvar-1))*sum((means[,32]-expected[4,5])**2)
variance55=(1/(nvar-1))*sum((means[,33]-expected[5,5])**2)
variance65=(1/(nvar-1))*sum((means[,34]-expected[6,5])**2)
variance75=(1/(nvar-1))*sum((means[,35]-expected[7,5])**2)

variance16=(1/(nvar-1))*sum((means[,36]-expected[1,6])**2)
variance26=(1/(nvar-1))*sum((means[,37]-expected[2,6])**2)
variance36=(1/(nvar-1))*sum((means[,38]-expected[3,6])**2)
variance46=(1/(nvar-1))*sum((means[,39]-expected[4,6])**2)
variance56=(1/(nvar-1))*sum((means[,40]-expected[5,6])**2)
variance66=(1/(nvar-1))*sum((means[,41]-expected[6,6])**2)
variance76=(1/(nvar-1))*sum((means[,42]-expected[7,6])**2)

variance=c(variance1,variance11,variance21,variance31,variance41,variance51, variance61,variance71,
variance12,variance22,variance32,variance42,variance52, variance62,variance72,
variance13,variance23,variance33,variance43,variance53, variance63,variance73,
variance14,variance24,variance34,variance44,variance54, variance64,variance74,
variance15,variance25,variance35,variance45,variance55, variance65,variance75,
variance16,variance26,variance36,variance46,variance56, variance66,variance76)

bias= expectedcol-popmean

relbias=(bias/expectedcol)*100

mse= variance + bias**2

cv= (sqrt(variance)/expectedcol)*100

s.e= (sqrt(variance)/sqrt(nvar))

LB= expectedcol - 1.96*s.e
UB= expectedcol + 1.96*s.e

result=cbind(popmean,expectedcol,relbias,s.e,cv,UB,mse)

#write.table(result,"result_anormal_1000.txt",sep="\t")
write.table(result,"million1_anormal_1000.txt",sep="\t")
rm(list = ls())

```

E4 Skewed T Distribution, 3 degrees of freedom

```
#####
#Scenario 1: N=1000000 N(100,25) n=1000 D=nsamp=20
#####

dagostino.pearson.test <- function(x) {
  # from Zar (1999), implemented by Doug Scofield, scofield at bio.indiana.edu
  DNAME <- deparse(substitute(x))
  n <- length(x)
  x2 <- x * x
  x3 <- x * x2
  x4 <- x * x3
  # compute Z_g1
  k3 <- ((n*sum(x3)) - (3*sum(x)*sum(x2)) + (2*(sum(x)^3)/n))/((n-1)*(n-2))
  g1 <- k3 / sqrt(var(x)^3)
  sqrtb1 <- ((n - 2)*g1) / sqrt(n*(n - 1))
  A <- sqrtb1 * sqrt(((n + 1)*(n + 3)) / (6*(n - 2)))
  B <- (3*(n*n + 27*n - 70)*(n+1)*(n+3)) / ((n-2)*(n+5)*(n+7)*(n+9))
  C <- sqrt(2*(B - 1)) - 1
  D <- sqrt(C)
  E <- 1 / sqrt(log(D))
  F <- A / sqrt(2/(C - 1))
  Zg1 <- E * log(F + sqrt(F*F + 1))
  # compute Z_g2
  G <- (24*n*(n-2)*(n-3)) / (((n+1)^2)*(n+3)*(n+5))
  k4 <- (((n*n*n + n*n)*sum(x4)) - (4*(n*n + n)*sum(x3)*sum(x)) -
    (3*(n*n - n)*sum(x2)^2) + (12*n*sum(x2)*sum(x)^2) - (6*sum(x)^4)) /
    (n*(n-1)*(n-2)*(n-3))
  g2 <- k4 / var(x)^2
  H <- ((n-2)*(n-3)*abs(g2)) / ((n+1)*(n-1)*sqrt(G))
  J <- ((6*(n*n - 5*n + 2)) / ((n+7)*(n+9))) * sqrt((6*(n+3)*(n+5)) /
    (n*(n-2)*(n-3)))
  K <- 6 + (8/J)*(2/J + sqrt(1 + 4/(J*J)))
  L <- (1 - 2/K) / (1 + H*sqrt(2/(K-4)))
  Zg2 <- (1 - 2/(9*K) - (L^(1/3))) / (sqrt(2/(9*K)))
  K2 <- Zg1*Zg1 + Zg2*Zg2
  pk2 <- pchisq(K2, 2, lower.tail=FALSE)
  RVAL <- list(statistic = c(K2 = K2), p.value = pk2, method =
    "D'Agostino-Pearson normality test\n\nK2 is distributed as Chi-squared
    with df=2", alternative = "distribution is not normal", data.name =
    DNAME)
  class(RVAL) <- "htest"
  #return(RVAL)
  return(K2)
}

#####

setwd("C:\\Users\\User\\Dropbox\\Review\\simulations_2019\\srs\\at3")

library(moments)
library(sn)

N=1000000      #finite population size
n=1000         #sample size
nsamp=20       #number of samples
k=nvar=1000

set.seed(123456)

#####
#####Assymetric t distribution, 3 degrees of freedom

v=3

xi=-0.7891
omega=0.746
```



```

combmeans75 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 7

combmeans16 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
combmeans26 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
combmeans36 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans46 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans56 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5
combmeans66 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 6
combmeans76 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 7

#####

for (k in 1:nvar){

for (j in 1:nsamp){

w1<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1) #store weights
w2<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w3<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w4<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w5<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w6<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w7<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)

}

##### Sampling from finite population#####

for (i in 1:nsamp) {

s[,i]<-sample(x,n)

means[i]<-mean(s[,i]) #sample mean for ith sample
variance[i]<-sum((s[,i]-means[i])^2)/(n-1) #sample variance
designvar[i]<- (1-(n/N))*(variance[i])/n #design consistent variance
cv[i]<-abs(sqrt(designvar[i])/means[i]) #coefficient of variation
skewness[i]<-skewness(s[,i])
kurtosis[i]<-kurtosis(s[,i])
DP[i]<-dagostino.pearson.test(s[,i]) #dagostino pearson statistic

w1[i]<-1/n #inverse of the sample size
w2[i]<-1/(designvar[i]) #inverse of the variance of the estimator of the mean
w3[i]=1 #simple average weights
w4[i]=1/cv[i] #inverse of cv
w5[i]=abs(1/skewness[i]) #inverse skewness
w6[i]=1/kurtosis[i] #inverse kurtosis
w7[i]=1/DP[i] #inverse dagostino pearson k_squared

combmeans1[i]<-round(sum(means[1:i]*w1[1:i])/sum(w1[1:i]),digits=9) #combine means using inverse sample size
combmeans2[i]<-round(sum(means[1:i]*w2[1:i])/sum(w2[1:i]),digits=9) # combine using inverse variance
combmeans3[i]<-round(sum(means[1:i]*w3[1:i])/sum(w3[1:i]),digits=9) # combine using a weight of 1
combmeans4[i]<-round(sum(means[1:i]*w4[1:i])/sum(w4[1:i]),digits=9) # combine using a weight of n
combmeans5[i]<-round(sum(means[1:i]*w5[1:i])/sum(w5[1:i]),digits=9) # combine using a weight of inverse CV
combmeans6[i]<-round(sum(means[1:i]*w6[1:i])/sum(w6[1:i]),digits=9) # combine using a weight of inverse skewness
combmeans7[i]<-round(sum(means[1:i]*w7[1:i])/sum(w7[1:i]),digits=9) # combine using a weight of inverse kurtosis

}

#####Measures of performance of the different estimators#####

sampmean[k]=means[1] # capturing the mean of 1 sample, where nsamp=1 reps are combined

combmeans11[k]=combmeans1[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans21[k]=combmeans2[2]
combmeans31[k]=combmeans3[2]
combmeans41[k]=combmeans4[2]

```

```

combmeans51[k]=combmeans5[2]
combmeans61[k]=combmeans6[2]
combmeans71[k]=combmeans7[2]

combmeans12[k]=combmeans1[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans22[k]=combmeans2[3]
combmeans32[k]=combmeans3[3]
combmeans42[k]=combmeans4[3]
combmeans52[k]=combmeans5[3]
combmeans62[k]=combmeans6[3]
combmeans72[k]=combmeans7[3]

combmeans13[k]=combmeans1[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans23[k]=combmeans2[5]
combmeans33[k]=combmeans3[5]
combmeans43[k]=combmeans4[5]
combmeans53[k]=combmeans5[5]
combmeans63[k]=combmeans6[5]
combmeans73[k]=combmeans7[5]

combmeans14[k]=combmeans1[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans24[k]=combmeans2[10]
combmeans34[k]=combmeans3[10]
combmeans44[k]=combmeans4[10]
combmeans54[k]=combmeans5[10]
combmeans64[k]=combmeans6[10]
combmeans74[k]=combmeans7[10]

combmeans15[k]=combmeans1[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans25[k]=combmeans2[15]
combmeans35[k]=combmeans3[15]
combmeans45[k]=combmeans4[15]
combmeans55[k]=combmeans5[15]
combmeans65[k]=combmeans6[15]
combmeans75[k]=combmeans7[15]

combmeans16[k]=combmeans1[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans26[k]=combmeans2[20]
combmeans36[k]=combmeans3[20]
combmeans46[k]=combmeans4[20]
combmeans56[k]=combmeans5[20]
combmeans66[k]=combmeans6[20]
combmeans76[k]=combmeans7[20]

}

expected1=mean(sampmean) #expected mean for one sample, not combining
expected=matrix(0,7,6)

expected[1,1]=mean(combmeans11) #strategy 1 nsamp=2
expected[2,1]=mean(combmeans21)
expected[3,1]=mean(combmeans31)
expected[4,1]=mean(combmeans41)
expected[5,1]=mean(combmeans51)
expected[6,1]=mean(combmeans61)
expected[7,1]=mean(combmeans71)

expected[1,2]=mean(combmeans12) #strategy 1 nsamp=3
expected[2,2]=mean(combmeans22)
expected[3,2]=mean(combmeans32)
expected[4,2]=mean(combmeans42)
expected[5,2]=mean(combmeans52)
expected[6,2]=mean(combmeans62)
expected[7,2]=mean(combmeans72)

expected[1,3]=mean(combmeans13) #strategy 1 nsamp=5
expected[2,3]=mean(combmeans23)
expected[3,3]=mean(combmeans33)
expected[4,3]=mean(combmeans43)

```

```
expected[5,3]=mean(combmeans53)
expected[6,3]=mean(combmeans63)
expected[7,3]=mean(combmeans73)
```

```
expected[1,4]=mean(combmeans14) #strategy 1 nsamp=10
expected[2,4]=mean(combmeans24)
expected[3,4]=mean(combmeans34)
expected[4,4]=mean(combmeans44)
expected[5,4]=mean(combmeans54)
expected[6,4]=mean(combmeans64)
expected[7,4]=mean(combmeans74)
```

```
expected[1,5]=mean(combmeans15) #strategy 1 nsamp=2
expected[2,5]=mean(combmeans25)
expected[3,5]=mean(combmeans35)
expected[4,5]=mean(combmeans45)
expected[5,5]=mean(combmeans55)
expected[6,5]=mean(combmeans65)
expected[7,5]=mean(combmeans75)
```

```
expected[1,6]=mean(combmeans16) #strategy 1 nsamp=20
expected[2,6]=mean(combmeans26)
expected[3,6]=mean(combmeans36)
expected[4,6]=mean(combmeans46)
expected[5,6]=mean(combmeans56)
expected[6,6]=mean(combmeans66)
expected[7,6]=mean(combmeans76)
```

```
expectedcol=c(expected1,expected[,1],expected[,2],expected[,3],expected[,4],expected[,5],expected[,6])
```

```
means=cbind(combmeans11,combmeans21, combmeans31,combmeans41,combmeans51,combmeans61,combmeans71,
combmeans12,combmeans22, combmeans32,combmeans42,combmeans52,combmeans62,combmeans72,
combmeans13,combmeans23, combmeans33,combmeans43,combmeans53,combmeans63,combmeans73,
combmeans14,combmeans24, combmeans34,combmeans44,combmeans54,combmeans64,combmeans74,
combmeans15,combmeans25, combmeans35,combmeans45,combmeans55,combmeans65,combmeans75,
combmeans16,combmeans26, combmeans36,combmeans46,combmeans56,combmeans66,combmeans76)
```

```
variance1=(1/(nvar))*sum((sampmean[,1]-expected1)**2)
```

```
variance11=(1/(nvar-1))*sum((means[,1]-expected[1,1])**2)
variance21=(1/(nvar-1))*sum((means[,2]-expected[2,1])**2)
variance31=(1/(nvar-1))*sum((means[,3]-expected[3,1])**2)
variance41=(1/(nvar-1))*sum((means[,4]-expected[4,1])**2)
variance51=(1/(nvar-1))*sum((means[,5]-expected[5,1])**2)
variance61=(1/(nvar-1))*sum((means[,6]-expected[6,1])**2)
variance71=(1/(nvar-1))*sum((means[,7]-expected[7,1])**2)
```

```
variance12=(1/(nvar-1))*sum((means[,8]-expected[1,2])**2)
variance22=(1/(nvar-1))*sum((means[,9]-expected[2,2])**2)
variance32=(1/(nvar-1))*sum((means[,10]-expected[3,2])**2)
variance42=(1/(nvar-1))*sum((means[,11]-expected[4,2])**2)
variance52=(1/(nvar-1))*sum((means[,12]-expected[5,2])**2)
variance62=(1/(nvar-1))*sum((means[,13]-expected[6,2])**2)
variance72=(1/(nvar-1))*sum((means[,14]-expected[7,2])**2)
```

```
variance13=(1/(nvar-1))*sum((means[,15]-expected[1,3])**2)
variance23=(1/(nvar-1))*sum((means[,16]-expected[2,3])**2)
variance33=(1/(nvar-1))*sum((means[,17]-expected[3,3])**2)
variance43=(1/(nvar-1))*sum((means[,18]-expected[4,3])**2)
variance53=(1/(nvar-1))*sum((means[,19]-expected[5,3])**2)
variance63=(1/(nvar-1))*sum((means[,20]-expected[6,3])**2)
variance73=(1/(nvar-1))*sum((means[,21]-expected[7,3])**2)
```

```
variance14=(1/(nvar-1))*sum((means[,22]-expected[1,4])**2)
variance24=(1/(nvar-1))*sum((means[,23]-expected[2,4])**2)
variance34=(1/(nvar-1))*sum((means[,24]-expected[3,4])**2)
variance44=(1/(nvar-1))*sum((means[,25]-expected[4,4])**2)
variance54=(1/(nvar-1))*sum((means[,26]-expected[5,4])**2)
variance64=(1/(nvar-1))*sum((means[,27]-expected[6,4])**2)
variance74=(1/(nvar-1))*sum((means[,28]-expected[7,4])**2)
```

```

variance15=(1/(nvar-1))*sum((means[,29]-expected[1,5])**2)
variance25=(1/(nvar-1))*sum((means[,30]-expected[2,5])**2)
variance35=(1/(nvar-1))*sum((means[,31]-expected[3,5])**2)
variance45=(1/(nvar-1))*sum((means[,32]-expected[4,5])**2)
variance55=(1/(nvar-1))*sum((means[,33]-expected[5,5])**2)
variance65=(1/(nvar-1))*sum((means[,34]-expected[6,5])**2)
variance75=(1/(nvar-1))*sum((means[,35]-expected[7,5])**2)

variance16=(1/(nvar-1))*sum((means[,36]-expected[1,6])**2)
variance26=(1/(nvar-1))*sum((means[,37]-expected[2,6])**2)
variance36=(1/(nvar-1))*sum((means[,38]-expected[3,6])**2)
variance46=(1/(nvar-1))*sum((means[,39]-expected[4,6])**2)
variance56=(1/(nvar-1))*sum((means[,40]-expected[5,6])**2)
variance66=(1/(nvar-1))*sum((means[,41]-expected[6,6])**2)
variance76=(1/(nvar-1))*sum((means[,42]-expected[7,6])**2)

variance=c(variance1,variance11,variance21,variance31,variance41,variance51, variance61,variance71,
variance12,variance22,variance32,variance42,variance52, variance62,variance72,
variance13,variance23,variance33,variance43,variance53, variance63,variance73,
variance14,variance24,variance34,variance44,variance54, variance64,variance74,
variance15,variance25,variance35,variance45,variance55, variance65,variance75,
variance16,variance26,variance36,variance46,variance56, variance66,variance76)

bias= expectedcol-popmean

relbias=(bias/expectedcol)*100

mse= variance + bias**2

cv= (sqrt(variance)/expectedcol)*100

s.e= (sqrt(variance)/sqrt(nvar))

LB= expectedcol - 1.96*s.e
UB= expectedcol + 1.96*s.e

result=cbind(popmean,expectedcol,relbias,s.e,cv,LB,UB,mse)

write.table(result,"million_at3_1000.txt",sep="\t")

rm(list = ls())

```

Appendix F

R code used in simulations for simple random sampling, variance estimators

F1 Normal Distribution

```
#####  
# Variance Estimation  
# author: "Loveness Dzikiti"  
# date: "20 April 2017"  
#  
#Scenario 1: N=1000000 N(100,25) n=1000 D=nsamp=20  
#####  
  
setwd("C:\\Users\\User\\Dropbox\\Review\\simulations_2019\\variance\\srs\\normal\\scenario1")  
  
dagostino.pearson.test <- function(x) {  
  # from Zar (1999), implemented by Doug Scofield, scofield at bio.indiana.edu  
  DNAME <- deparse(substitute(x))  
  n <- length(x)  
  x2 <- x * x  
  x3 <- x * x2  
  x4 <- x * x3  
  # compute Z_g1  
  k3 <- ((n*sum(x3)) - (3*sum(x)*sum(x2)) + (2*(sum(x)^3)/n))/((n-1)*(n-2))  
  g1 <- k3 / sqrt(var(x)^3)  
  sqrtb1 <- ((n - 2)*g1) / sqrt(n*(n - 1))  
  A <- sqrtb1 * sqrt(((n + 1)*(n + 3)) / (6*(n - 2)))  
  B <- (3*(n*n + 27*n - 70)*(n+1)*(n+3)) / ((n-2)*(n+5)*(n+7)*(n+9))  
  C <- sqrt(2*(B - 1)) - 1  
  D <- sqrt(C)  
  E <- 1 / sqrt(log(D))  
  F <- A / sqrt(2/(C - 1))  
  Zg1 <- E * log(F + sqrt(F*F + 1))  
  # compute Z_g2  
  G <- (24*n*(n-2)*(n-3)) / (((n+1)^2)*(n+3)*(n+5))  
  k4 <- (((n*n*n + n*n)*sum(x4)) - (4*(n*n + n)*sum(x3)*sum(x)) -  
    (3*(n*n - n)*sum(x2)^2) + (12*n*sum(x2)*sum(x)^2) - (6*sum(x)^4)) /  
    (n*(n-1)*(n-2)*(n-3))  
  g2 <- k4 / var(x)^2  
  H <- ((n-2)*(n-3)*abs(g2)) / ((n+1)*(n-1)*sqrt(G))  
  J <- ((6*(n*n - 5*n + 2)) / ((n+7)*(n+9))) * sqrt((6*(n+3)*(n+5)) /  
    (n*(n-2)*(n-3)))  
  K <- 6 + (8/J)*(2/J + sqrt(1 + 4/(J*J)))  
  L <- (1 - 2/K) / (1 + H*sqrt(2/(K-4)))  
  Zg2 <- (1 - 2/(9*K) - (L^(1/3))) / (sqrt(2/(9*K)))  
  K2 <- Zg1*Zg1 + Zg2*Zg2  
  pk2 <- pchisq(K2, 2, lower.tail=FALSE)  
  RVAL <- list(statistic = c(K2 = K2), p.value = pk2, method =  
    "D'Agostino-Pearson normality test\n\nK2 is distributed as Chi-squared  
with df=2", alternative = "distribution is not normal", data.name =  
    DNAME)  
  class(RVAL) <- "htest"  
  #return(RVAL)  
  return(K2) # adapting the test to return K2 instead of the usual results stored in RVAL  
}  
  
library(moments)  
  
N=1000000 #finite population size  
n=1000 #sample size  
nsamp=20 #number of samples  
k=nvar=1000  
  
set.seed(123456)  
  
s <- matrix(c(rep.int(NA,n)), nrow = n, ncol = nsamp) #store samples  
means <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample means  
variance <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample variance
```

```

designvar <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store design consistent variance
design <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store design corrected variance
cv <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample coefficient of variation
skewness <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample skewness
kurtosis <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample skewness
DP <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample dagostino k_squared

```

```

combmeans1 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 1
combmeans2 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 2
combmeans3 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 3
combmeans4 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 4
combmeans5 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 5
combmeans6 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 6
combmeans7 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 7

```

```

varest1 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=1) #store means combined weighting strategy 1

```

```

varest2_1 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7) #varest, nsamp=2
varest2_2 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)
varest2_3 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)
varest2_4 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)

```

```

varest3_1 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7) #varest, nsamp=3
varest3_2 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)
varest3_3 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)
varest3_4 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)

```

```

varest5_1 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7) #varest, nsamp=5
varest5_2 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)
varest5_3 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)
varest5_4 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)

```

```

varest10_1 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7) #varest, nsamp=10
varest10_2 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)
varest10_3 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)
varest10_4 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)

```

```

varest15_1 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)
varest15_2 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)
varest15_3 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)
varest15_4 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)

```

```

varest20_1 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)
varest20_2 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)
varest20_3 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)
varest20_4 <- matrix(c(rep.int(NA,nvar)),nrow=nvar,ncol=7)

```

```

#####

```

```

x<-rnorm(N,100,5) #simulate data from normal distribution mean=100 sd=5

```

```

for (k in 1:nvar){

```

```

  for (j in 1:nsamp){

```

```

    w1<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1) #store weights
    w2<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
    w3<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
    w4<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
    w5<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
    w6<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
    w7<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)

```

```

  }

```

```

# Sampling from finite population

```

```

for (i in 1:nsamp) {

s[,i]<-sample(x,n)

means[i]<-mean(s[,i])           #sample mean for ith sample
#variance[i]<-sum((s[,i]-means[i])^2)/(n-1)   #sample variance
variance[i]<-var(s[,i])
designvar[i]<-(1-(n/N))*(variance[i])/n      #design consistent variance
design[i]<-(1-(n/N))*(variance[i])          #design correction expression 5.3
cv[i]<-abs(sqrt(designvar[i])/means[i])     #coefficient of variation
skewness[i]<-skewness(s[,i])
kurtosis[i]<-kurtosis(s[,i])
DP[i]<-dagostino.pearson.test(s[,i])       #dagostino pearson statistic

w1[i]<-n           #inverse of the sample size
w2[i]<-1/(designvar[i]) #inverse of the variance of the estimator of the mean
w3[i]=1           #simple average weights
w4[i]=1/cv[i]     #inverse of cv
w5[i]=abs(1/skewness[i]) #inverse skewness
w6[i]=1/kurtosis[i] #inverse kurtosis
w7[i]=1/DP[i]     #inverse dagostino pearson k_squared

combmeans1[i]<-round(sum(means[1:i]*w1[1:i])/sum(w1[1:i]),digits=9) #combine means using inverse sample size
combmeans2[i]<-round(sum(means[1:i]*w2[1:i])/sum(w2[1:i]),digits=9) # combine using inverse variance
combmeans3[i]<-round(sum(means[1:i]*w3[1:i])/sum(w3[1:i]),digits=9) # combine using a weight of 1
combmeans4[i]<-round(sum(means[1:i]*w4[1:i])/sum(w4[1:i]),digits=9) # combine using a weight of n
combmeans5[i]<-round(sum(means[1:i]*w5[1:i])/sum(w5[1:i]),digits=9) # combine using a weight of inverse CV
combmeans6[i]<-round(sum(means[1:i]*w6[1:i])/sum(w6[1:i]),digits=9) # combine using a weight of inverse skewness
combmeans7[i]<-round(sum(means[1:i]*w7[1:i])/sum(w7[1:i]),digits=9) # combine using a weight of inverse kurtosis

}

popmean=mean(x) #finite population mean

varest1[k]= designvar[1]

varest2_1[k,1]=(1/(2**2))*sum(designvar[1:2])
varest2_1[k,2]=(1/(2**2))*sum(designvar[1:2])
varest2_1[k,3]=(1/(2**2))*sum(designvar[1:2])
varest2_1[k,4]=(1/(2**2))*sum(designvar[1:2])
varest2_1[k,5]=(1/(2**2))*sum(designvar[1:2])
varest2_1[k,6]=(1/(2**2))*sum(designvar[1:2])
varest2_1[k,7]=(1/(2**2))*sum(designvar[1:2])

varest3_1[k,1]=(1/(3**2))*sum(designvar[1:3])
varest3_1[k,2]=(1/(3**2))*sum(designvar[1:3])
varest3_1[k,3]=(1/(3**2))*sum(designvar[1:3])
varest3_1[k,4]=(1/(3**2))*sum(designvar[1:3])
varest3_1[k,5]=(1/(3**2))*sum(designvar[1:3])
varest3_1[k,6]=(1/(3**2))*sum(designvar[1:3])
varest3_1[k,7]=(1/(3**2))*sum(designvar[1:3])

varest5_1[k,1]=(1/(5**2))*sum(designvar[1:5])
varest5_1[k,2]=(1/(5**2))*sum(designvar[1:5])
varest5_1[k,3]=(1/(5**2))*sum(designvar[1:5])
varest5_1[k,4]=(1/(5**2))*sum(designvar[1:5])
varest5_1[k,5]=(1/(5**2))*sum(designvar[1:5])
varest5_1[k,6]=(1/(5**2))*sum(designvar[1:5])
varest5_1[k,7]=(1/(5**2))*sum(designvar[1:5])

varest10_1[k,1]=(1/(10**2))*sum(designvar[1:10])
varest10_1[k,2]=(1/(10**2))*sum(designvar[1:10])
varest10_1[k,3]=(1/(10**2))*sum(designvar[1:10])
varest10_1[k,4]=(1/(10**2))*sum(designvar[1:10])
varest10_1[k,5]=(1/(10**2))*sum(designvar[1:10])
varest10_1[k,6]=(1/(10**2))*sum(designvar[1:10])
varest10_1[k,7]=(1/(10**2))*sum(designvar[1:10])

```

$varest15_1[k,1]=1/(15**2)*sum(designvar[1:15])$
 $varest15_1[k,2]=1/(15**2)*sum(designvar[1:15])$
 $varest15_1[k,3]=1/(15**2)*sum(designvar[1:15])$
 $varest15_1[k,4]=1/(15**2)*sum(designvar[1:15])$
 $varest15_1[k,5]=1/(15**2)*sum(designvar[1:15])$
 $varest15_1[k,6]=1/(15**2)*sum(designvar[1:15])$
 $varest15_1[k,7]=1/(15**2)*sum(designvar[1:15])$

$varest20_1[k,1]=1/(nsamp**2)*sum(designvar[1:nsamp])$
 $varest20_1[k,2]=1/(nsamp**2)*sum(designvar[1:nsamp])$
 $varest20_1[k,3]=1/(nsamp**2)*sum(designvar[1:nsamp])$
 $varest20_1[k,4]=1/(nsamp**2)*sum(designvar[1:nsamp])$
 $varest20_1[k,5]=1/(nsamp**2)*sum(designvar[1:nsamp])$
 $varest20_1[k,6]=1/(nsamp**2)*sum(designvar[1:nsamp])$
 $varest20_1[k,7]=1/(nsamp**2)*sum(designvar[1:nsamp])$

$varest2_2[k,1]=1/sum(w1[1:2])$
 $varest2_2[k,2]=1/sum(w2[1:2])$
 $varest2_2[k,3]=1/sum(w3[1:2])$
 $varest2_2[k,4]=1/sum(w4[1:2])$
 $varest2_2[k,5]=1/sum(w5[1:2])$
 $varest2_2[k,6]=1/sum(w6[1:2])$
 $varest2_2[k,7]=1/sum(w7[1:2])$

$varest3_2[k,1]=1/sum(w1[1:3])$
 $varest3_2[k,2]=1/sum(w2[1:3])$
 $varest3_2[k,3]=1/sum(w3[1:3])$
 $varest3_2[k,4]=1/sum(w4[1:3])$
 $varest3_2[k,5]=1/sum(w5[1:3])$
 $varest3_2[k,6]=1/sum(w6[1:3])$
 $varest3_2[k,7]=1/sum(w7[1:3])$

$varest5_2[k,1]=1/sum(w1[1:5])$
 $varest5_2[k,2]=1/sum(w2[1:5])$
 $varest5_2[k,3]=1/sum(w3[1:5])$
 $varest5_2[k,4]=1/sum(w4[1:5])$
 $varest5_2[k,5]=1/sum(w5[1:5])$
 $varest5_2[k,6]=1/sum(w6[1:5])$
 $varest5_2[k,7]=1/sum(w7[1:5])$

$varest10_2[k,1]=1/sum(w1[1:10])$
 $varest10_2[k,2]=1/sum(w2[1:10])$
 $varest10_2[k,3]=1/sum(w3[1:10])$
 $varest10_2[k,4]=1/sum(w4[1:10])$
 $varest10_2[k,5]=1/sum(w5[1:10])$
 $varest10_2[k,6]=1/sum(w6[1:10])$
 $varest10_2[k,7]=1/sum(w7[1:10])$

$varest15_2[k,1]=1/sum(w1[1:15])$
 $varest15_2[k,2]=1/sum(w2[1:15])$
 $varest15_2[k,3]=1/sum(w3[1:15])$
 $varest15_2[k,4]=1/sum(w4[1:15])$
 $varest15_2[k,5]=1/sum(w5[1:15])$
 $varest15_2[k,6]=1/sum(w6[1:15])$
 $varest15_2[k,7]=1/sum(w7[1:15])$

$varest20_2[k,1]=1/sum(w1[1:nsamp])$
 $varest20_2[k,2]=1/sum(w2[1:nsamp])$
 $varest20_2[k,3]=1/sum(w3[1:nsamp])$
 $varest20_2[k,4]=1/sum(w4[1:nsamp])$
 $varest20_2[k,5]=1/sum(w5[1:nsamp])$
 $varest20_2[k,6]=1/sum(w6[1:nsamp])$
 $varest20_2[k,7]=1/sum(w7[1:nsamp])$

$varest2_3[k,1]=1/(2-1)*sum(w1[1:2]*(combmeans1[1:2]-means[1:2])**2)/sum(w1[1:2])$
 $varest2_3[k,2]=1/(2-1)*sum(w2[1:2]*(combmeans2[1:2]-means[1:2])**2)/sum(w2[1:2])$
 $varest2_3[k,3]=1/(2-1)*sum(w3[1:2]*(combmeans3[1:2]-means[1:2])**2)/sum(w3[1:2])$
 $varest2_3[k,4]=1/(2-1)*sum(w4[1:2]*(combmeans4[1:2]-means[1:2])**2)/sum(w4[1:2])$
 $varest2_3[k,5]=1/(2-1)*sum(w5[1:2]*(combmeans5[1:2]-means[1:2])**2)/sum(w5[1:2])$
 $varest2_3[k,6]=1/(2-1)*sum(w6[1:2]*(combmeans6[1:2]-means[1:2])**2)/sum(w6[1:2])$


```

varest10_4[k,3]=(1/(sum(w3[1:10])**2)*sum(((w3[1:10])**2)*(combmeans3[1:10]-means[1:10])**2)
varest10_4[k,4]=(1/(sum(w4[1:10])**2)*sum(((w4[1:10])**2)*(combmeans4[1:10]-means[1:10])**2)
varest10_4[k,5]=(1/(sum(w5[1:10])**2)*sum(((w5[1:10])**2)*(combmeans5[1:10]-means[1:10])**2)
varest10_4[k,6]=(1/(sum(w6[1:10])**2)*sum(((w6[1:10])**2)*(combmeans6[1:10]-means[1:10])**2)
varest10_4[k,7]=(1/(sum(w7[1:10])**2)*sum(((w7[1:10])**2)*(combmeans7[1:10]-means[1:10])**2)

```

```

varest15_4[k,1]=(1/(sum(w1[1:15])**2)*sum(((w1[1:15])**2)*(combmeans1[1:15]-means[1:15])**2)
varest15_4[k,2]=(1/(sum(w2[1:15])**2)*sum(((w2[1:15])**2)*(combmeans2[1:15]-means[1:15])**2)
varest15_4[k,3]=(1/(sum(w3[1:15])**2)*sum(((w3[1:15])**2)*(combmeans3[1:15]-means[1:15])**2)
varest15_4[k,4]=(1/(sum(w4[1:15])**2)*sum(((w4[1:15])**2)*(combmeans4[1:15]-means[1:15])**2)
varest15_4[k,5]=(1/(sum(w5[1:15])**2)*sum(((w5[1:15])**2)*(combmeans5[1:15]-means[1:15])**2)
varest15_4[k,6]=(1/(sum(w6[1:15])**2)*sum(((w6[1:15])**2)*(combmeans6[1:15]-means[1:15])**2)
varest15_4[k,7]=(1/(sum(w7[1:15])**2)*sum(((w7[1:15])**2)*(combmeans7[1:15]-means[1:15])**2)

```

```

varest20_4[k,1]=(1/(sum(w1[1:20])**2)*sum(((w1[1:20])**2)*(combmeans1[1:20]-means[1:20])**2)
varest20_4[k,2]=(1/(sum(w2[1:20])**2)*sum(((w2[1:20])**2)*(combmeans2[1:20]-means[1:20])**2)
varest20_4[k,3]=(1/(sum(w3[1:20])**2)*sum(((w3[1:20])**2)*(combmeans3[1:20]-means[1:20])**2)
varest20_4[k,4]=(1/(sum(w4[1:20])**2)*sum(((w4[1:20])**2)*(combmeans4[1:20]-means[1:20])**2)
varest20_4[k,5]=(1/(sum(w5[1:20])**2)*sum(((w5[1:20])**2)*(combmeans5[1:20]-means[1:20])**2)
varest20_4[k,6]=(1/(sum(w6[1:20])**2)*sum(((w6[1:20])**2)*(combmeans6[1:20]-means[1:20])**2)
varest20_4[k,7]=(1/(sum(w7[1:20])**2)*sum(((w7[1:20])**2)*(combmeans7[1:20]-means[1:20])**2)

```

```

}

```

```

varest2_1=cbind(varest1, varest2_1) # to include results from 1 sample

```

```

write.table(varest2_1,"varest_1_scenario1_2.txt",sep="\t")
write.table(varest3_1,"varest_1_scenario1_3.txt",sep="\t")
write.table(varest5_1,"varest_1_scenario1_5.txt",sep="\t")
write.table(varest10_1,"varest_1_scenario1_10.txt",sep="\t")
write.table(varest15_1,"varest_1_scenario1_15.txt",sep="\t")
write.table(varest20_1,"varest_1_scenario1_20.txt",sep="\t")

```

```

write.table(varest2_2,"varest_2_scenario1_2.txt",sep="\t")
write.table(varest3_2,"varest_2_scenario1_3.txt",sep="\t")
write.table(varest5_2,"varest_2_scenario1_5.txt",sep="\t")
write.table(varest10_2,"varest_2_scenario1_10.txt",sep="\t")
write.table(varest15_2,"varest_2_scenario1_15.txt",sep="\t")
write.table(varest20_2,"varest_2_scenario1_20.txt",sep="\t")

```

```

write.table(varest2_3,"varest_3_scenario1_2.txt",sep="\t")
write.table(varest3_3,"varest_3_scenario1_3.txt",sep="\t")
write.table(varest5_3,"varest_3_scenario1_5.txt",sep="\t")
write.table(varest10_3,"varest_3_scenario1_10.txt",sep="\t")
write.table(varest15_3,"varest_3_scenario1_15.txt",sep="\t")
write.table(varest20_3,"varest_3_scenario1_20.txt",sep="\t")

```

```

write.table(varest2_4,"varest_4_scenario1_2.txt",sep="\t")
write.table(varest3_4,"varest_4_scenario1_3.txt",sep="\t")
write.table(varest5_4,"varest_4_scenario1_5.txt",sep="\t")
write.table(varest10_4,"varest_4_scenario1_10.txt",sep="\t")
write.table(varest15_4,"varest_4_scenario1_15.txt",sep="\t")
write.table(varest20_4,"varest_4_scenario1_20.txt",sep="\t")

```

Appendix G

R code used in simulations for stratified random sampling, point estimators

G1 Normal Distribution

```
#####
# Title: Stratified Fox inspired simulations_Normal distribution
# author: "Loveness Dzikiti"
# date: "22 September 2017"
=====
#
#Scenario 1: N=1000000 N(100,25) Nstrata=200000 n=1000 D=nsamp=20
#####

setwd("C:\\Users\\User\\Dropbox\\Review\\simulations_2019\\stratified\\normal")

#install.packages("survey")

library(moments)
library(sampling)
library(survey)

N=1000000      #finite population size
L=5           # number of strata in population
Nstrata=N/L
nsamp=20      #number of samples or replicates
k=nvar=1000

set.seed(123456)

#####
##### Generating a stratified population

x<-c(rnorm(Nstrata,105,4),rnorm(Nstrata,103,3),rnorm(Nstrata,100,3),rnorm(Nstrata,97,3),rnorm(Nstrata,95,4)) #simulate stratified
data from normal distribution

strata_id<-c(rep('1',Nstrata),rep('2',Nstrata),rep('3',Nstrata),rep('4',Nstrata),rep('5',Nstrata))
popx<-data.frame(1:1000000,strata_id,x)
colnames(popx)<-c('ID','group','x')

pop.des <- svydesign(id=~1,strata=~group,data=popx)
pop.mean=data.frame(svymean(x,pop.des))
popmean <- pop.mean[1,1]#finite population mean

#####

mean_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample means calculated as srs
variance_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample variance calculated as srs
cv_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample coefficient of variation calculated as srs
skewness_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample skewness calculated as srs
kurtosis_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample kurtosis calculated as srs

means <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample means
variance <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample variance
designvar <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store design consistent variance
cv <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample coefficient of variation
skewness <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample skewness
kurtosis <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample kurtosis
meff_var<-matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample design effect

combmeans1 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 1
combmeans2 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 2
combmeans3 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 3
combmeans4 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 4
combmeans5 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 5
combmeans6 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 6
```



```

combmeans65 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 6
combmeans75 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 7
combmeans85 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
combmeans95 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
combmeans105 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans115 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans125 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5

```

```

combmeans16 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
combmeans26 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
combmeans36 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans46 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans56 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5
combmeans66 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 6
combmeans76 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 7
combmeans86 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
combmeans96 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
combmeans106 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans116 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans126 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5

```

```
#####
```

```
for (k in 1:nvar){
```

```
for (j in 1:nsamp){
```

```

w1<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1) #store weights
w2<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w3<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w4<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w5<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w6<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w7<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w8<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1) #store weights
w9<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w10<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w11<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w12<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
}

```

```
#####
# Stratified random Sampling from a stratified finite population
#####
```

```

#Calculating sample sizes for strata
L=5 #number of strata
propn<-c(0.00144,0.00072,0.00072,0.00072,0.00144)
Nstrata<-N/L #strata size for equal strata sizes in population

```

```

samps1<-ceiling(Nstrata*propn)
n<-samps1[1]+samps1[2]+samps1[3]+samps1[4]+samps1[5]

```

```
fpc=rep(c(Nstrata,Nstrata,Nstrata,Nstrata,Nstrata),samps1)
```

```
for (i in 1:nsamp) {
```

```

s1<-sample(popx$x[1:200000],samps1[1])
s2<-sample(popx$x[200001:400000],samps1[2])
s3<-sample(popx$x[400001:600000],samps1[3])
s4<-sample(popx$x[600001:800000],samps1[4])
s5<-sample(popx$x[800001:1000000],samps1[5])

```

```
sampstrat<-c(rep('1',samps1[1]),rep('2',samps1[2]),rep('3',samps1[3]),rep('4',samps1[4]),rep('5',samps1[5]))
```

```

s<-data.frame(sampstrat,c(s1,s2,s3,s4,s5))
colnames(s)<-c('sampstrat','x')

#sample1=strata(popx, stratanames=c("group"), size=sampsi, method=c("srswor"), description=FALSE)
#sample=getdata(popx,sample1)

#####
# Naive weights
#####

mean_naive[i]=mean(s$x)
variance_naive[i]=var(s$x)
cv_naive[i]=abs(sqrt(variance_naive[i])/mean_naive[i])
skewness_naive[i]<-skewness(s$x)
kurtosis_naive[i]<-kurtosis(s$x)

#####
# Design conscious weights
#####

plan=svydesign(~1,strata=~sampstrat,data=s,fpc=~fpc) #probs=~probn is not the cluster samp. prob.
mean_st=svymean(s$x,plan)
mean_st=as.data.frame(mean_st)

means[i]<-mean_st[1,1]          #sample mean for ith sample
variance[i]<-mean_st[1,2]^2     #sample variance, design consistent variance
cv[i]<-abs(sqrt(variance[i])/means[i]) #coefficient of variation

# skewness[i]<-skewness(s[,i])
moments = svymean(~l(x^4) + l(x^3) +
                 l(x^2) + l(x), plan)

skew=data.frame(svycontrast(moments,
                             quote(`l(x^3)` - 3*`l(x^2)`*`l(x)` +
                                     3*`l(x)`*`l(x)^2` - `l(x)^3`)/
                                     (`l(x^2)` - `l(x)^2`^1.5)))
skewness[i]<-skew[1,1]

#kurtosis[i]<-kurtosis(s[,i])

kurt=data.frame(svycontrast(moments,
                             quote(`l(x^4)` -
                                     4*`l(x^3)`*`l(x)` +
                                     6*`l(x^2)`*`l(x)^2` -
                                     4*`l(x)`*`l(x)^3` +
                                     `l(x)^4`/
                                     (`l(x^2)` -
                                     `l(x)^2`^2)))
kurtosis[i]<-kurt[1,1]

meff.1 <- data.frame(svymean(s$x,plan,deff=TRUE))
meff_var[i]=variance[i]/((1-(n/N))*(var(s$x)/n))

w1[i]=1/n          #inverse of the sample size

w2[i]<-1/(variance[i]) #inverse of the design variance of the estimator of the mean
w3[i]<-1/(variance_naive[i]) #inverse of the naive variance of the estimator of the mean

w4[i]=1          #simple average weights

w5[i]=1/cv[i]    #inverse of cv
w6[i]=1/cv_naive[i] #inverse of naive cv

w7[i]=abs(1/skewness[i]) #inverse skewness
w8[i]=abs(1/skewness_naive[i]) #inverse skewness

```

```

w9[i]=1/kurtosis[i] #inverse kurtosis
w10[i]=1/kurtosis_naive[i] #inverse kurtosis

w11[i]=meff_var[i]/variance[i] #in srs deff=1 so coincides with inverse of variance
w12[i]=1/meff_var[i]

combmeans1[i]<-round(sum(means[1:i]*w1[1:i])/sum(w1[1:i]),digits=9) #combine means using inverse sample size
combmeans2[i]<-round(sum(means[1:i]*w2[1:i])/sum(w2[1:i]),digits=9) # combine using inverse variance
combmeans3[i]<-round(sum(means[1:i]*w3[1:i])/sum(w3[1:i]),digits=9) # combine using inverse variance naive
combmeans4[i]<-round(sum(means[1:i]*w4[1:i])/sum(w4[1:i]),digits=9) # combine using a weight of 1
combmeans5[i]<-round(sum(means[1:i]*w5[1:i])/sum(w5[1:i]),digits=9) # combine using a weight of inverse CV
combmeans6[i]<-round(sum(means[1:i]*w6[1:i])/sum(w6[1:i]),digits=9) # combine using a weight of inverse CV naive
combmeans7[i]<-round(sum(means[1:i]*w7[1:i])/sum(w7[1:i]),digits=9) # combine using a weight of inverse skewness
combmeans8[i]<-round(sum(means[1:i]*w8[1:i])/sum(w8[1:i]),digits=9) #combine using a weight of inverse skewness naive
combmeans9[i]<-round(sum(means[1:i]*w9[1:i])/sum(w9[1:i]),digits=9) # combine using inverse kurtosis
combmeans10[i]<-round(sum(means[1:i]*w10[1:i])/sum(w10[1:i]),digits=9) # combine using inverse kurtosis naive
combmeans11[i]<-round(sum(means[1:i]*w11[1:i])/sum(w11[1:i]),digits=9) # combine using a weight of meff/var
combmeans12[i]<-round(sum(means[1:i]*w12[1:i])/sum(w12[1:i]),digits=9) # combine using a weight of 1/meff

}

sampmean[k]=means[1] # capturing the combined mean, where nsamp reps are combined, all weighting strategies

combmeans11[k]=combmeans1[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans21[k]=combmeans2[2]
combmeans31[k]=combmeans3[2]
combmeans41[k]=combmeans4[2]
combmeans51[k]=combmeans5[2]
combmeans61[k]=combmeans6[2]
combmeans71[k]=combmeans7[2]
combmeans81[k]=combmeans8[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans91[k]=combmeans9[2]
combmeans101[k]=combmeans10[2]
combmeans111[k]=combmeans11[2]
combmeans121[k]=combmeans12[2]

combmeans12[k]=combmeans1[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans22[k]=combmeans2[3]
combmeans32[k]=combmeans3[3]
combmeans42[k]=combmeans4[3]
combmeans52[k]=combmeans5[3]
combmeans62[k]=combmeans6[3]
combmeans72[k]=combmeans7[3]
combmeans82[k]=combmeans8[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans92[k]=combmeans9[3]
combmeans102[k]=combmeans10[3]
combmeans112[k]=combmeans11[3]
combmeans122[k]=combmeans12[3]

combmeans13[k]=combmeans1[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans23[k]=combmeans2[5]
combmeans33[k]=combmeans3[5]
combmeans43[k]=combmeans4[5]
combmeans53[k]=combmeans5[5]
combmeans63[k]=combmeans6[5]
combmeans73[k]=combmeans7[5]
combmeans83[k]=combmeans8[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans93[k]=combmeans9[5]
combmeans103[k]=combmeans10[5]
combmeans113[k]=combmeans11[5]
combmeans123[k]=combmeans12[5]

combmeans14[k]=combmeans1[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans24[k]=combmeans2[10]

```

```

combmeans34[k]=combmeans3[10]
combmeans44[k]=combmeans4[10]
combmeans54[k]=combmeans5[10]
combmeans64[k]=combmeans6[10]
combmeans74[k]=combmeans7[10]
combmeans84[k]=combmeans8[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans94[k]=combmeans9[10]
combmeans104[k]=combmeans10[10]
combmeans114[k]=combmeans11[10]
combmeans124[k]=combmeans12[10]

combmeans15[k]=combmeans1[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans25[k]=combmeans2[15]
combmeans35[k]=combmeans3[15]
combmeans45[k]=combmeans4[15]
combmeans55[k]=combmeans5[15]
combmeans65[k]=combmeans6[15]
combmeans75[k]=combmeans7[15]
combmeans85[k]=combmeans8[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans95[k]=combmeans9[15]
combmeans105[k]=combmeans10[15]
combmeans115[k]=combmeans11[15]
combmeans125[k]=combmeans12[15]

combmeans16[k]=combmeans1[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans26[k]=combmeans2[20]
combmeans36[k]=combmeans3[20]
combmeans46[k]=combmeans4[20]
combmeans56[k]=combmeans5[20]
combmeans66[k]=combmeans6[20]
combmeans76[k]=combmeans7[20]
combmeans86[k]=combmeans8[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans96[k]=combmeans9[20]
combmeans106[k]=combmeans10[20]
combmeans116[k]=combmeans11[20]
combmeans126[k]=combmeans12[20]

}

expected1=mean(sampmean) #expected mean for one sample, not combining

expected=matrix(0,12,6)

expected[1,1]=mean(combmeans11) #strategy 1 nsamp=2
expected[2,1]=mean(combmeans21)
expected[3,1]=mean(combmeans31)
expected[4,1]=mean(combmeans41)
expected[5,1]=mean(combmeans51)
expected[6,1]=mean(combmeans61)
expected[7,1]=mean(combmeans71)
expected[8,1]=mean(combmeans81)
expected[9,1]=mean(combmeans91)
expected[10,1]=mean(combmeans101)
expected[11,1]=mean(combmeans111)
expected[12,1]=mean(combmeans121)

expected[1,2]=mean(combmeans12) #strategy 1 nsamp=3
expected[2,2]=mean(combmeans22)
expected[3,2]=mean(combmeans32)
expected[4,2]=mean(combmeans42)
expected[5,2]=mean(combmeans52)
expected[6,2]=mean(combmeans62)
expected[7,2]=mean(combmeans72)
expected[8,2]=mean(combmeans82)
expected[9,2]=mean(combmeans92)
expected[10,2]=mean(combmeans102)
expected[11,2]=mean(combmeans112)

```

```

expected[12,2]=mean(combmeans122)

expected[1,3]=mean(combmeans13) #strategy 1 nsamp=5
expected[2,3]=mean(combmeans23)
expected[3,3]=mean(combmeans33)
expected[4,3]=mean(combmeans43)
expected[5,3]=mean(combmeans53)
expected[6,3]=mean(combmeans63)
expected[7,3]=mean(combmeans73)
expected[8,3]=mean(combmeans83)
expected[9,3]=mean(combmeans93)
expected[10,3]=mean(combmeans103)
expected[11,3]=mean(combmeans113)
expected[12,3]=mean(combmeans123)

expected[1,4]=mean(combmeans14) #strategy 1 nsamp=10
expected[2,4]=mean(combmeans24)
expected[3,4]=mean(combmeans34)
expected[4,4]=mean(combmeans44)
expected[5,4]=mean(combmeans54)
expected[6,4]=mean(combmeans64)
expected[7,4]=mean(combmeans74)
expected[8,4]=mean(combmeans84)
expected[9,4]=mean(combmeans94)
expected[10,4]=mean(combmeans104)
expected[11,4]=mean(combmeans114)
expected[12,4]=mean(combmeans124)

expected[1,5]=mean(combmeans15) #strategy 1 nsamp=15
expected[2,5]=mean(combmeans25)
expected[3,5]=mean(combmeans35)
expected[4,5]=mean(combmeans45)
expected[5,5]=mean(combmeans55)
expected[6,5]=mean(combmeans65)
expected[7,5]=mean(combmeans75)
expected[8,5]=mean(combmeans85)
expected[9,5]=mean(combmeans95)
expected[10,5]=mean(combmeans105)
expected[11,5]=mean(combmeans115)
expected[12,5]=mean(combmeans125)

expected[1,6]=mean(combmeans16) #strategy 1 nsamp=20
expected[2,6]=mean(combmeans26)
expected[3,6]=mean(combmeans36)
expected[4,6]=mean(combmeans46)
expected[5,6]=mean(combmeans56)
expected[6,6]=mean(combmeans66)
expected[7,6]=mean(combmeans76)
expected[8,6]=mean(combmeans86)
expected[9,6]=mean(combmeans96)
expected[10,6]=mean(combmeans106)
expected[11,6]=mean(combmeans116)
expected[12,6]=mean(combmeans126)

expectedcol=c(expected1,expected[,1],expected[,2],expected[,3],expected[,4],expected[,5],expected[,6])

means=cbind(combmeans11,combmeans21,
combmeans31,combmeans41,combmeans51,combmeans61,combmeans71,combmeans81,combmeans91,
combmeans101,combmeans111,combmeans121,
combmeans12,combmeans22, combmeans32,combmeans42,combmeans52,combmeans62,combmeans72,combmeans82,combmeans92,
combmeans102,combmeans112,combmeans122,
combmeans13,combmeans23, combmeans33,combmeans43,combmeans53,combmeans63,combmeans73,combmeans83,combmeans93,
combmeans103,combmeans113,combmeans123,
combmeans14,combmeans24, combmeans34,combmeans44,combmeans54,combmeans64,combmeans74,combmeans84,combmeans94,
combmeans104,combmeans114,combmeans124,
combmeans15,combmeans25, combmeans35,combmeans45,combmeans55,combmeans65,combmeans75,combmeans85,combmeans95,
combmeans105,combmeans115,combmeans125,
combmeans16,combmeans26, combmeans36,combmeans46,combmeans56,combmeans66,combmeans76,combmeans86,combmeans96,
combmeans106,combmeans116,combmeans126)

```



```

variance56=(1/nvar)*sum((means[,65]-expected[5,6])**2)
variance66=(1/nvar)*sum((means[,66]-expected[6,6])**2)
variance76=(1/nvar)*sum((means[,67]-expected[7,6])**2)
variance86=(1/nvar)*sum((means[,68]-expected[8,6])**2)
variance96=(1/nvar)*sum((means[,69]-expected[9,6])**2)
variance106=(1/nvar)*sum((means[,70]-expected[10,6])**2)
variance116=(1/nvar)*sum((means[,71]-expected[11,6])**2)
variance126=(1/nvar)*sum((means[,72]-expected[12,6])**2)

```

```

variance=c(variance1,variance11,variance21,variance31,variance41,variance51,
variance61,variance71,variance81,variance91,variance101,variance111,variance121,
variance12,variance22,variance32,variance42,variance52,
variance62,variance72,variance82,variance92,variance102,variance112,variance122,
variance13,variance23,variance33,variance43,variance53,
variance63,variance73,variance83,variance93,variance103,variance113,variance123,
variance14,variance24,variance34,variance44,variance54,
variance64,variance74,variance84,variance94,variance104,variance114,variance124,
variance15,variance25,variance35,variance45,variance55,
variance65,variance75,variance85,variance95,variance105,variance115,variance125,
variance16,variance26,variance36,variance46,variance56,
variance66,variance76,variance86,variance96,variance106,variance116,variance126)

```

```

bias= expectedcol-popmean

```

```

relbias=(bias/expectedcol)*100

```

```

mse= variance + bias**2

```

```

cv= (sqrt(variance)/expectedcol)*100

```

```

s.e= (sqrt(variance)/sqrt(nvar))

```

```

LB= expectedcol - 1.96*s.e

```

```

UB= expectedcol + 1.96*s.e

```

```

result=cbind(popmean,expectedcol,relbias,s.e,cv,UB,mse)

```

```

write.table(result,"result_normal_scenario1_1000.txt",sep="\t")

```

```

write.table(meff_var,"meff_normal_scenario1_1000.txt",sep="\t")

```

G2 T Distribution, 3 degrees of freedom

```
#####
# Title: Stratified Fox inspired simulations_Normal distribution
# author: "Loveness Dzikiti"
# date: "22 September 2017"
#####
#####
#
#Scenario 1: N=100000 t3(100,25) Nstrata=200000 n=1000 D=nsamp=20
#####

setwd("C:\\Users\\User\\Dropbox\\Review\\simulations_2019\\stratified\\t3")
#install.packages("survey")

library(moments)
library(sampling)
library(survey)

N=1000000      #finite population size
L=5            # number of strata in population
Nstrata=N/L
nsamp=20       #number of samples or replicates
k=nvar=1000

set.seed(123456)

#####
# Generating a stratified population
v=3
r=2

mu1=105
sigma1=4

mu2=103
sigma2=3

mu3=100
sigma3=3

mu4=97
sigma4=3

mu5=95
sigma5=4

Var<-v^(r/2)*((gamma(1/2*(r+1))*gamma(1/2*(v-r)))/(gamma(0.5)*gamma(1/2*v))) # variance of t-distribution

x<-rt(N,v)      #simulate data from a t distribution random mean and variance

x<-x/sqrt(Var)

x[1:Nstrata]=x[1:Nstrata]*sigma1+mu1
x[(Nstrata+1):(2*Nstrata)]=x[(Nstrata+1):(2*Nstrata)]*sigma2+mu2
x[(2*Nstrata+1):(3*Nstrata)]=x[(2*Nstrata+1):(3*Nstrata)]*sigma3+mu3
x[(3*Nstrata+1):(4*Nstrata)]=x[(3*Nstrata+1):(4*Nstrata)]*sigma4+mu4
x[(4*Nstrata+1):(5*Nstrata)]=x[(4*Nstrata+1):(5*Nstrata)]*sigma5+mu5

#####

strata_id<-c(rep('1',Nstrata),rep('2',Nstrata),rep('3',Nstrata),rep('4',Nstrata),rep('5',Nstrata))
popx<-data.frame(1:100000,strata_id,x)
colnames(popx)<-c('ID','group','x')

pop.des <- svydesign(id=~1,strata=~group,data=popx)
pop.mean=data.frame(svymean(x,pop.des))
popmean <- pop.mean[1,1]#finite population mean
```



```

#####Calculating sample sizes for strata
L=5 #number of strata
propn<-c(0.00144,0.00072,0.00072,0.00072,0.00144)
Nstrata<-N/L #strata size for equal strata sizes in population

samps<-ceiling(Nstrata*propn)
n<-samps[1]+samps[2]+samps[3]+samps[4]+samps[5]

fpc=rep(c(Nstrata,Nstrata,Nstrata,Nstrata,Nstrata),samps)

for (i in 1:nsamp) {

s1<-sample(popx$x[1:200000],samps[1])
s2<-sample(popx$x[200001:400000],samps[2])
s3<-sample(popx$x[400001:600000],samps[3])
s4<-sample(popx$x[600001:800000],samps[4])
s5<-sample(popx$x[800001:1000000],samps[5])

sampstrat<-c(rep('1',samps[1]),rep('2',samps[2]),rep('3',samps[3]),rep('4',samps[4]),rep('5',samps[5]))

s<-data.frame(sampstrat,c(s1,s2,s3,s4,s5))
colnames(s)<-c('sampstrat','x')

#####
# Naive weights
#####

mean_naive[i]=mean(s$x)
variance_naive[i]=var(s$x)
cv_naive[i]=abs(sqrt(variance_naive[i])/mean_naive[i])
skewness_naive[i]<-skewness(s$x)
kurtosis_naive[i]<-kurtosis(s$x)

#####
# Design conscious weights
#####

plan=svydesign(~1,strata=~sampstrat,data=s,fpc=~fpc) #probs=~propn is not the cluster samp. prob.
mean_st=svymean(s$x,plan)
mean_st=as.data.frame(mean_st)

means[i]<-mean_st[1,1] #sample mean for ith sample
variance[i]<-mean_st[1,2]^2 #sample variance, design consistent variance
cv[i]<-abs(sqrt(variance[i])/means[i]) #coefficient of variation

# skewness[i]<-skewness(s[,i])
moments = svymean(~l(x^4) + l(x^3) +
l(x^2) + l(x), plan)

skew=data.frame(svycontrast(moments,
quote('l(x^3)' - 3*l(x^2)*l(x)' +
3*l(x)*l(x)^2 - l(x)^3)/
('l(x^2)' - l(x)^2)^1.5)))
skewness[i]<-skew[1,1]

#kurtosis[i]<-kurtosis(s[,i])

kurt=data.frame(svycontrast(moments,
quote('l(x^4)' -
4*l(x^3)*l(x)' +
6*l(x^2)*l(x)^2 -
4*l(x)*l(x)^3 +
l(x)^4)/
('l(x^2)' -
l(x)^2)^2)))

kurtosis[i]<-kurt[1,1]

```

```

meff.1 <- data.frame(svymean(s$x,plan,deff=TRUE))
meff_var[i]=variance[i]/((1-(n/N))*(var(s$x)/n))

w1[i]=1/n      #inverse of the sample size

w2[i]<-1/(variance[i]) #inverse of the design variance of the estimator of the mean
w3[i]<-1/(variance_naive[i]) #inverse of the naive variance of the estimator of the mean

w4[i]=1      #simple average weights

w5[i]=1/cv[i] #inverse of cv
w6[i]=1/cv_naive[i] #inverse of naive cv

w7[i]=abs(1/skewness[i]) #inverse skewness
w8[i]=abs(1/skewness_naive[i]) #inverse skewness

w9[i]=1/kurtosis[i] #inverse kurtosis
w10[i]=1/kurtosis_naive[i] #inverse kurtosis

w11[i]=meff_var[i]/variance[i] #in srs deff=1 so coincides with inverse of variance
w12[i]=1/meff_var[i]

combmeans1[i]<-round(sum(means[1:i]*w1[1:i])/sum(w1[1:i]),digits=9) #combine means using inverse sample size
combmeans2[i]<-round(sum(means[1:i]*w2[1:i])/sum(w2[1:i]),digits=9) # combine using inverse variance
combmeans3[i]<-round(sum(means[1:i]*w3[1:i])/sum(w3[1:i]),digits=9) # combine using inverse variance naive
combmeans4[i]<-round(sum(means[1:i]*w4[1:i])/sum(w4[1:i]),digits=9) # combine using a weight of 1
combmeans5[i]<-round(sum(means[1:i]*w5[1:i])/sum(w5[1:i]),digits=9) # combine using a weight of inverse CV
combmeans6[i]<-round(sum(means[1:i]*w6[1:i])/sum(w6[1:i]),digits=9) # combine using a weight of inverse CV naive
combmeans7[i]<-round(sum(means[1:i]*w7[1:i])/sum(w7[1:i]),digits=9) # combine using a weight of inverse skewness
combmeans8[i]<-round(sum(means[1:i]*w8[1:i])/sum(w8[1:i]),digits=9) #combine using a weight of inverse skewness naive
combmeans9[i]<-round(sum(means[1:i]*w9[1:i])/sum(w9[1:i]),digits=9) # combine using inverse kurtosis
combmeans10[i]<-round(sum(means[1:i]*w10[1:i])/sum(w10[1:i]),digits=9) # combine using inverse kurtosis naive
combmeans11[i]<-round(sum(means[1:i]*w11[1:i])/sum(w11[1:i]),digits=9) # combine using a weight of meff/var
combmeans12[i]<-round(sum(means[1:i]*w12[1:i])/sum(w12[1:i]),digits=9) # combine using a weight of 1/meff

}

sampmean[k]=means[1] # capturing the combined mean, where nsamp reps are combined, all weighting strategies

combmeans11[k]=combmeans1[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans21[k]=combmeans2[2]
combmeans31[k]=combmeans3[2]
combmeans41[k]=combmeans4[2]
combmeans51[k]=combmeans5[2]
combmeans61[k]=combmeans6[2]
combmeans71[k]=combmeans7[2]
combmeans81[k]=combmeans8[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans91[k]=combmeans9[2]
combmeans101[k]=combmeans10[2]
combmeans111[k]=combmeans11[2]
combmeans121[k]=combmeans12[2]

combmeans12[k]=combmeans1[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans22[k]=combmeans2[3]
combmeans32[k]=combmeans3[3]
combmeans42[k]=combmeans4[3]
combmeans52[k]=combmeans5[3]
combmeans62[k]=combmeans6[3]
combmeans72[k]=combmeans7[3]
combmeans82[k]=combmeans8[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans92[k]=combmeans9[3]
combmeans102[k]=combmeans10[3]
combmeans112[k]=combmeans11[3]
combmeans122[k]=combmeans12[3]

```

```

combmeans13[k]=combmeans1[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans23[k]=combmeans2[5]
combmeans33[k]=combmeans3[5]
combmeans43[k]=combmeans4[5]
combmeans53[k]=combmeans5[5]
combmeans63[k]=combmeans6[5]
combmeans73[k]=combmeans7[5]
combmeans83[k]=combmeans8[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans93[k]=combmeans9[5]
combmeans103[k]=combmeans10[5]
combmeans113[k]=combmeans11[5]
combmeans123[k]=combmeans12[5]

combmeans14[k]=combmeans1[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans24[k]=combmeans2[10]
combmeans34[k]=combmeans3[10]
combmeans44[k]=combmeans4[10]
combmeans54[k]=combmeans5[10]
combmeans64[k]=combmeans6[10]
combmeans74[k]=combmeans7[10]
combmeans84[k]=combmeans8[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans94[k]=combmeans9[10]
combmeans104[k]=combmeans10[10]
combmeans114[k]=combmeans11[10]
combmeans124[k]=combmeans12[10]

combmeans15[k]=combmeans1[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans25[k]=combmeans2[15]
combmeans35[k]=combmeans3[15]
combmeans45[k]=combmeans4[15]
combmeans55[k]=combmeans5[15]
combmeans65[k]=combmeans6[15]
combmeans75[k]=combmeans7[15]
combmeans85[k]=combmeans8[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans95[k]=combmeans9[15]
combmeans105[k]=combmeans10[15]
combmeans115[k]=combmeans11[15]
combmeans125[k]=combmeans12[15]

combmeans16[k]=combmeans1[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans26[k]=combmeans2[20]
combmeans36[k]=combmeans3[20]
combmeans46[k]=combmeans4[20]
combmeans56[k]=combmeans5[20]
combmeans66[k]=combmeans6[20]
combmeans76[k]=combmeans7[20]
combmeans86[k]=combmeans8[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans96[k]=combmeans9[20]
combmeans106[k]=combmeans10[20]
combmeans116[k]=combmeans11[20]
combmeans126[k]=combmeans12[20]

}

expected1=mean(sampmean) #expected mean for one sample, not combining

expected=matrix(0,12,6)

expected[1,1]=mean(combmeans11) #strategy 1 nsamp=2
expected[2,1]=mean(combmeans21)
expected[3,1]=mean(combmeans31)
expected[4,1]=mean(combmeans41)
expected[5,1]=mean(combmeans51)
expected[6,1]=mean(combmeans61)

```

```

expected[7,1]=mean(combmeans71)
expected[8,1]=mean(combmeans81)
expected[9,1]=mean(combmeans91)
expected[10,1]=mean(combmeans101)
expected[11,1]=mean(combmeans111)
expected[12,1]=mean(combmeans121)

expected[1,2]=mean(combmeans12) #strategy 1 nsamp=3
expected[2,2]=mean(combmeans22)
expected[3,2]=mean(combmeans32)
expected[4,2]=mean(combmeans42)
expected[5,2]=mean(combmeans52)
expected[6,2]=mean(combmeans62)
expected[7,2]=mean(combmeans72)
expected[8,2]=mean(combmeans82)
expected[9,2]=mean(combmeans92)
expected[10,2]=mean(combmeans102)
expected[11,2]=mean(combmeans112)
expected[12,2]=mean(combmeans122)

expected[1,3]=mean(combmeans13) #strategy 1 nsamp=5
expected[2,3]=mean(combmeans23)
expected[3,3]=mean(combmeans33)
expected[4,3]=mean(combmeans43)
expected[5,3]=mean(combmeans53)
expected[6,3]=mean(combmeans63)
expected[7,3]=mean(combmeans73)
expected[8,3]=mean(combmeans83)
expected[9,3]=mean(combmeans93)
expected[10,3]=mean(combmeans103)
expected[11,3]=mean(combmeans113)
expected[12,3]=mean(combmeans123)

expected[1,4]=mean(combmeans14) #strategy 1 nsamp=10
expected[2,4]=mean(combmeans24)
expected[3,4]=mean(combmeans34)
expected[4,4]=mean(combmeans44)
expected[5,4]=mean(combmeans54)
expected[6,4]=mean(combmeans64)
expected[7,4]=mean(combmeans74)
expected[8,4]=mean(combmeans84)
expected[9,4]=mean(combmeans94)
expected[10,4]=mean(combmeans104)
expected[11,4]=mean(combmeans114)
expected[12,4]=mean(combmeans124)

expected[1,5]=mean(combmeans15) #strategy 1 nsamp=15
expected[2,5]=mean(combmeans25)
expected[3,5]=mean(combmeans35)
expected[4,5]=mean(combmeans45)
expected[5,5]=mean(combmeans55)
expected[6,5]=mean(combmeans65)
expected[7,5]=mean(combmeans75)
expected[8,5]=mean(combmeans85)
expected[9,5]=mean(combmeans95)
expected[10,5]=mean(combmeans105)
expected[11,5]=mean(combmeans115)
expected[12,5]=mean(combmeans125)

expected[1,6]=mean(combmeans16) #strategy 1 nsamp=20
expected[2,6]=mean(combmeans26)
expected[3,6]=mean(combmeans36)
expected[4,6]=mean(combmeans46)
expected[5,6]=mean(combmeans56)
expected[6,6]=mean(combmeans66)
expected[7,6]=mean(combmeans76)
expected[8,6]=mean(combmeans86)
expected[9,6]=mean(combmeans96)
expected[10,6]=mean(combmeans106)
expected[11,6]=mean(combmeans116)
expected[12,6]=mean(combmeans126)

```

expectedcol=c(expected1,expected[,1],expected[,2],expected[,3],expected[,4],expected[,5],expected[,6])

means=cbind(combmeans11,combmeans21,
combmeans31,combmeans41,combmeans51,combmeans61,combmeans71,combmeans81,combmeans91,
combmeans101,combmeans111,combmeans121,
combmeans12,combmeans22, combmeans32,combmeans42,combmeans52,combmeans62,combmeans72,combmeans82,combmeans92,
combmeans102,combmeans112,combmeans122,
combmeans13,combmeans23, combmeans33,combmeans43,combmeans53,combmeans63,combmeans73,combmeans83,combmeans93,
combmeans103,combmeans113,combmeans123,
combmeans14,combmeans24, combmeans34,combmeans44,combmeans54,combmeans64,combmeans74,combmeans84,combmeans94,
combmeans104,combmeans114,combmeans124,
combmeans15,combmeans25, combmeans35,combmeans45,combmeans55,combmeans65,combmeans75,combmeans85,combmeans95,
combmeans105,combmeans115,combmeans125,
combmeans16,combmeans26, combmeans36,combmeans46,combmeans56,combmeans66,combmeans76,combmeans86,combmeans96,
combmeans106,combmeans116,combmeans126)

variance1=(1/nvar)*sum((sampmean[,1]-expected1)**2)

variance11=(1/nvar)*sum((means[,1]-expected[1,1])**2)
variance21=(1/nvar)*sum((means[,2]-expected[2,1])**2)
variance31=(1/nvar)*sum((means[,3]-expected[3,1])**2)
variance41=(1/nvar)*sum((means[,4]-expected[4,1])**2)
variance51=(1/nvar)*sum((means[,5]-expected[5,1])**2)
variance61=(1/nvar)*sum((means[,6]-expected[6,1])**2)
variance71=(1/nvar)*sum((means[,7]-expected[7,1])**2)
variance81=(1/nvar)*sum((means[,8]-expected[8,1])**2)
variance91=(1/nvar)*sum((means[,9]-expected[9,1])**2)
variance101=(1/nvar)*sum((means[,10]-expected[10,1])**2)
variance111=(1/nvar)*sum((means[,11]-expected[11,1])**2)
variance121=(1/nvar)*sum((means[,12]-expected[12,1])**2)

variance12=(1/nvar)*sum((means[,13]-expected[1,2])**2)
variance22=(1/nvar)*sum((means[,14]-expected[2,2])**2)
variance32=(1/nvar)*sum((means[,15]-expected[3,2])**2)
variance42=(1/nvar)*sum((means[,16]-expected[4,2])**2)
variance52=(1/nvar)*sum((means[,17]-expected[5,2])**2)
variance62=(1/nvar)*sum((means[,18]-expected[6,2])**2)
variance72=(1/nvar)*sum((means[,19]-expected[7,2])**2)
variance82=(1/nvar)*sum((means[,20]-expected[8,2])**2)
variance92=(1/nvar)*sum((means[,21]-expected[9,2])**2)
variance102=(1/nvar)*sum((means[,22]-expected[10,2])**2)
variance112=(1/nvar)*sum((means[,23]-expected[11,2])**2)
variance122=(1/nvar)*sum((means[,24]-expected[12,2])**2)

variance13=(1/nvar)*sum((means[,25]-expected[1,3])**2)
variance23=(1/nvar)*sum((means[,26]-expected[2,3])**2)
variance33=(1/nvar)*sum((means[,27]-expected[3,3])**2)
variance43=(1/nvar)*sum((means[,28]-expected[4,3])**2)
variance53=(1/nvar)*sum((means[,29]-expected[5,3])**2)
variance63=(1/nvar)*sum((means[,30]-expected[6,3])**2)
variance73=(1/nvar)*sum((means[,31]-expected[7,3])**2)
variance83=(1/nvar)*sum((means[,32]-expected[8,3])**2)
variance93=(1/nvar)*sum((means[,33]-expected[9,3])**2)
variance103=(1/nvar)*sum((means[,34]-expected[10,3])**2)
variance113=(1/nvar)*sum((means[,35]-expected[11,3])**2)
variance123=(1/nvar)*sum((means[,36]-expected[12,3])**2)

variance14=(1/nvar)*sum((means[,37]-expected[1,4])**2)
variance24=(1/nvar)*sum((means[,38]-expected[2,4])**2)
variance34=(1/nvar)*sum((means[,39]-expected[3,4])**2)
variance44=(1/nvar)*sum((means[,40]-expected[4,4])**2)
variance54=(1/nvar)*sum((means[,41]-expected[5,4])**2)
variance64=(1/nvar)*sum((means[,42]-expected[6,4])**2)
variance74=(1/nvar)*sum((means[,43]-expected[7,4])**2)
variance84=(1/nvar)*sum((means[,44]-expected[8,4])**2)
variance94=(1/nvar)*sum((means[,45]-expected[9,4])**2)
variance104=(1/nvar)*sum((means[,46]-expected[10,4])**2)
variance114=(1/nvar)*sum((means[,47]-expected[11,4])**2)
variance124=(1/nvar)*sum((means[,48]-expected[12,4])**2)

```

variance15=(1/nvar)*sum((means[,49]-expected[1,5])**2)
variance25=(1/nvar)*sum((means[,50]-expected[2,5])**2)
variance35=(1/nvar)*sum((means[,51]-expected[3,5])**2)
variance45=(1/nvar)*sum((means[,52]-expected[4,5])**2)
variance55=(1/nvar)*sum((means[,53]-expected[5,5])**2)
variance65=(1/nvar)*sum((means[,54]-expected[6,5])**2)
variance75=(1/nvar)*sum((means[,55]-expected[7,5])**2)
variance85=(1/nvar)*sum((means[,56]-expected[8,5])**2)
variance95=(1/nvar)*sum((means[,57]-expected[9,5])**2)
variance105=(1/nvar)*sum((means[,58]-expected[10,5])**2)
variance115=(1/nvar)*sum((means[,59]-expected[11,5])**2)
variance125=(1/nvar)*sum((means[,60]-expected[12,5])**2)

```

```

variance16=(1/nvar)*sum((means[,61]-expected[1,6])**2)
variance26=(1/nvar)*sum((means[,62]-expected[2,6])**2)
variance36=(1/nvar)*sum((means[,63]-expected[3,6])**2)
variance46=(1/nvar)*sum((means[,64]-expected[4,6])**2)
variance56=(1/nvar)*sum((means[,65]-expected[5,6])**2)
variance66=(1/nvar)*sum((means[,66]-expected[6,6])**2)
variance76=(1/nvar)*sum((means[,67]-expected[7,6])**2)
variance86=(1/nvar)*sum((means[,68]-expected[8,6])**2)
variance96=(1/nvar)*sum((means[,69]-expected[9,6])**2)
variance106=(1/nvar)*sum((means[,70]-expected[10,6])**2)
variance116=(1/nvar)*sum((means[,71]-expected[11,6])**2)
variance126=(1/nvar)*sum((means[,72]-expected[12,6])**2)

```

```

variance=c(variance1,variance11,variance21,variance31,variance41,variance51,
variance61,variance71,variance81,variance91,variance101,variance111,variance121,
variance12,variance22,variance32,variance42,variance52,
variance62,variance72,variance82,variance92,variance102,variance112,variance122,
variance13,variance23,variance33,variance43,variance53,
variance63,variance73,variance83,variance93,variance103,variance113,variance123,
variance14,variance24,variance34,variance44,variance54,
variance64,variance74,variance84,variance94,variance104,variance114,variance124,
variance15,variance25,variance35,variance45,variance55,
variance65,variance75,variance85,variance95,variance105,variance115,variance125,
variance16,variance26,variance36,variance46,variance56,
variance66,variance76,variance86,variance96,variance106,variance116,variance126)

```

```
bias= expectedcol-popmean
```

```
relbias=(bias/expectedcol)*100
```

```
mse= variance + bias**2
```

```
cv= (sqrt(variance)/expectedcol)*100
```

```
s.e= (sqrt(variance)/sqrt(nvar))
```

```
LB= expectedcol - 1.96*s.e
```

```
UB= expectedcol + 1.96*s.e
```

```
result=cbind(popmean,expectedcol,relbias,s.e,cv,LB,UB,mse)
```

```
write.table(result,"result_t3_scenario1.txt",sep="\t")
```

```
write.table(meff_var,"meff_t3_scenario1.txt",sep="\t")
```

G3 Skewed Normal Distribution

```
#####
# Title: Fox inspired simulations_Normal distribution
# author: "Loveness Dzikiti"
# date: "19 September 2017"
#####
#Scenario 1: N=1000000 SN(100,25) n=1000 D=nsamp=20
#####

setwd("C:\\Users\\User\\Dropbox\\Review\\simulations_2019\\stratified\\anormal")

library(moments)
library(sampling)
library(survey)
library(sn)

N=1000000      #finite population size
L=5            # number of strata in population
Nstrata=N/L
nsamp=20      #number of samples
k=nvar=1000

set.seed(123456)

#####
#####
##### Generating Asymmetric Normal Distribution      #

alpha=3        # asymmetry parameter for strong asymmetry

xi=-1.15      #
omega=1.541    #

x <- rsn(N, xi, omega, alpha)

x1=x[1:200000]
x2=x[200001:400000]
x3=x[400001:600000]
x4=x[600001:800000]
x5=x[800001:1000000]

mu1=105
sigma1=4

mu2=103
sigma2=3

mu3=100
sigma3=3

mu4=97
sigma4=3

mu5=95
sigma5=4

x=c((x1*sigma1)+mu1, (x2*sigma2)+mu2, (x3*sigma3)+mu3, (x4*sigma4)+mu4, (x5*sigma5)+mu5)
strata_id<-c(rep('1',Nstrata),rep('2',Nstrata),rep('3',Nstrata),rep('4',Nstrata),rep('5',Nstrata))
popx<-data.frame(1:100000, strata_id,x)
colnames(popx)<-c('ID','group','x')

pop.des <- svydesign(id=~1,strata=~group,data=popx)
pop.mean=data.frame(svymean(x,pop.des))
popmean <- pop.mean[1,1]#finite population mean
```



```

#####Calculating sample sizes for strata
L=5 #number of strata
propn<-c(0.00144,0.00072,0.00072,0.00072,0.00144)
Nstrata<-N/L #strata size for equal strata sizes in population

sampsiz<-ceiling(Nstrata*propn)
n<-sampsiz[1]+sampsiz[2]+sampsiz[3]+sampsiz[4]+sampsiz[5]

fpc=rep(c(Nstrata,Nstrata,Nstrata,Nstrata,Nstrata),sampsiz)

for (i in 1:nsamp) {

s1<-sample(popx$sx[1:200000],sampsiz[1])
s2<-sample(popx$sx[200001:400000],sampsiz[2])
s3<-sample(popx$sx[400001:600000],sampsiz[3])
s4<-sample(popx$sx[600001:800000],sampsiz[4])
s5<-sample(popx$sx[800001:1000000],sampsiz[5])

sampstrat<-c(rep('1',sampsiz[1]),rep('2',sampsiz[2]),rep('3',sampsiz[3]),rep('4',sampsiz[4]),rep('5',sampsiz[5]))

s<-data.frame(sampstrat,c(s1,s2,s3,s4,s5))
colnames(s)<-c('sampstrat','x')

#####
# Naive weights
#####

mean_naive[i]=mean(s$x)
variance_naive[i]=var(s$x)
cv_naive[i]=abs(sqrt(variance_naive[i])/mean_naive[i])
skewness_naive[i]<-skewness(s$x)
kurtosis_naive[i]<-kurtosis(s$x)

#####
# Design conscious weights
#####

plan=svydesign(~1,strata=~sampstrat,data=s,fpc=~fpc) #probs=~propn is not the cluster samp. prob.
mean_st=svymean(s$x,plan)
mean_st=as.data.frame(mean_st)

means[i]<-mean_st[1,1] #sample mean for ith sample
variance[i]<-mean_st[1,2]^2 #sample variance, design consistent variance
cv[i]<-abs(sqrt(variance[i])/means[i]) #coefficient of variation

# Design consistent skewness
moments = svymean(~l(x^4) + l(x^3) +
l(x^2) + l(x), plan)

skew=data.frame(svycontrast(moments,
quote('l(x^3)' - 3*l(x^2)*l(x)' +
3*l(x)*l(x)^2 - l(x)^3)/
('l(x^2)' - l(x)^2)^1.5)))
skewness[i]<-skew[1,1]

# Design consistent

kurt=data.frame(svycontrast(moments,
quote('l(x^4)' -
4*l(x^3)*l(x)' +
6*l(x^2)*l(x)^2 -
4*l(x)*l(x)^3 +
l(x)^4)/
('l(x^2)' -
l(x)^2)^2)))

kurtosis[i]<-kurt[1,1]

```

```

meff.1 <- data.frame(svymean(s$x,plan,deff=TRUE))
meff_var[i]=variance[i]/((1-(n/N))*(var(s$x)/n))

w1[i]=1/n      #inverse of the sample size

w2[i]<-1/(variance[i]) #inverse of the design variance of the estimator of the mean
w3[i]<-1/(variance_naive[i]) #inverse of the naive variance of the estimator of the mean

w4[i]=1      #simple average weights

w5[i]=1/cv[i] #inverse of cv
w6[i]=1/cv_naive[i] #inverse of naive cv

w7[i]=abs(1/skewness[i]) #inverse skewness
w8[i]=abs(1/skewness_naive[i]) #inverse skewness

w9[i]=1/kurtosis[i] #inverse kurtosis
w10[i]=1/kurtosis_naive[i] #inverse kurtosis

w11[i]=meff_var[i]/variance[i] #in srs deff=1 so coincides with inverse of variance
w12[i]=1/meff_var[i]

combmeans1[i]<-round(sum(means[1:i]*w1[1:i])/sum(w1[1:i]),digits=9) #combine means using inverse sample size
combmeans2[i]<-round(sum(means[1:i]*w2[1:i])/sum(w2[1:i]),digits=9) # combine using inverse variance
combmeans3[i]<-round(sum(means[1:i]*w3[1:i])/sum(w3[1:i]),digits=9) # combine using inverse variance naive
combmeans4[i]<-round(sum(means[1:i]*w4[1:i])/sum(w4[1:i]),digits=9) # combine using a weight of 1
combmeans5[i]<-round(sum(means[1:i]*w5[1:i])/sum(w5[1:i]),digits=9) # combine using a weight of inverse CV
combmeans6[i]<-round(sum(means[1:i]*w6[1:i])/sum(w6[1:i]),digits=9) # combine using a weight of inverse CV naive
combmeans7[i]<-round(sum(means[1:i]*w7[1:i])/sum(w7[1:i]),digits=9) # combine using a weight of inverse skewness
combmeans8[i]<-round(sum(means[1:i]*w8[1:i])/sum(w8[1:i]),digits=9) #combine using a weight of inverse skewness naive
combmeans9[i]<-round(sum(means[1:i]*w9[1:i])/sum(w9[1:i]),digits=9) # combine using inverse kurtosis
combmeans10[i]<-round(sum(means[1:i]*w10[1:i])/sum(w10[1:i]),digits=9) # combine using inverse kurtosis naive
combmeans11[i]<-round(sum(means[1:i]*w11[1:i])/sum(w11[1:i]),digits=9) # combine using a weight of meff/var
combmeans12[i]<-round(sum(means[1:i]*w12[1:i])/sum(w12[1:i]),digits=9) # combine using a weight of 1/meff

}

sampmean[k]=means[1] # capturing the combined mean, where nsamp reps are combined, all weighting strategies

combmeans11[k]=combmeans1[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans21[k]=combmeans2[2]
combmeans31[k]=combmeans3[2]
combmeans41[k]=combmeans4[2]
combmeans51[k]=combmeans5[2]
combmeans61[k]=combmeans6[2]
combmeans71[k]=combmeans7[2]
combmeans81[k]=combmeans8[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans91[k]=combmeans9[2]
combmeans101[k]=combmeans10[2]
combmeans111[k]=combmeans11[2]
combmeans121[k]=combmeans12[2]

combmeans12[k]=combmeans1[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans22[k]=combmeans2[3]
combmeans32[k]=combmeans3[3]
combmeans42[k]=combmeans4[3]
combmeans52[k]=combmeans5[3]
combmeans62[k]=combmeans6[3]
combmeans72[k]=combmeans7[3]
combmeans82[k]=combmeans8[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans92[k]=combmeans9[3]
combmeans102[k]=combmeans10[3]
combmeans112[k]=combmeans11[3]
combmeans122[k]=combmeans12[3]

```

```

combmeans13[k]=combmeans1[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans23[k]=combmeans2[5]
combmeans33[k]=combmeans3[5]
combmeans43[k]=combmeans4[5]
combmeans53[k]=combmeans5[5]
combmeans63[k]=combmeans6[5]
combmeans73[k]=combmeans7[5]
combmeans83[k]=combmeans8[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans93[k]=combmeans9[5]
combmeans103[k]=combmeans10[5]
combmeans113[k]=combmeans11[5]
combmeans123[k]=combmeans12[5]

combmeans14[k]=combmeans1[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans24[k]=combmeans2[10]
combmeans34[k]=combmeans3[10]
combmeans44[k]=combmeans4[10]
combmeans54[k]=combmeans5[10]
combmeans64[k]=combmeans6[10]
combmeans74[k]=combmeans7[10]
combmeans84[k]=combmeans8[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans94[k]=combmeans9[10]
combmeans104[k]=combmeans10[10]
combmeans114[k]=combmeans11[10]
combmeans124[k]=combmeans12[10]

combmeans15[k]=combmeans1[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans25[k]=combmeans2[15]
combmeans35[k]=combmeans3[15]
combmeans45[k]=combmeans4[15]
combmeans55[k]=combmeans5[15]
combmeans65[k]=combmeans6[15]
combmeans75[k]=combmeans7[15]
combmeans85[k]=combmeans8[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans95[k]=combmeans9[15]
combmeans105[k]=combmeans10[15]
combmeans115[k]=combmeans11[15]
combmeans125[k]=combmeans12[15]

combmeans16[k]=combmeans1[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans26[k]=combmeans2[20]
combmeans36[k]=combmeans3[20]
combmeans46[k]=combmeans4[20]
combmeans56[k]=combmeans5[20]
combmeans66[k]=combmeans6[20]
combmeans76[k]=combmeans7[20]
combmeans86[k]=combmeans8[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans96[k]=combmeans9[20]
combmeans106[k]=combmeans10[20]
combmeans116[k]=combmeans11[20]
combmeans126[k]=combmeans12[20]

}

expected1=mean(sampmean) #expected mean for one sample, not combining

expected=matrix(0,12,6)

expected[1,1]=mean(combmeans11) #strategy 1 nsamp=2
expected[2,1]=mean(combmeans21)
expected[3,1]=mean(combmeans31)
expected[4,1]=mean(combmeans41)
expected[5,1]=mean(combmeans51)
expected[6,1]=mean(combmeans61)

```

```

expected[7,1]=mean(combmeans71)
expected[8,1]=mean(combmeans81)
expected[9,1]=mean(combmeans91)
expected[10,1]=mean(combmeans101)
expected[11,1]=mean(combmeans111)
expected[12,1]=mean(combmeans121)

expected[1,2]=mean(combmeans12) #strategy 1 nsamp=3
expected[2,2]=mean(combmeans22)
expected[3,2]=mean(combmeans32)
expected[4,2]=mean(combmeans42)
expected[5,2]=mean(combmeans52)
expected[6,2]=mean(combmeans62)
expected[7,2]=mean(combmeans72)
expected[8,2]=mean(combmeans82)
expected[9,2]=mean(combmeans92)
expected[10,2]=mean(combmeans102)
expected[11,2]=mean(combmeans112)
expected[12,2]=mean(combmeans122)

expected[1,3]=mean(combmeans13) #strategy 1 nsamp=5
expected[2,3]=mean(combmeans23)
expected[3,3]=mean(combmeans33)
expected[4,3]=mean(combmeans43)
expected[5,3]=mean(combmeans53)
expected[6,3]=mean(combmeans63)
expected[7,3]=mean(combmeans73)
expected[8,3]=mean(combmeans83)
expected[9,3]=mean(combmeans93)
expected[10,3]=mean(combmeans103)
expected[11,3]=mean(combmeans113)
expected[12,3]=mean(combmeans123)

expected[1,4]=mean(combmeans14) #strategy 1 nsamp=10
expected[2,4]=mean(combmeans24)
expected[3,4]=mean(combmeans34)
expected[4,4]=mean(combmeans44)
expected[5,4]=mean(combmeans54)
expected[6,4]=mean(combmeans64)
expected[7,4]=mean(combmeans74)
expected[8,4]=mean(combmeans84)
expected[9,4]=mean(combmeans94)
expected[10,4]=mean(combmeans104)
expected[11,4]=mean(combmeans114)
expected[12,4]=mean(combmeans124)

expected[1,5]=mean(combmeans15) #strategy 1 nsamp=15
expected[2,5]=mean(combmeans25)
expected[3,5]=mean(combmeans35)
expected[4,5]=mean(combmeans45)
expected[5,5]=mean(combmeans55)
expected[6,5]=mean(combmeans65)
expected[7,5]=mean(combmeans75)
expected[8,5]=mean(combmeans85)
expected[9,5]=mean(combmeans95)
expected[10,5]=mean(combmeans105)
expected[11,5]=mean(combmeans115)
expected[12,5]=mean(combmeans125)

expected[1,6]=mean(combmeans16) #strategy 1 nsamp=20
expected[2,6]=mean(combmeans26)
expected[3,6]=mean(combmeans36)
expected[4,6]=mean(combmeans46)
expected[5,6]=mean(combmeans56)
expected[6,6]=mean(combmeans66)
expected[7,6]=mean(combmeans76)
expected[8,6]=mean(combmeans86)
expected[9,6]=mean(combmeans96)
expected[10,6]=mean(combmeans106)
expected[11,6]=mean(combmeans116)
expected[12,6]=mean(combmeans126)

```

expectedcol=c(expected1,expected[,1],expected[,2],expected[,3],expected[,4],expected[,5],expected[,6])

means=cbind(combmeans11,combmeans21,
combmeans31,combmeans41,combmeans51,combmeans61,combmeans71,combmeans81,combmeans91,
combmeans101,combmeans111,combmeans121,
combmeans12,combmeans22, combmeans32,combmeans42,combmeans52,combmeans62,combmeans72,combmeans82,combmeans92,
combmeans102,combmeans112,combmeans122,
combmeans13,combmeans23, combmeans33,combmeans43,combmeans53,combmeans63,combmeans73,combmeans83,combmeans93,
combmeans103,combmeans113,combmeans123,
combmeans14,combmeans24, combmeans34,combmeans44,combmeans54,combmeans64,combmeans74,combmeans84,combmeans94,
combmeans104,combmeans114,combmeans124,
combmeans15,combmeans25, combmeans35,combmeans45,combmeans55,combmeans65,combmeans75,combmeans85,combmeans95,
combmeans105,combmeans115,combmeans125,
combmeans16,combmeans26, combmeans36,combmeans46,combmeans56,combmeans66,combmeans76,combmeans86,combmeans96,
combmeans106,combmeans116,combmeans126)

variance1=(1/nvar)*sum((sampmean[,1]-expected1)**2)

variance11=(1/nvar)*sum((means[,1]-expected[1,1])**2)
variance21=(1/nvar)*sum((means[,2]-expected[2,1])**2)
variance31=(1/nvar)*sum((means[,3]-expected[3,1])**2)
variance41=(1/nvar)*sum((means[,4]-expected[4,1])**2)
variance51=(1/nvar)*sum((means[,5]-expected[5,1])**2)
variance61=(1/nvar)*sum((means[,6]-expected[6,1])**2)
variance71=(1/nvar)*sum((means[,7]-expected[7,1])**2)
variance81=(1/nvar)*sum((means[,8]-expected[8,1])**2)
variance91=(1/nvar)*sum((means[,9]-expected[9,1])**2)
variance101=(1/nvar)*sum((means[,10]-expected[10,1])**2)
variance111=(1/nvar)*sum((means[,11]-expected[11,1])**2)
variance121=(1/nvar)*sum((means[,12]-expected[12,1])**2)

variance12=(1/nvar)*sum((means[,13]-expected[1,2])**2)
variance22=(1/nvar)*sum((means[,14]-expected[2,2])**2)
variance32=(1/nvar)*sum((means[,15]-expected[3,2])**2)
variance42=(1/nvar)*sum((means[,16]-expected[4,2])**2)
variance52=(1/nvar)*sum((means[,17]-expected[5,2])**2)
variance62=(1/nvar)*sum((means[,18]-expected[6,2])**2)
variance72=(1/nvar)*sum((means[,19]-expected[7,2])**2)
variance82=(1/nvar)*sum((means[,20]-expected[8,2])**2)
variance92=(1/nvar)*sum((means[,21]-expected[9,2])**2)
variance102=(1/nvar)*sum((means[,22]-expected[10,2])**2)
variance112=(1/nvar)*sum((means[,23]-expected[11,2])**2)
variance122=(1/nvar)*sum((means[,24]-expected[12,2])**2)

variance13=(1/nvar)*sum((means[,25]-expected[1,3])**2)
variance23=(1/nvar)*sum((means[,26]-expected[2,3])**2)
variance33=(1/nvar)*sum((means[,27]-expected[3,3])**2)
variance43=(1/nvar)*sum((means[,28]-expected[4,3])**2)
variance53=(1/nvar)*sum((means[,29]-expected[5,3])**2)
variance63=(1/nvar)*sum((means[,30]-expected[6,3])**2)
variance73=(1/nvar)*sum((means[,31]-expected[7,3])**2)
variance83=(1/nvar)*sum((means[,32]-expected[8,3])**2)
variance93=(1/nvar)*sum((means[,33]-expected[9,3])**2)
variance103=(1/nvar)*sum((means[,34]-expected[10,3])**2)
variance113=(1/nvar)*sum((means[,35]-expected[11,3])**2)
variance123=(1/nvar)*sum((means[,36]-expected[12,3])**2)

variance14=(1/nvar)*sum((means[,37]-expected[1,4])**2)
variance24=(1/nvar)*sum((means[,38]-expected[2,4])**2)
variance34=(1/nvar)*sum((means[,39]-expected[3,4])**2)
variance44=(1/nvar)*sum((means[,40]-expected[4,4])**2)
variance54=(1/nvar)*sum((means[,41]-expected[5,4])**2)
variance64=(1/nvar)*sum((means[,42]-expected[6,4])**2)
variance74=(1/nvar)*sum((means[,43]-expected[7,4])**2)
variance84=(1/nvar)*sum((means[,44]-expected[8,4])**2)
variance94=(1/nvar)*sum((means[,45]-expected[9,4])**2)
variance104=(1/nvar)*sum((means[,46]-expected[10,4])**2)
variance114=(1/nvar)*sum((means[,47]-expected[11,4])**2)
variance124=(1/nvar)*sum((means[,48]-expected[12,4])**2)

```

variance15=(1/nvar)*sum((means[,49]-expected[1,5])**2)
variance25=(1/nvar)*sum((means[,50]-expected[2,5])**2)
variance35=(1/nvar)*sum((means[,51]-expected[3,5])**2)
variance45=(1/nvar)*sum((means[,52]-expected[4,5])**2)
variance55=(1/nvar)*sum((means[,53]-expected[5,5])**2)
variance65=(1/nvar)*sum((means[,54]-expected[6,5])**2)
variance75=(1/nvar)*sum((means[,55]-expected[7,5])**2)
variance85=(1/nvar)*sum((means[,56]-expected[8,5])**2)
variance95=(1/nvar)*sum((means[,57]-expected[9,5])**2)
variance105=(1/nvar)*sum((means[,58]-expected[10,5])**2)
variance115=(1/nvar)*sum((means[,59]-expected[11,5])**2)
variance125=(1/nvar)*sum((means[,60]-expected[12,5])**2)

```

```

variance16=(1/nvar)*sum((means[,61]-expected[1,6])**2)
variance26=(1/nvar)*sum((means[,62]-expected[2,6])**2)
variance36=(1/nvar)*sum((means[,63]-expected[3,6])**2)
variance46=(1/nvar)*sum((means[,64]-expected[4,6])**2)
variance56=(1/nvar)*sum((means[,65]-expected[5,6])**2)
variance66=(1/nvar)*sum((means[,66]-expected[6,6])**2)
variance76=(1/nvar)*sum((means[,67]-expected[7,6])**2)
variance86=(1/nvar)*sum((means[,68]-expected[8,6])**2)
variance96=(1/nvar)*sum((means[,69]-expected[9,6])**2)
variance106=(1/nvar)*sum((means[,70]-expected[10,6])**2)
variance116=(1/nvar)*sum((means[,71]-expected[11,6])**2)
variance126=(1/nvar)*sum((means[,72]-expected[12,6])**2)

```

```

variance=c(variance1,variance11,variance21,variance31,variance41,variance51,
variance61,variance71,variance81,variance91,variance101,variance111,variance121,
variance12,variance22,variance32,variance42,variance52,
variance62,variance72,variance82,variance92,variance102,variance112,variance122,
variance13,variance23,variance33,variance43,variance53,
variance63,variance73,variance83,variance93,variance103,variance113,variance123,
variance14,variance24,variance34,variance44,variance54,
variance64,variance74,variance84,variance94,variance104,variance114,variance124,
variance15,variance25,variance35,variance45,variance55,
variance65,variance75,variance85,variance95,variance105,variance115,variance125,
variance16,variance26,variance36,variance46,variance56,
variance66,variance76,variance86,variance96,variance106,variance116,variance126)

```

```
bias= expectedcol-popmean
```

```
relbias=(bias/expectedcol)*100
```

```
mse= variance + bias**2
```

```
cv= (sqrt(variance)/expectedcol)*100
```

```
s.e= (sqrt(variance)/sqrt(nvar))
```

```
LB= expectedcol - 1.96*s.e
```

```
UB= expectedcol + 1.96*s.e
```

```
result=cbind(popmean,expectedcol,relbias,s.e,cv,LB,UB,mse)
```

```
write.table(result,"result_anormal_scenario1.txt",sep="\t")
```

```
write.table(meff_var,"meff_anormal_scenario1.txt",sep="\t")
```

G4 Skewed T Distribution, 3 degrees of freedom

```
#####
# Title: Stratified Fox inspired simulations_Normal distribution
# author: "Loveness Dzikiti"
# date: "22 September 2017"
#####

setwd("C:\\Users\\User\\Dropbox\\Review\\simulations_2019\\stratified\\at3")
#install.packages("survey")

library(moments)
library(sampling)
library(survey)
library(sn)

N=1000000      #finite population size
L=5           # number of strata in population
Nstrata=N/L
nsamp=20      #number of samples or replicates
k=nvar=1000

set.seed(123456)

#####
# Generating a stratified population
v=3
xi=-0.7891
omega=0.746
alpha=3

x<- rst(N, xi, omega, alpha, v)

x1=x[1:200000]
x2=x[200001:400000]
x3=x[400001:600000]
x4=x[600001:800000]
x5=x[800001:1000000]

mu1=105
sigma1=4

mu2=103
sigma2=3

mu3=100
sigma3=3

mu4=97
sigma4=3

mu5=95
sigma5=4

x=c(x1*sigma1+mu1, (x2*sigma2)+mu2, (x3*sigma3)+mu3, (x4*sigma4)+mu4, (x5*sigma5)+mu5)
strata_id<-c(rep('1',Nstrata),rep('2',Nstrata),rep('3',Nstrata),rep('4',Nstrata),rep('5',Nstrata))
popx<-data.frame(1:1000000,strata_id,x)
colnames(popx)<-c('ID','group','x')

pop.des <- svydesign(id=~1,strata=~group,data=popx)
pop.mean=data.frame(svymean(x,pop.des))
popmean <- pop.mean[1,1]#finite population mean

#####

mean_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample means calculated as srs
variance_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample variance calculated as srs
cv_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample coefficient of variation calculated as srs
skewness_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample skewness calculated as srs
kurtosis_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample kurtosis calculated as srs
```



```

combmeans34 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans44 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans54 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5
combmeans64 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 6
combmeans74 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 7
combmeans84 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
combmeans94 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
combmeans104 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans114 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans124 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5

```

```

combmeans15 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
combmeans25 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
combmeans35 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans45 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans55 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5
combmeans65 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 6
combmeans75 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 7
combmeans85 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
combmeans95 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
combmeans105 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans115 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans125 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5

```

```

combmeans16 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
combmeans26 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
combmeans36 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans46 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans56 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5
combmeans66 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 6
combmeans76 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 7
combmeans86 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
combmeans96 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
combmeans106 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans116 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans126 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5

```

```
#####
```

```
for (k in 1:nvar){
```

```
for (j in 1:nsamp){
```

```

w1<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1) #store weights
w2<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w3<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w4<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w5<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w6<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w7<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w8<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1) #store weights
w9<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w10<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w11<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w12<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
}

```

```
##### Sampling from stratified finite population#####
```

```

#####Calculating sample sizes for strata
L=5 #number of strata
propn<-c(0.00144,0.00072,0.00072,0.00072,0.00144)
Nstrata<-N/L #strata size for equal strata sizes in population

sampsiz<-ceiling(Nstrata*propn)
n<-sampsiz[1]+sampsiz[2]+sampsiz[3]+sampsiz[4]+sampsiz[5]

```

```

fpc=rep(c(Nstrata,Nstrata,Nstrata,Nstrata,Nstrata),sampsi)

for (i in 1:nsamp) {

s1<-sample(popx$x[1:200000],sampsi[1])
s2<-sample(popx$x[200001:400000],sampsi[2])
s3<-sample(popx$x[400001:600000],sampsi[3])
s4<-sample(popx$x[600001:800000],sampsi[4])
s5<-sample(popx$x[800001:1000000],sampsi[5])

sampstrat<-c(rep('1',sampsi[1]),rep('2',sampsi[2]),rep('3',sampsi[3]),rep('4',sampsi[4]),rep('5',sampsi[5]))

s<-data.frame(sampstrat,c(s1,s2,s3,s4,s5))
colnames(s)<-c('sampstrat','x')

#####
# Naive weights
#####

mean_naive[i]=mean(s$x)
variance_naive[i]=var(s$x)
cv_naive[i]=abs(sqrt(variance_naive[i])/mean_naive[i])
skewness_naive[i]<-skewness(s$x)
kurtosis_naive[i]<-kurtosis(s$x)

#####
# Design conscious weights
#####

plan=svydesign(~1,strata=~sampstrat,data=s,fpc=~fpc) #probs=~propn is not the cluster samp. prob.
mean_st=svymean(s$x,plan)
mean_st=as.data.frame(mean_st)

means[i]<-mean_st[1,1]          #sample mean for ith sample
variance[i]<-mean_st[1,2]^2    #sample variance, design consistent variance
cv[i]<-abs(sqrt(variance[i])/means[i])    #coefficient of variation

# skewness[i]<-skewness(s[,i])
moments = svymean(~I(x^4) + I(x^3) +
                 I(x^2) + I(x), plan)

skew=data.frame(svycontrast(moments,
  quote('I(x^3)' - 3*I(x^2)*I(x)' +
        3*I(x)*I(x)^2 - I(x)^3)/
        ('I(x^2)' - I(x)^2)^1.5)))
skewness[i]<-skew[1,1]

#kurtosis[i]<-kurtosis(s[,i])

kurt=data.frame(svycontrast(moments,
  quote('I(x^4)' -
        4*I(x^3)*I(x)' +
        6*I(x^2)*I(x)^2 -
        4*I(x)*I(x)^3 +
        I(x)^4)/
        ('I(x^2)' -
        I(x)^2)^2)))

kurtosis[i]<-kurt[1,1]

meff.1 <- data.frame(svymean(s$x,plan,deff=TRUE))
meff_var[i]=variance[i]/((1-(n/N))*(var(s$x)/n))

w1[i]=1/n          #inverse of the sample size

w2[i]<-1/(variance[i]) #inverse of the design variance of the estimator of the mean
w3[i]<-1/(variance_naive[i]) #inverse of the naive variance of the estimator of the mean

```

```

w4[i]=1          #simple average weights

w5[i]=1/cv[i]    #inverse of cv
w6[i]=1/cv_naive[i]  #inverse of naive cv

w7[i]=abs(1/skewness[i])  #inverse skewness
w8[i]=abs(1/skewness_naive[i])  #inverse skewness

w9[i]=1/kurtosis[i]  #inverse kurtosis
w10[i]=1/kurtosis_naive[i]  #inverse kurtosis

w11[i]=meff_var[i]/variance[i] #in srs deff=1 so coincides with inverse of variance
w12[i]=1/meff_var[i]

combmeans1[i]<-round(sum(means[1:i]*w1[1:i])/sum(w1[1:i]),digits=9) #combine means using inverse sample size
combmeans2[i]<-round(sum(means[1:i]*w2[1:i])/sum(w2[1:i]),digits=9) # combine using inverse variance
combmeans3[i]<-round(sum(means[1:i]*w3[1:i])/sum(w3[1:i]),digits=9) # combine using inverse variance naive
combmeans4[i]<-round(sum(means[1:i]*w4[1:i])/sum(w4[1:i]),digits=9) # combine using a weight of 1
combmeans5[i]<-round(sum(means[1:i]*w5[1:i])/sum(w5[1:i]),digits=9) # combine using a weight of inverse CV
combmeans6[i]<-round(sum(means[1:i]*w6[1:i])/sum(w6[1:i]),digits=9) # combine using a weight of inverse CV naive
combmeans7[i]<-round(sum(means[1:i]*w7[1:i])/sum(w7[1:i]),digits=9) # combine using a weight of inverse skewness
combmeans8[i]<-round(sum(means[1:i]*w8[1:i])/sum(w8[1:i]),digits=9) #combine using a weight of inverse skewness naive
combmeans9[i]<-round(sum(means[1:i]*w9[1:i])/sum(w9[1:i]),digits=9) # combine using inverse kurtosis
combmeans10[i]<-round(sum(means[1:i]*w10[1:i])/sum(w10[1:i]),digits=9) # combine using inverse kurtosis naive
combmeans11[i]<-round(sum(means[1:i]*w11[1:i])/sum(w11[1:i]),digits=9) # combine using a weight of meff/var
combmeans12[i]<-round(sum(means[1:i]*w12[1:i])/sum(w12[1:i]),digits=9) # combine using a weight of 1/meff

}

sampmean[k]=means[1] # capturing the combined mean, where nsamp reps are combined, all weighting strategies

combmeans11[k]=combmeans1[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans21[k]=combmeans2[2]
combmeans31[k]=combmeans3[2]
combmeans41[k]=combmeans4[2]
combmeans51[k]=combmeans5[2]
combmeans61[k]=combmeans6[2]
combmeans71[k]=combmeans7[2]
combmeans81[k]=combmeans8[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans91[k]=combmeans9[2]
combmeans101[k]=combmeans10[2]
combmeans111[k]=combmeans11[2]
combmeans121[k]=combmeans12[2]

combmeans12[k]=combmeans1[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans22[k]=combmeans2[3]
combmeans32[k]=combmeans3[3]
combmeans42[k]=combmeans4[3]
combmeans52[k]=combmeans5[3]
combmeans62[k]=combmeans6[3]
combmeans72[k]=combmeans7[3]
combmeans82[k]=combmeans8[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans92[k]=combmeans9[3]
combmeans102[k]=combmeans10[3]
combmeans112[k]=combmeans11[3]
combmeans122[k]=combmeans12[3]

combmeans13[k]=combmeans1[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans23[k]=combmeans2[5]
combmeans33[k]=combmeans3[5]
combmeans43[k]=combmeans4[5]
combmeans53[k]=combmeans5[5]
combmeans63[k]=combmeans6[5]
combmeans73[k]=combmeans7[5]

```

```

combmeans83[k]=combmeans8[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans93[k]=combmeans9[5]
combmeans103[k]=combmeans10[5]
combmeans113[k]=combmeans11[5]
combmeans123[k]=combmeans12[5]

combmeans14[k]=combmeans1[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans24[k]=combmeans2[10]
combmeans34[k]=combmeans3[10]
combmeans44[k]=combmeans4[10]
combmeans54[k]=combmeans5[10]
combmeans64[k]=combmeans6[10]
combmeans74[k]=combmeans7[10]
combmeans84[k]=combmeans8[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans94[k]=combmeans9[10]
combmeans104[k]=combmeans10[10]
combmeans114[k]=combmeans11[10]
combmeans124[k]=combmeans12[10]

combmeans15[k]=combmeans1[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans25[k]=combmeans2[15]
combmeans35[k]=combmeans3[15]
combmeans45[k]=combmeans4[15]
combmeans55[k]=combmeans5[15]
combmeans65[k]=combmeans6[15]
combmeans75[k]=combmeans7[15]
combmeans85[k]=combmeans8[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans95[k]=combmeans9[15]
combmeans105[k]=combmeans10[15]
combmeans115[k]=combmeans11[15]
combmeans125[k]=combmeans12[15]

combmeans16[k]=combmeans1[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans26[k]=combmeans2[20]
combmeans36[k]=combmeans3[20]
combmeans46[k]=combmeans4[20]
combmeans56[k]=combmeans5[20]
combmeans66[k]=combmeans6[20]
combmeans76[k]=combmeans7[20]
combmeans86[k]=combmeans8[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans96[k]=combmeans9[20]
combmeans106[k]=combmeans10[20]
combmeans116[k]=combmeans11[20]
combmeans126[k]=combmeans12[20]

}

expected1=mean(sampmean) #expected mean for one sample, not combining

expected=matrix(0,12,6)

expected[1,1]=mean(combmeans11) #strategy 1 nsamp=2
expected[2,1]=mean(combmeans21)
expected[3,1]=mean(combmeans31)
expected[4,1]=mean(combmeans41)
expected[5,1]=mean(combmeans51)
expected[6,1]=mean(combmeans61)
expected[7,1]=mean(combmeans71)
expected[8,1]=mean(combmeans81)
expected[9,1]=mean(combmeans91)
expected[10,1]=mean(combmeans101)
expected[11,1]=mean(combmeans111)
expected[12,1]=mean(combmeans121)

expected[1,2]=mean(combmeans12) #strategy 1 nsamp=3
expected[2,2]=mean(combmeans22)

```

```

expected[3,2]=mean(combmeans32)
expected[4,2]=mean(combmeans42)
expected[5,2]=mean(combmeans52)
expected[6,2]=mean(combmeans62)
expected[7,2]=mean(combmeans72)
expected[8,2]=mean(combmeans82)
expected[9,2]=mean(combmeans92)
expected[10,2]=mean(combmeans102)
expected[11,2]=mean(combmeans112)
expected[12,2]=mean(combmeans122)

expected[1,3]=mean(combmeans13) #strategy 1 nsamp=5
expected[2,3]=mean(combmeans23)
expected[3,3]=mean(combmeans33)
expected[4,3]=mean(combmeans43)
expected[5,3]=mean(combmeans53)
expected[6,3]=mean(combmeans63)
expected[7,3]=mean(combmeans73)
expected[8,3]=mean(combmeans83)
expected[9,3]=mean(combmeans93)
expected[10,3]=mean(combmeans103)
expected[11,3]=mean(combmeans113)
expected[12,3]=mean(combmeans123)

expected[1,4]=mean(combmeans14) #strategy 1 nsamp=10
expected[2,4]=mean(combmeans24)
expected[3,4]=mean(combmeans34)
expected[4,4]=mean(combmeans44)
expected[5,4]=mean(combmeans54)
expected[6,4]=mean(combmeans64)
expected[7,4]=mean(combmeans74)
expected[8,4]=mean(combmeans84)
expected[9,4]=mean(combmeans94)
expected[10,4]=mean(combmeans104)
expected[11,4]=mean(combmeans114)
expected[12,4]=mean(combmeans124)

expected[1,5]=mean(combmeans15) #strategy 1 nsamp=15
expected[2,5]=mean(combmeans25)
expected[3,5]=mean(combmeans35)
expected[4,5]=mean(combmeans45)
expected[5,5]=mean(combmeans55)
expected[6,5]=mean(combmeans65)
expected[7,5]=mean(combmeans75)
expected[8,5]=mean(combmeans85)
expected[9,5]=mean(combmeans95)
expected[10,5]=mean(combmeans105)
expected[11,5]=mean(combmeans115)
expected[12,5]=mean(combmeans125)

expected[1,6]=mean(combmeans16) #strategy 1 nsamp=20
expected[2,6]=mean(combmeans26)
expected[3,6]=mean(combmeans36)
expected[4,6]=mean(combmeans46)
expected[5,6]=mean(combmeans56)
expected[6,6]=mean(combmeans66)
expected[7,6]=mean(combmeans76)
expected[8,6]=mean(combmeans86)
expected[9,6]=mean(combmeans96)
expected[10,6]=mean(combmeans106)
expected[11,6]=mean(combmeans116)
expected[12,6]=mean(combmeans126)

expectedcol=c(expected1,expected[,1],expected[,2],expected[,3],expected[,4],expected[,5],expected[,6])

means=cbind(combmeans11,combmeans21,
combmeans31,combmeans41,combmeans51,combmeans61,combmeans71,combmeans81,combmeans91,
combmeans101,combmeans111,combmeans121,
combmeans12,combmeans22, combmeans32,combmeans42,combmeans52,combmeans62,combmeans72,combmeans82,combmeans92,
combmeans102,combmeans112,combmeans122,

```

combmeans13,combmeans23, combmeans33,combmeans43,combmeans53,combmeans63,combmeans73,combmeans83,combmeans93,
 combmeans103,combmeans113,combmeans123,
 combmeans14,combmeans24, combmeans34,combmeans44,combmeans54,combmeans64,combmeans74,combmeans84,combmeans94,
 combmeans104,combmeans114,combmeans124,
 combmeans15,combmeans25, combmeans35,combmeans45,combmeans55,combmeans65,combmeans75,combmeans85,combmeans95,
 combmeans105,combmeans115,combmeans125,
 combmeans16,combmeans26, combmeans36,combmeans46,combmeans56,combmeans66,combmeans76,combmeans86,combmeans96,
 combmeans106,combmeans116,combmeans126)

variance1=(1/nvar)*sum((sampmean[,1]-expected1)**2)

variance11=(1/nvar)*sum((means[,1]-expected[1,1])**2)
 variance21=(1/nvar)*sum((means[,2]-expected[2,1])**2)
 variance31=(1/nvar)*sum((means[,3]-expected[3,1])**2)
 variance41=(1/nvar)*sum((means[,4]-expected[4,1])**2)
 variance51=(1/nvar)*sum((means[,5]-expected[5,1])**2)
 variance61=(1/nvar)*sum((means[,6]-expected[6,1])**2)
 variance71=(1/nvar)*sum((means[,7]-expected[7,1])**2)
 variance81=(1/nvar)*sum((means[,8]-expected[8,1])**2)
 variance91=(1/nvar)*sum((means[,9]-expected[9,1])**2)
 variance101=(1/nvar)*sum((means[,10]-expected[10,1])**2)
 variance111=(1/nvar)*sum((means[,11]-expected[11,1])**2)
 variance121=(1/nvar)*sum((means[,12]-expected[12,1])**2)

variance12=(1/nvar)*sum((means[,13]-expected[1,2])**2)
 variance22=(1/nvar)*sum((means[,14]-expected[2,2])**2)
 variance32=(1/nvar)*sum((means[,15]-expected[3,2])**2)
 variance42=(1/nvar)*sum((means[,16]-expected[4,2])**2)
 variance52=(1/nvar)*sum((means[,17]-expected[5,2])**2)
 variance62=(1/nvar)*sum((means[,18]-expected[6,2])**2)
 variance72=(1/nvar)*sum((means[,19]-expected[7,2])**2)
 variance82=(1/nvar)*sum((means[,20]-expected[8,2])**2)
 variance92=(1/nvar)*sum((means[,21]-expected[9,2])**2)
 variance102=(1/nvar)*sum((means[,22]-expected[10,2])**2)
 variance112=(1/nvar)*sum((means[,23]-expected[11,2])**2)
 variance122=(1/nvar)*sum((means[,24]-expected[12,2])**2)

variance13=(1/nvar)*sum((means[,25]-expected[1,3])**2)
 variance23=(1/nvar)*sum((means[,26]-expected[2,3])**2)
 variance33=(1/nvar)*sum((means[,27]-expected[3,3])**2)
 variance43=(1/nvar)*sum((means[,28]-expected[4,3])**2)
 variance53=(1/nvar)*sum((means[,29]-expected[5,3])**2)
 variance63=(1/nvar)*sum((means[,30]-expected[6,3])**2)
 variance73=(1/nvar)*sum((means[,31]-expected[7,3])**2)
 variance83=(1/nvar)*sum((means[,32]-expected[8,3])**2)
 variance93=(1/nvar)*sum((means[,33]-expected[9,3])**2)
 variance103=(1/nvar)*sum((means[,34]-expected[10,3])**2)
 variance113=(1/nvar)*sum((means[,35]-expected[11,3])**2)
 variance123=(1/nvar)*sum((means[,36]-expected[12,3])**2)

variance14=(1/nvar)*sum((means[,37]-expected[1,4])**2)
 variance24=(1/nvar)*sum((means[,38]-expected[2,4])**2)
 variance34=(1/nvar)*sum((means[,39]-expected[3,4])**2)
 variance44=(1/nvar)*sum((means[,40]-expected[4,4])**2)
 variance54=(1/nvar)*sum((means[,41]-expected[5,4])**2)
 variance64=(1/nvar)*sum((means[,42]-expected[6,4])**2)
 variance74=(1/nvar)*sum((means[,43]-expected[7,4])**2)
 variance84=(1/nvar)*sum((means[,44]-expected[8,4])**2)
 variance94=(1/nvar)*sum((means[,45]-expected[9,4])**2)
 variance104=(1/nvar)*sum((means[,46]-expected[10,4])**2)
 variance114=(1/nvar)*sum((means[,47]-expected[11,4])**2)
 variance124=(1/nvar)*sum((means[,48]-expected[12,4])**2)

variance15=(1/nvar)*sum((means[,49]-expected[1,5])**2)
 variance25=(1/nvar)*sum((means[,50]-expected[2,5])**2)
 variance35=(1/nvar)*sum((means[,51]-expected[3,5])**2)
 variance45=(1/nvar)*sum((means[,52]-expected[4,5])**2)
 variance55=(1/nvar)*sum((means[,53]-expected[5,5])**2)
 variance65=(1/nvar)*sum((means[,54]-expected[6,5])**2)
 variance75=(1/nvar)*sum((means[,55]-expected[7,5])**2)

```

variance85=(1/nvar)*sum((means[,56]-expected[8,5])**2)
variance95=(1/nvar)*sum((means[,57]-expected[9,5])**2)
variance105=(1/nvar)*sum((means[,58]-expected[10,5])**2)
variance115=(1/nvar)*sum((means[,59]-expected[11,5])**2)
variance125=(1/nvar)*sum((means[,60]-expected[12,5])**2)

variance16=(1/nvar)*sum((means[,61]-expected[1,6])**2)
variance26=(1/nvar)*sum((means[,62]-expected[2,6])**2)
variance36=(1/nvar)*sum((means[,63]-expected[3,6])**2)
variance46=(1/nvar)*sum((means[,64]-expected[4,6])**2)
variance56=(1/nvar)*sum((means[,65]-expected[5,6])**2)
variance66=(1/nvar)*sum((means[,66]-expected[6,6])**2)
variance76=(1/nvar)*sum((means[,67]-expected[7,6])**2)
variance86=(1/nvar)*sum((means[,68]-expected[8,6])**2)
variance96=(1/nvar)*sum((means[,69]-expected[9,6])**2)
variance106=(1/nvar)*sum((means[,70]-expected[10,6])**2)
variance116=(1/nvar)*sum((means[,71]-expected[11,6])**2)
variance126=(1/nvar)*sum((means[,72]-expected[12,6])**2)

variance=c(variance1,variance11,variance21,variance31,variance41,variance51,
variance61,variance71,variance81,variance91,variance101,variance111,variance121,
variance12,variance22,variance32,variance42,variance52,
variance62,variance72,variance82,variance92,variance102,variance112,variance122,
variance13,variance23,variance33,variance43,variance53,
variance63,variance73,variance83,variance93,variance103,variance113,variance123,
variance14,variance24,variance34,variance44,variance54,
variance64,variance74,variance84,variance94,variance104,variance114,variance124,
variance15,variance25,variance35,variance45,variance55,
variance65,variance75,variance85,variance95,variance105,variance115,variance125,
variance16,variance26,variance36,variance46,variance56,
variance66,variance76,variance86,variance96,variance106,variance116,variance126)

bias= expectedcol-popmean

relbias=(bias/expectedcol)*100

mse= variance + bias**2

cv= (sqrt(variance)/expectedcol)*100

s.e= (sqrt(variance)/sqrt(nvar))

LB= expectedcol - 1.96*s.e
UB= expectedcol + 1.96*s.e

result=cbind(popmean,expectedcol,relbias,s.e,cv,LB,UB,mse)

write.table(result,"result_at3_scenario1.txt",sep="\t")
write.table(meff_var,"meff_at3_scenario1.txt",sep="\t")

```

Appendix H

R code used in simulations for cluster random sampling, point estimators

H1 Normal Distribution

```
#####  
# Variance Estimation: Normal distribution  
# author: "Loveness Dzikiti"  
# date: "20 April 2017"  
#####  
  
setwd("C:\\Users\\Loveness\\Dropbox\\Review\\simulations_2019\\cluster\\normal")  
  
library(moments)  
library(sampling)  
library(survey)  
  
N=1000000      #finite population size  
nsamp=20       #number of samples or replicates  
k=nvar=100  
  
n=1000  
clsiz=250  
  
set.seed(123456)  
  
mean_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample means calculated as srs  
variance_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample variance calculated as srs  
cv_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample coefficient of variation calculated as srs  
skewness_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample skewness calculated as srs  
kurtosis_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample kurtosis calculated as srs  
  
means <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample means  
variance <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample variance  
cv <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample coefficient of variation  
skewness <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample skewness  
kurtosis <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample kurtosis  
meff_var<-matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample design effect  
  
combmeans1 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 1  
combmeans2 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 2  
combmeans3 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 3  
combmeans4 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 4  
combmeans5 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 5  
combmeans6 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 6  
combmeans7 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 7  
combmeans8 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 8  
combmeans9 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 9  
combmeans10<- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 10  
combmeans11<- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 11  
combmeans12<- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 12  
  
samppmean <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store mean of 1 sample, no combining  
  
combmeans11 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1  
combmeans21 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2  
combmeans31 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3  
combmeans41 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4  
combmeans51 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5  
combmeans61 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 6  
combmeans71 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 7  
combmeans81 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1  
combmeans91 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2  
combmeans101 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3  
combmeans111 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4  
combmeans121 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5
```


#####

Generating a clustered population#####

```
mu <- matrix(c(rep(100,N)), nrow = N, ncol = 1) #store superpopulation mean
e_ij <- rnorm(N,0,sqrt(20)) #individual characteristic
cl <- rnorm(4000,0,sqrt(5)) #cluster characteristic

v_j <-
c(rep(cl[1],clsize),rep(cl[2],clsize),rep(cl[3],clsize),rep(cl[4],clsize),rep(cl[5],clsize),rep(cl[6],clsize),rep(cl[7],clsize),rep(cl[8],clsize),rep(cl[9],cl
size),rep(cl[10],clsize),
rep(cl[11],clsize),rep(cl[12],clsize),rep(cl[13],clsize),rep(cl[14],clsize),rep(cl[15],clsize),rep(cl[16],clsize),rep(cl[17],clsize),rep(cl[18],clsize),re
p(cl[19],clsize),rep(cl[20],clsize),
rep(cl[21],clsize),rep(cl[22],clsize),rep(cl[23],clsize),rep(cl[24],clsize),rep(cl[25],clsize),rep(cl[26],clsize),rep(cl[27],clsize),rep(cl[28],clsize),re
p(cl[29],clsize),rep(cl[30],clsize),
rep(cl[31],clsize),rep(cl[32],clsize),rep(cl[33],clsize),rep(cl[34],clsize),rep(cl[35],clsize),rep(cl[36],clsize),rep(cl[37],clsize),rep(cl[38],clsize),re
p(cl[39],clsize),rep(cl[40],clsize),
rep(cl[41],clsize),rep(cl[42],clsize),rep(cl[43],clsize),rep(cl[44],clsize),rep(cl[45],clsize),rep(cl[46],clsize),rep(cl[47],clsize),rep(cl[48],clsize),re
p(cl[49],clsize),rep(cl[50],clsize),
rep(cl[51],clsize),rep(cl[52],clsize),rep(cl[53],clsize),rep(cl[54],clsize),rep(cl[55],clsize),rep(cl[56],clsize),rep(cl[57],clsize),rep(cl[58],clsize),re
p(cl[59],clsize),rep(cl[60],clsize),
rep(cl[61],clsize),rep(cl[62],clsize),rep(cl[63],clsize),rep(cl[64],clsize),rep(cl[65],clsize),rep(cl[66],clsize),rep(cl[67],clsize),rep(cl[68],clsize),re
p(cl[69],clsize),rep(cl[70],clsize),
rep(cl[71],clsize),rep(cl[72],clsize),rep(cl[73],clsize),rep(cl[74],clsize),rep(cl[75],clsize),rep(cl[76],clsize),rep(cl[77],clsize),rep(cl[78],clsize),re
p(cl[79],clsize),rep(cl[80],clsize),
rep(cl[81],clsize),rep(cl[82],clsize),rep(cl[83],clsize),rep(cl[84],clsize),rep(cl[85],clsize),rep(cl[86],clsize),rep(cl[87],clsize),rep(cl[88],clsize),re
p(cl[89],clsize),rep(cl[90],clsize),
rep(cl[91],clsize),rep(cl[92],clsize),rep(cl[93],clsize),rep(cl[94],clsize),rep(cl[95],clsize),rep(cl[96],clsize),rep(cl[97],clsize),rep(cl[98],clsize),re
p(cl[99],clsize),rep(cl[100],clsize),
rep(cl[101],clsize),rep(cl[102],clsize),rep(cl[103],clsize),rep(cl[104],clsize),rep(cl[105],clsize),rep(cl[106],clsize),rep(cl[107],clsize),rep(cl[108
],clsize),rep(cl[109],clsize),rep(cl[110],clsize),
rep(cl[111],clsize),rep(cl[112],clsize),rep(cl[113],clsize),rep(cl[114],clsize),rep(cl[115],clsize),rep(cl[116],clsize),rep(cl[117],clsize),rep(cl[118
],clsize),rep(cl[119],clsize),rep(cl[120],clsize),
rep(cl[121],clsize),rep(cl[122],clsize),rep(cl[123],clsize),rep(cl[124],clsize),rep(cl[125],clsize),rep(cl[126],clsize),rep(cl[127],clsize),rep(cl[128
],clsize),rep(cl[129],clsize),rep(cl[130],clsize),
rep(cl[131],clsize),rep(cl[132],clsize),rep(cl[133],clsize),rep(cl[134],clsize),rep(cl[135],clsize),rep(cl[136],clsize),rep(cl[137],clsize),rep(cl[138
],clsize),rep(cl[139],clsize),rep(cl[140],clsize),
rep(cl[141],clsize),rep(cl[142],clsize),rep(cl[143],clsize),rep(cl[144],clsize),rep(cl[145],clsize),rep(cl[146],clsize),rep(cl[147],clsize),rep(cl[148
],clsize),rep(cl[149],clsize),rep(cl[150],clsize),
rep(cl[151],clsize),rep(cl[152],clsize),rep(cl[153],clsize),rep(cl[154],clsize),rep(cl[155],clsize),rep(cl[156],clsize),rep(cl[157],clsize),rep(cl[158
],clsize),rep(cl[159],clsize),rep(cl[160],clsize),
rep(cl[161],clsize),rep(cl[162],clsize),rep(cl[163],clsize),rep(cl[164],clsize),rep(cl[165],clsize),rep(cl[166],clsize),rep(cl[167],clsize),rep(cl[168
],clsize),rep(cl[169],clsize),rep(cl[170],clsize),
rep(cl[171],clsize),rep(cl[172],clsize),rep(cl[173],clsize),rep(cl[174],clsize),rep(cl[175],clsize),rep(cl[176],clsize),rep(cl[177],clsize),rep(cl[178
],clsize),rep(cl[179],clsize),rep(cl[180],clsize),
rep(cl[181],clsize),rep(cl[182],clsize),rep(cl[183],clsize),rep(cl[184],clsize),rep(cl[185],clsize),rep(cl[186],clsize),rep(cl[187],clsize),rep(cl[188
],clsize),rep(cl[189],clsize),rep(cl[190],clsize),
rep(cl[191],clsize),rep(cl[192],clsize),rep(cl[193],clsize),rep(cl[194],clsize),rep(cl[195],clsize),rep(cl[196],clsize),rep(cl[197],clsize),rep(cl[198
],clsize),rep(cl[199],clsize),rep(cl[200],clsize),
rep(cl[201],clsize),rep(cl[202],clsize),rep(cl[203],clsize),rep(cl[204],clsize),rep(cl[205],clsize),rep(cl[206],clsize),rep(cl[207],clsize),rep(cl[208
],clsize),rep(cl[209],clsize),rep(cl[210],clsize),
rep(cl[211],clsize),rep(cl[212],clsize),rep(cl[213],clsize),rep(cl[214],clsize),rep(cl[215],clsize),rep(cl[216],clsize),rep(cl[217],clsize),rep(cl[218
],clsize),rep(cl[219],clsize),rep(cl[220],clsize),
rep(cl[221],clsize),rep(cl[222],clsize),rep(cl[223],clsize),rep(cl[224],clsize),rep(cl[225],clsize),rep(cl[226],clsize),rep(cl[227],clsize),rep(cl[228
],clsize),rep(cl[229],clsize),rep(cl[230],clsize),
rep(cl[231],clsize),rep(cl[232],clsize),rep(cl[233],clsize),rep(cl[234],clsize),rep(cl[235],clsize),rep(cl[236],clsize),rep(cl[237],clsize),rep(cl[238
],clsize),rep(cl[239],clsize),rep(cl[240],clsize),
rep(cl[241],clsize),rep(cl[242],clsize),rep(cl[243],clsize),rep(cl[244],clsize),rep(cl[245],clsize),rep(cl[246],clsize),rep(cl[247],clsize),rep(cl[248
],clsize),rep(cl[249],clsize),rep(cl[250],clsize),
rep(cl[251],clsize),rep(cl[252],clsize),rep(cl[253],clsize),rep(cl[254],clsize),rep(cl[255],clsize),rep(cl[256],clsize),rep(cl[257],clsize),rep(cl[258
],clsize),rep(cl[259],clsize),rep(cl[260],clsize),
rep(cl[261],clsize),rep(cl[262],clsize),rep(cl[263],clsize),rep(cl[264],clsize),rep(cl[265],clsize),rep(cl[266],clsize),rep(cl[267],clsize),rep(cl[268
],clsize),rep(cl[269],clsize),rep(cl[270],clsize),
rep(cl[271],clsize),rep(cl[272],clsize),rep(cl[273],clsize),rep(cl[274],clsize),rep(cl[275],clsize),rep(cl[276],clsize),rep(cl[277],clsize),rep(cl[278
],clsize),rep(cl[279],clsize),rep(cl[280],clsize),
rep(cl[281],clsize),rep(cl[282],clsize),rep(cl[283],clsize),rep(cl[284],clsize),rep(cl[285],clsize),rep(cl[286],clsize),rep(cl[287],clsize),rep(cl[288
],clsize),rep(cl[289],clsize),rep(cl[290],clsize),
rep(cl[291],clsize),rep(cl[292],clsize),rep(cl[293],clsize),rep(cl[294],clsize),rep(cl[295],clsize),rep(cl[296],clsize),rep(cl[297],clsize),rep(cl[298
],clsize),rep(cl[299],clsize),rep(cl[300],clsize),
```



```

popx1<-data.frame(cl_id,x)
popx1$unit<-1:nrow(popx1)

popx <- data.frame(popx1$unit,cl_id,mu,v_j,e_ij,x)
colnames(popx)<-c('unit_id','cl_id','mu','v_j','e_ij','x')

pop.des=svydesign(data=popx,ids=~cl_id+unit_id, nest=TRUE)
pop.mean=data.frame(svymean(x,pop.des))
popmean <- pop.mean[1,1]#finite population mean

##### Weights storage #####

for (k in 1:nvar){
for (j in 1:nsamp){

w1<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1) #store weights
w2<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w3<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w4<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w5<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w6<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w7<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w8<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1) #store weights
w9<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w10<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w11<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w12<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
}

#####
#####
# Two stage Cluster Random Sampling from a clustered finite population
#####

n2 = rep(20,50)    # select 20 units from 50 clusters

for (i in 1:nsamp) {

sample = mstage(popx, stage=list("cluster","cluster"), varnames=list("cl_id", "unit_id"),size=list(500,n2), method=list("srswor","srswor"))
#unique(sample[[1]]$cl_id)
#unique(sample[[2]]$unit_id)  #double checking

s=getdata(popx,sample)[[2]]

#####
# Naive weights
#####

mean_naive[i]=mean(s$x)
variance_naive[i]=var(s$x)
cv_naive[i]=abs(sqrt(variance_naive[i])/mean_naive[i])
skewness_naive[i]<-skewness(s$x)
kurtosis_naive[i]<-kurtosis(s$x)

#####
# Design conscious weights
#####

plan=svydesign(data=s,ids=~cl_id+unit_id, nest=TRUE)
mean_st=svymean(~x,plan)
mean_st=as.data.frame(mean_st)

means[i]<-mean_st[1,1]          #sample mean for ith sample
variance[i]<-mean_st[1,2]^2     #sample variance, design consistent variance
cv[i]<-abs(sqrt(variance[i])/means[i])    #coefficient of variation

# skewness[i]<-skewness(s[,i])
moments = svymean(~l(x^4) + l(x^3) +

```

```

l(x^2) + l(x), plan)

skew=data.frame(svycontrast(moments,
  quote('l(x^3)` - 3*l(x^2)`*l(x)` +
    3*l(x)`*l(x)^2 - `l(x)^`^3)/
    ('l(x^2)` - `l(x)^`^2^1.5)))
skewness[i]<-skew[1,1]

#kurtosis[i]<-kurtosis(s[,i])

kurt=data.frame(svycontrast(moments,
  quote('l(x^4)` -
    4*l(x^3)`*l(x)` +
    6*l(x^2)`*l(x)^2 -
    4*l(x)`*l(x)^3 +
    `l(x)^`^4)/
    ('l(x^2)` -
    `l(x)^`^2^2)))

kurtosis[i]<-kurt[1,1]

meff.1 <- data.frame(svymean(s$x,plan,deff=TRUE))
meff_var[i]=variance[i]/((1-(n/N))*(var(s$x)/n))

#####Weights calculation#####

w1[i]=n      #inverse of the sample size

w2[i]<-1/(variance[i]) #inverse of the design variance of the estimator of the mean
w3[i]<-1/(variance_naive[i]) #inverse of the naive variance of the estimator of the mean

w4[i]=1      #simple average weights

w5[i]=1/cv[i] #inverse of cv
w6[i]=1/cv_naive[i] #inverse of naive cv

w7[i]=abs(1/skewness[i]) #inverse skewness
w8[i]=abs(1/skewness_naive[i]) #inverse skewness

w9[i]=1/kurtosis[i] #inverse kurtosis
w10[i]=1/kurtosis_naive[i] #inverse kurtosis

w11[i]=meff_var[i]/variance[i] #in srs deff=1 so coincides with inverse of variance
w12[i]=1/meff_var[i]

combmeans1[i]<-round(sum(means[1:i]*w1[1:i])/sum(w1[1:i]),digits=9) #combine means using inverse sample size
combmeans2[i]<-round(sum(means[1:i]*w2[1:i])/sum(w2[1:i]),digits=9) # combine using inverse variance
combmeans3[i]<-round(sum(means[1:i]*w3[1:i])/sum(w3[1:i]),digits=9) # combine using inverse variance naive
combmeans4[i]<-round(sum(means[1:i]*w4[1:i])/sum(w4[1:i]),digits=9) # combine using a weight of 1
combmeans5[i]<-round(sum(means[1:i]*w5[1:i])/sum(w5[1:i]),digits=9) # combine using a weight of inverse CV
combmeans6[i]<-round(sum(means[1:i]*w6[1:i])/sum(w6[1:i]),digits=9) # combine using a weight of inverse CV naive
combmeans7[i]<-round(sum(means[1:i]*w7[1:i])/sum(w7[1:i]),digits=9) # combine using a weight of inverse skewness
combmeans8[i]<-round(sum(means[1:i]*w8[1:i])/sum(w8[1:i]),digits=9) #combine using a weight of inverse skewness naive
combmeans9[i]<-round(sum(means[1:i]*w9[1:i])/sum(w9[1:i]),digits=9) # combine using inverse kurtosis
combmeans10[i]<-round(sum(means[1:i]*w10[1:i])/sum(w10[1:i]),digits=9) # combine using inverse kurtosis naive
combmeans11[i]<-round(sum(means[1:i]*w11[1:i])/sum(w11[1:i]),digits=9) # combine using a weight of meff/var
combmeans12[i]<-round(sum(means[1:i]*w12[1:i])/sum(w12[1:i]),digits=9) # combine using a weight of 1/meff

}

sampmean[k]=means[1] # capturing the combined mean, where nsamp reps are combined, all weighting strategies

combmeans11[k]=combmeans1[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans21[k]=combmeans2[2]
combmeans31[k]=combmeans3[2]

```

combmeans41[k]=combmeans4[2]
combmeans51[k]=combmeans5[2]
combmeans61[k]=combmeans6[2]
combmeans71[k]=combmeans7[2]
combmeans81[k]=combmeans8[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans91[k]=combmeans9[2]
combmeans101[k]=combmeans10[2]
combmeans111[k]=combmeans11[2]
combmeans121[k]=combmeans12[2]

combmeans12[k]=combmeans1[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans22[k]=combmeans2[3]
combmeans32[k]=combmeans3[3]
combmeans42[k]=combmeans4[3]
combmeans52[k]=combmeans5[3]
combmeans62[k]=combmeans6[3]
combmeans72[k]=combmeans7[3]
combmeans82[k]=combmeans8[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans92[k]=combmeans9[3]
combmeans102[k]=combmeans10[3]
combmeans112[k]=combmeans11[3]
combmeans122[k]=combmeans12[3]

combmeans13[k]=combmeans1[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans23[k]=combmeans2[5]
combmeans33[k]=combmeans3[5]
combmeans43[k]=combmeans4[5]
combmeans53[k]=combmeans5[5]
combmeans63[k]=combmeans6[5]
combmeans73[k]=combmeans7[5]
combmeans83[k]=combmeans8[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans93[k]=combmeans9[5]
combmeans103[k]=combmeans10[5]
combmeans113[k]=combmeans11[5]
combmeans123[k]=combmeans12[5]

combmeans14[k]=combmeans1[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans24[k]=combmeans2[10]
combmeans34[k]=combmeans3[10]
combmeans44[k]=combmeans4[10]
combmeans54[k]=combmeans5[10]
combmeans64[k]=combmeans6[10]
combmeans74[k]=combmeans7[10]
combmeans84[k]=combmeans8[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans94[k]=combmeans9[10]
combmeans104[k]=combmeans10[10]
combmeans114[k]=combmeans11[10]
combmeans124[k]=combmeans12[10]

combmeans15[k]=combmeans1[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans25[k]=combmeans2[15]
combmeans35[k]=combmeans3[15]
combmeans45[k]=combmeans4[15]
combmeans55[k]=combmeans5[15]
combmeans65[k]=combmeans6[15]
combmeans75[k]=combmeans7[15]
combmeans85[k]=combmeans8[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans95[k]=combmeans9[15]
combmeans105[k]=combmeans10[15]
combmeans115[k]=combmeans11[15]
combmeans125[k]=combmeans12[15]

combmeans16[k]=combmeans1[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans26[k]=combmeans2[20]
combmeans36[k]=combmeans3[20]
combmeans46[k]=combmeans4[20]

```

combmeans56[k]=combmeans5[20]
combmeans66[k]=combmeans6[20]
combmeans76[k]=combmeans7[20]
combmeans86[k]=combmeans8[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans96[k]=combmeans9[20]
combmeans106[k]=combmeans10[20]
combmeans116[k]=combmeans11[20]
combmeans126[k]=combmeans12[20]

}
expected1=mean(sampmean)      #expected mean for one sample, not combining

expected=matrix(0,12,6)

expected[1,1]=mean(combmeans11) #strategy 1 nsamp=2
expected[2,1]=mean(combmeans21)
expected[3,1]=mean(combmeans31)
expected[4,1]=mean(combmeans41)
expected[5,1]=mean(combmeans51)
expected[6,1]=mean(combmeans61)
expected[7,1]=mean(combmeans71)
expected[8,1]=mean(combmeans81)
expected[9,1]=mean(combmeans91)
expected[10,1]=mean(combmeans101)
expected[11,1]=mean(combmeans111)
expected[12,1]=mean(combmeans121)

expected[1,2]=mean(combmeans12) #strategy 1 nsamp=3
expected[2,2]=mean(combmeans22)
expected[3,2]=mean(combmeans32)
expected[4,2]=mean(combmeans42)
expected[5,2]=mean(combmeans52)
expected[6,2]=mean(combmeans62)
expected[7,2]=mean(combmeans72)
expected[8,2]=mean(combmeans82)
expected[9,2]=mean(combmeans92)
expected[10,2]=mean(combmeans102)
expected[11,2]=mean(combmeans112)
expected[12,2]=mean(combmeans122)

expected[1,3]=mean(combmeans13) #strategy 1 nsamp=5
expected[2,3]=mean(combmeans23)
expected[3,3]=mean(combmeans33)
expected[4,3]=mean(combmeans43)
expected[5,3]=mean(combmeans53)
expected[6,3]=mean(combmeans63)
expected[7,3]=mean(combmeans73)
expected[8,3]=mean(combmeans83)
expected[9,3]=mean(combmeans93)
expected[10,3]=mean(combmeans103)
expected[11,3]=mean(combmeans113)
expected[12,3]=mean(combmeans123)

expected[1,4]=mean(combmeans14) #strategy 1 nsamp=10
expected[2,4]=mean(combmeans24)
expected[3,4]=mean(combmeans34)
expected[4,4]=mean(combmeans44)
expected[5,4]=mean(combmeans54)
expected[6,4]=mean(combmeans64)
expected[7,4]=mean(combmeans74)
expected[8,4]=mean(combmeans84)
expected[9,4]=mean(combmeans94)
expected[10,4]=mean(combmeans104)
expected[11,4]=mean(combmeans114)
expected[12,4]=mean(combmeans124)

expected[1,5]=mean(combmeans15) #strategy 1 nsamp=15
expected[2,5]=mean(combmeans25)
expected[3,5]=mean(combmeans35)
expected[4,5]=mean(combmeans45)
expected[5,5]=mean(combmeans55)

```

```

expected[6,5]=mean(combmeans65)
expected[7,5]=mean(combmeans75)
expected[8,5]=mean(combmeans85)
expected[9,5]=mean(combmeans95)
expected[10,5]=mean(combmeans105)
expected[11,5]=mean(combmeans115)
expected[12,5]=mean(combmeans125)

```

```

expected[1,6]=mean(combmeans16) #strategy 1 nsamp=20
expected[2,6]=mean(combmeans26)
expected[3,6]=mean(combmeans36)
expected[4,6]=mean(combmeans46)
expected[5,6]=mean(combmeans56)
expected[6,6]=mean(combmeans66)
expected[7,6]=mean(combmeans76)
expected[8,6]=mean(combmeans86)
expected[9,6]=mean(combmeans96)
expected[10,6]=mean(combmeans106)
expected[11,6]=mean(combmeans116)
expected[12,6]=mean(combmeans126)

```

```

expectedcol=c(expected1,expected[,1],expected[,2],expected[,3],expected[,4],expected[,5],expected[,6])

```

```

means=cbind(combmeans11,combmeans21,
combmeans31,combmeans41,combmeans51,combmeans61,combmeans71,combmeans81,combmeans91,
combmeans101,combmeans111,combmeans121,
combmeans12,combmeans22, combmeans32,combmeans42,combmeans52,combmeans62,combmeans72,combmeans82,combmeans92,
combmeans102,combmeans112,combmeans122,
combmeans13,combmeans23, combmeans33,combmeans43,combmeans53,combmeans63,combmeans73,combmeans83,combmeans93,
combmeans103,combmeans113,combmeans123,
combmeans14,combmeans24, combmeans34,combmeans44,combmeans54,combmeans64,combmeans74,combmeans84,combmeans94,
combmeans104,combmeans114,combmeans124,
combmeans15,combmeans25, combmeans35,combmeans45,combmeans55,combmeans65,combmeans75,combmeans85,combmeans95,
combmeans105,combmeans115,combmeans125,
combmeans16,combmeans26, combmeans36,combmeans46,combmeans56,combmeans66,combmeans76,combmeans86,combmeans96,
combmeans106,combmeans116,combmeans126)

```

```

variance1=(1/nvar)*sum((sampmean[,1]-expected1)**2)

```

```

variance11=(1/nvar)*sum((means[,1]-expected[1,1])**2)
variance21=(1/nvar)*sum((means[,2]-expected[2,1])**2)
variance31=(1/nvar)*sum((means[,3]-expected[3,1])**2)
variance41=(1/nvar)*sum((means[,4]-expected[4,1])**2)
variance51=(1/nvar)*sum((means[,5]-expected[5,1])**2)
variance61=(1/nvar)*sum((means[,6]-expected[6,1])**2)
variance71=(1/nvar)*sum((means[,7]-expected[7,1])**2)
variance81=(1/nvar)*sum((means[,8]-expected[8,1])**2)
variance91=(1/nvar)*sum((means[,9]-expected[9,1])**2)
variance101=(1/nvar)*sum((means[,10]-expected[10,1])**2)
variance111=(1/nvar)*sum((means[,11]-expected[11,1])**2)
variance121=(1/nvar)*sum((means[,12]-expected[12,1])**2)

```

```

variance12=(1/nvar)*sum((means[,13]-expected[1,2])**2)
variance22=(1/nvar)*sum((means[,14]-expected[2,2])**2)
variance32=(1/nvar)*sum((means[,15]-expected[3,2])**2)
variance42=(1/nvar)*sum((means[,16]-expected[4,2])**2)
variance52=(1/nvar)*sum((means[,17]-expected[5,2])**2)
variance62=(1/nvar)*sum((means[,18]-expected[6,2])**2)
variance72=(1/nvar)*sum((means[,19]-expected[7,2])**2)
variance82=(1/nvar)*sum((means[,20]-expected[8,2])**2)
variance92=(1/nvar)*sum((means[,21]-expected[9,2])**2)
variance102=(1/nvar)*sum((means[,22]-expected[10,2])**2)
variance112=(1/nvar)*sum((means[,23]-expected[11,2])**2)
variance122=(1/nvar)*sum((means[,24]-expected[12,2])**2)

```

```

variance13=(1/nvar)*sum((means[,25]-expected[1,3])**2)
variance23=(1/nvar)*sum((means[,26]-expected[2,3])**2)
variance33=(1/nvar)*sum((means[,27]-expected[3,3])**2)
variance43=(1/nvar)*sum((means[,28]-expected[4,3])**2)
variance53=(1/nvar)*sum((means[,29]-expected[5,3])**2)

```

$\text{variance63}=(1/\text{nvar})*\text{sum}((\text{means}[30]-\text{expected}[6,3])**2)$
 $\text{variance73}=(1/\text{nvar})*\text{sum}((\text{means}[31]-\text{expected}[7,3])**2)$
 $\text{variance83}=(1/\text{nvar})*\text{sum}((\text{means}[32]-\text{expected}[8,3])**2)$
 $\text{variance93}=(1/\text{nvar})*\text{sum}((\text{means}[33]-\text{expected}[9,3])**2)$
 $\text{variance103}=(1/\text{nvar})*\text{sum}((\text{means}[34]-\text{expected}[10,3])**2)$
 $\text{variance113}=(1/\text{nvar})*\text{sum}((\text{means}[35]-\text{expected}[11,3])**2)$
 $\text{variance123}=(1/\text{nvar})*\text{sum}((\text{means}[36]-\text{expected}[12,3])**2)$

$\text{variance14}=(1/\text{nvar})*\text{sum}((\text{means}[37]-\text{expected}[1,4])**2)$
 $\text{variance24}=(1/\text{nvar})*\text{sum}((\text{means}[38]-\text{expected}[2,4])**2)$
 $\text{variance34}=(1/\text{nvar})*\text{sum}((\text{means}[39]-\text{expected}[3,4])**2)$
 $\text{variance44}=(1/\text{nvar})*\text{sum}((\text{means}[40]-\text{expected}[4,4])**2)$
 $\text{variance54}=(1/\text{nvar})*\text{sum}((\text{means}[41]-\text{expected}[5,4])**2)$
 $\text{variance64}=(1/\text{nvar})*\text{sum}((\text{means}[42]-\text{expected}[6,4])**2)$
 $\text{variance74}=(1/\text{nvar})*\text{sum}((\text{means}[43]-\text{expected}[7,4])**2)$
 $\text{variance84}=(1/\text{nvar})*\text{sum}((\text{means}[44]-\text{expected}[8,4])**2)$
 $\text{variance94}=(1/\text{nvar})*\text{sum}((\text{means}[45]-\text{expected}[9,4])**2)$
 $\text{variance104}=(1/\text{nvar})*\text{sum}((\text{means}[46]-\text{expected}[10,4])**2)$
 $\text{variance114}=(1/\text{nvar})*\text{sum}((\text{means}[47]-\text{expected}[11,4])**2)$
 $\text{variance124}=(1/\text{nvar})*\text{sum}((\text{means}[48]-\text{expected}[12,4])**2)$

$\text{variance15}=(1/\text{nvar})*\text{sum}((\text{means}[49]-\text{expected}[1,5])**2)$
 $\text{variance25}=(1/\text{nvar})*\text{sum}((\text{means}[50]-\text{expected}[2,5])**2)$
 $\text{variance35}=(1/\text{nvar})*\text{sum}((\text{means}[51]-\text{expected}[3,5])**2)$
 $\text{variance45}=(1/\text{nvar})*\text{sum}((\text{means}[52]-\text{expected}[4,5])**2)$
 $\text{variance55}=(1/\text{nvar})*\text{sum}((\text{means}[53]-\text{expected}[5,5])**2)$
 $\text{variance65}=(1/\text{nvar})*\text{sum}((\text{means}[54]-\text{expected}[6,5])**2)$
 $\text{variance75}=(1/\text{nvar})*\text{sum}((\text{means}[55]-\text{expected}[7,5])**2)$
 $\text{variance85}=(1/\text{nvar})*\text{sum}((\text{means}[56]-\text{expected}[8,5])**2)$
 $\text{variance95}=(1/\text{nvar})*\text{sum}((\text{means}[57]-\text{expected}[9,5])**2)$
 $\text{variance105}=(1/\text{nvar})*\text{sum}((\text{means}[58]-\text{expected}[10,5])**2)$
 $\text{variance115}=(1/\text{nvar})*\text{sum}((\text{means}[59]-\text{expected}[11,5])**2)$
 $\text{variance125}=(1/\text{nvar})*\text{sum}((\text{means}[60]-\text{expected}[12,5])**2)$

$\text{variance16}=(1/\text{nvar})*\text{sum}((\text{means}[61]-\text{expected}[1,6])**2)$
 $\text{variance26}=(1/\text{nvar})*\text{sum}((\text{means}[62]-\text{expected}[2,6])**2)$
 $\text{variance36}=(1/\text{nvar})*\text{sum}((\text{means}[63]-\text{expected}[3,6])**2)$
 $\text{variance46}=(1/\text{nvar})*\text{sum}((\text{means}[64]-\text{expected}[4,6])**2)$
 $\text{variance56}=(1/\text{nvar})*\text{sum}((\text{means}[65]-\text{expected}[5,6])**2)$
 $\text{variance66}=(1/\text{nvar})*\text{sum}((\text{means}[66]-\text{expected}[6,6])**2)$
 $\text{variance76}=(1/\text{nvar})*\text{sum}((\text{means}[67]-\text{expected}[7,6])**2)$
 $\text{variance86}=(1/\text{nvar})*\text{sum}((\text{means}[68]-\text{expected}[8,6])**2)$
 $\text{variance96}=(1/\text{nvar})*\text{sum}((\text{means}[69]-\text{expected}[9,6])**2)$
 $\text{variance106}=(1/\text{nvar})*\text{sum}((\text{means}[70]-\text{expected}[10,6])**2)$
 $\text{variance116}=(1/\text{nvar})*\text{sum}((\text{means}[71]-\text{expected}[11,6])**2)$
 $\text{variance126}=(1/\text{nvar})*\text{sum}((\text{means}[72]-\text{expected}[12,6])**2)$

$\text{variance}=\text{c}(\text{variance1},\text{variance11},\text{variance21},\text{variance31},\text{variance41},\text{variance51},$
 $\text{variance61},\text{variance71},\text{variance81},\text{variance91},\text{variance101},\text{variance111},\text{variance121},$
 $\text{variance12},\text{variance22},\text{variance32},\text{variance42},\text{variance52},$
 $\text{variance62},\text{variance72},\text{variance82},\text{variance92},\text{variance102},\text{variance112},\text{variance122},$
 $\text{variance13},\text{variance23},\text{variance33},\text{variance43},\text{variance53},$
 $\text{variance63},\text{variance73},\text{variance83},\text{variance93},\text{variance103},\text{variance113},\text{variance123},$
 $\text{variance14},\text{variance24},\text{variance34},\text{variance44},\text{variance54},$
 $\text{variance64},\text{variance74},\text{variance84},\text{variance94},\text{variance104},\text{variance114},\text{variance124},$
 $\text{variance15},\text{variance25},\text{variance35},\text{variance45},\text{variance55},$
 $\text{variance65},\text{variance75},\text{variance85},\text{variance95},\text{variance105},\text{variance115},\text{variance125},$
 $\text{variance16},\text{variance26},\text{variance36},\text{variance46},\text{variance56},$
 $\text{variance66},\text{variance76},\text{variance86},\text{variance96},\text{variance106},\text{variance116},\text{variance126})$

$\text{bias}=\text{expectedcol}-\text{popmean}$

$\text{relbias}=(\text{bias}/\text{expectedcol})*100$

$\text{mse}=\text{variance}+\text{bias}**2$

$\text{cv}=(\text{sqrt}(\text{variance})/\text{expectedcol})*100$

$\text{s.e}=(\text{sqrt}(\text{variance})/\text{sqrt}(\text{nvar}))$

```
LB= expectedcol - 1.96*s.e
UB= expectedcol + 1.96*s.e

result=cbind(popmean,expectedcol,relbias,s.e,cv,LB,UB,mse)

write.table(result,"result_scenario1_100.txt",sep="\t")
write.table(meff_var,"meff_scenario1_100.txt",sep="\t")
```

H2 T Distribution, 3 degrees of freedom

```
#####
# Variance Estimation: Normal distribution
# author: "Loveness Dzikiti"
# date: "20 April 2017"
#####

#####
# Scenario 1: N=1000000 N(100,25) n=1000 D=20
#####
# Generating a clustered population
# X_ij= mu + v_j + e_ij

rm(list=ls())

setwd("C:\\Users\\Loveness\\Dropbox\\Review\\simulations_2019\\cluster\\t3")

library(moments)
library(sampling)
library(survey)

N=1000000      #finite population size
n=1000         #sample size
nsamp=20       #number of samplesclsize=250
clsiz=250
k=nvar=100

set.seed(123456)

#####storage matrices#####

mean_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample means calculated as srs
variance_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample variance calculated as srs
cv_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample coefficient of variation calculated as srs
skewness_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample skewness calculated as srs
kurtosis_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample kurtosis calculated as srs

means <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample means
variance <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample variance
designvar <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store design consistent variance
cv <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample coefficient of variation
skewness <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample skewness
kurtosis <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample kurtosis
meff_var<-matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample design effect

combmeans1 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 1
combmeans2 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 2
combmeans3 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 3
combmeans4 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 4
combmeans5 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 5
combmeans6 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 6
combmeans7 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 7
combmeans8 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 8
combmeans9 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 9
combmeans10<- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 10
combmeans11<- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 11
combmeans12<- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 12

sampmean <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store mean of 1 sample, no combining

combmeans11 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
combmeans21 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
combmeans31 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans41 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans51 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5
combmeans61 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 6
combmeans71 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 7
combmeans81 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1
```



```

combmeans106 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3
combmeans116 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans126 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5

```

```
#####
```

```
### Generating a clustered population#####
```

```

v=3
r=2
nclus=4000

```

```

mu1=0
sigma1=sqrt(20)

```

```

mu2=0
sigma2=sqrt(5)

```

```
Var<-v^(r/2)*((gamma(1/2*(r+1))*gamma(1/2*(v-r)))/(gamma(0.5)*gamma(1/2*v))) # variance of t-distribution
```

```

x1<-rt(N,v) #simulate data from a t distribution random mean and variance
x1<-x1/sqrt(Var)

```

```

e_jj=(x1*sigma1)+mu1
#####

```

```

x2<-rt(nclus,v)
x2<-x2/sqrt(Var)

```

```
cl=(x2*sigma2)+mu2
```

```
mu <- matrix(c(rep(100,N)), nrow = N, ncol = 1) #store superpopulation mean
```

```

v_j <-
c(rep(cl[1],clsiz),rep(cl[2],clsiz),rep(cl[3],clsiz),rep(cl[4],clsiz),rep(cl[5],clsiz),rep(cl[6],clsiz),rep(cl[7],clsiz),rep(cl[8],clsiz),rep(cl[9],clsiz),rep(cl[10],clsiz),
rep(cl[11],clsiz),rep(cl[12],clsiz),rep(cl[13],clsiz),rep(cl[14],clsiz),rep(cl[15],clsiz),rep(cl[16],clsiz),rep(cl[17],clsiz),rep(cl[18],clsiz),rep(cl[19],clsiz),rep(cl[20],clsiz),
rep(cl[21],clsiz),rep(cl[22],clsiz),rep(cl[23],clsiz),rep(cl[24],clsiz),rep(cl[25],clsiz),rep(cl[26],clsiz),rep(cl[27],clsiz),rep(cl[28],clsiz),rep(cl[29],clsiz),rep(cl[30],clsiz),
rep(cl[31],clsiz),rep(cl[32],clsiz),rep(cl[33],clsiz),rep(cl[34],clsiz),rep(cl[35],clsiz),rep(cl[36],clsiz),rep(cl[37],clsiz),rep(cl[38],clsiz),rep(cl[39],clsiz),rep(cl[40],clsiz),
rep(cl[41],clsiz),rep(cl[42],clsiz),rep(cl[43],clsiz),rep(cl[44],clsiz),rep(cl[45],clsiz),rep(cl[46],clsiz),rep(cl[47],clsiz),rep(cl[48],clsiz),rep(cl[49],clsiz),rep(cl[50],clsiz),
rep(cl[51],clsiz),rep(cl[52],clsiz),rep(cl[53],clsiz),rep(cl[54],clsiz),rep(cl[55],clsiz),rep(cl[56],clsiz),rep(cl[57],clsiz),rep(cl[58],clsiz),rep(cl[59],clsiz),rep(cl[60],clsiz),
rep(cl[61],clsiz),rep(cl[62],clsiz),rep(cl[63],clsiz),rep(cl[64],clsiz),rep(cl[65],clsiz),rep(cl[66],clsiz),rep(cl[67],clsiz),rep(cl[68],clsiz),rep(cl[69],clsiz),rep(cl[70],clsiz),
rep(cl[71],clsiz),rep(cl[72],clsiz),rep(cl[73],clsiz),rep(cl[74],clsiz),rep(cl[75],clsiz),rep(cl[76],clsiz),rep(cl[77],clsiz),rep(cl[78],clsiz),rep(cl[79],clsiz),rep(cl[80],clsiz),
rep(cl[81],clsiz),rep(cl[82],clsiz),rep(cl[83],clsiz),rep(cl[84],clsiz),rep(cl[85],clsiz),rep(cl[86],clsiz),rep(cl[87],clsiz),rep(cl[88],clsiz),rep(cl[89],clsiz),rep(cl[90],clsiz),
rep(cl[91],clsiz),rep(cl[92],clsiz),rep(cl[93],clsiz),rep(cl[94],clsiz),rep(cl[95],clsiz),rep(cl[96],clsiz),rep(cl[97],clsiz),rep(cl[98],clsiz),rep(cl[99],clsiz),rep(cl[100],clsiz),
rep(cl[101],clsiz),rep(cl[102],clsiz),rep(cl[103],clsiz),rep(cl[104],clsiz),rep(cl[105],clsiz),rep(cl[106],clsiz),rep(cl[107],clsiz),rep(cl[108],clsiz),rep(cl[109],clsiz),rep(cl[110],clsiz),
rep(cl[111],clsiz),rep(cl[112],clsiz),rep(cl[113],clsiz),rep(cl[114],clsiz),rep(cl[115],clsiz),rep(cl[116],clsiz),rep(cl[117],clsiz),rep(cl[118],clsiz),rep(cl[119],clsiz),rep(cl[120],clsiz),
rep(cl[121],clsiz),rep(cl[122],clsiz),rep(cl[123],clsiz),rep(cl[124],clsiz),rep(cl[125],clsiz),rep(cl[126],clsiz),rep(cl[127],clsiz),rep(cl[128],clsiz),rep(cl[129],clsiz),rep(cl[130],clsiz),
rep(cl[131],clsiz),rep(cl[132],clsiz),rep(cl[133],clsiz),rep(cl[134],clsiz),rep(cl[135],clsiz),rep(cl[136],clsiz),rep(cl[137],clsiz),rep(cl[138],clsiz),rep(cl[139],clsiz),rep(cl[140],clsiz),
rep(cl[141],clsiz),rep(cl[142],clsiz),rep(cl[143],clsiz),rep(cl[144],clsiz),rep(cl[145],clsiz),rep(cl[146],clsiz),rep(cl[147],clsiz),rep(cl[148],clsiz),rep(cl[149],clsiz),rep(cl[150],clsiz),
rep(cl[151],clsiz),rep(cl[152],clsiz),rep(cl[153],clsiz),rep(cl[154],clsiz),rep(cl[155],clsiz),rep(cl[156],clsiz),rep(cl[157],clsiz),rep(cl[158],clsiz),rep(cl[159],clsiz),rep(cl[160],clsiz),
rep(cl[161],clsiz),rep(cl[162],clsiz),rep(cl[163],clsiz),rep(cl[164],clsiz),rep(cl[165],clsiz),rep(cl[166],clsiz),rep(cl[167],clsiz),rep(cl[168],clsiz),rep(cl[169],clsiz),rep(cl[170],clsiz),
rep(cl[171],clsiz),rep(cl[172],clsiz),rep(cl[173],clsiz),rep(cl[174],clsiz),rep(cl[175],clsiz),rep(cl[176],clsiz),rep(cl[177],clsiz),rep(cl[178],clsiz),rep(cl[179],clsiz),rep(cl[180],clsiz),

```



```

rep('3901',clsiz),rep('3902',clsiz),rep('3903',clsiz),rep('3904',clsiz),rep('3905',clsiz),rep('3906',clsiz),rep('3907',clsiz),rep('3908',clsiz
),rep('3909',clsiz),rep('3910',clsiz),
rep('3911',clsiz),rep('3912',clsiz),rep('3913',clsiz),rep('3914',clsiz),rep('3915',clsiz),rep('3916',clsiz),rep('3917',clsiz),rep('3918',clsiz
),rep('3919',clsiz),rep('3920',clsiz),
rep('3921',clsiz),rep('3922',clsiz),rep('3923',clsiz),rep('3924',clsiz),rep('3925',clsiz),rep('3926',clsiz),rep('3927',clsiz),rep('3928',clsiz
),rep('3929',clsiz),rep('3930',clsiz),
rep('3931',clsiz),rep('3932',clsiz),rep('3933',clsiz),rep('3934',clsiz),rep('3935',clsiz),rep('3936',clsiz),rep('3937',clsiz),rep('3938',clsiz
),rep('3939',clsiz),rep('3940',clsiz),
rep('3941',clsiz),rep('3942',clsiz),rep('3943',clsiz),rep('3944',clsiz),rep('3945',clsiz),rep('3946',clsiz),rep('3947',clsiz),rep('3948',clsiz
),rep('3949',clsiz),rep('3950',clsiz),
rep('3951',clsiz),rep('3952',clsiz),rep('3953',clsiz),rep('3954',clsiz),rep('3955',clsiz),rep('3956',clsiz),rep('3957',clsiz),rep('3958',clsiz
),rep('3959',clsiz),rep('3960',clsiz),
rep('3961',clsiz),rep('3962',clsiz),rep('3963',clsiz),rep('3964',clsiz),rep('3965',clsiz),rep('3966',clsiz),rep('3967',clsiz),rep('3968',clsiz
),rep('3969',clsiz),rep('3970',clsiz),
rep('3971',clsiz),rep('3972',clsiz),rep('3973',clsiz),rep('3974',clsiz),rep('3975',clsiz),rep('3976',clsiz),rep('3977',clsiz),rep('3978',clsiz
),rep('3979',clsiz),rep('3980',clsiz),
rep('3981',clsiz),rep('3982',clsiz),rep('3983',clsiz),rep('3984',clsiz),rep('3985',clsiz),rep('3986',clsiz),rep('3987',clsiz),rep('3988',clsiz
),rep('3989',clsiz),rep('3990',clsiz),
rep('3991',clsiz),rep('3992',clsiz),rep('3993',clsiz),rep('3994',clsiz),rep('3995',clsiz),rep('3996',clsiz),rep('3997',clsiz),rep('3998',clsiz
),rep('3999',clsiz),rep('4000',clsiz)

```

```
)
```

```
popx1<-data.frame(cl_id,x)
popx1$unit<-1:nrow(popx1)
```

```
popx <- data.frame(popx1$unit,cl_id,mu,v_j,e_ij,x)
colnames(popx)<-c('unit_id','cl_id','mu','v_j','e_ij','x')
```

```
pop.des=svydesign(data=popx,ids=~cl_id+unit_id, nest=TRUE)
pop.mean=data.frame(svymean(x,pop.des))
popmean <- pop.mean[1,1]#finite population mean
```

```
##### Weights storage #####
for (k in 1:nvar){
```

```
for (j in 1:nsamp){
```

```

w1<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1) #store weights
w2<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w3<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w4<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w5<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w6<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w7<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w8<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1) #store weights
w9<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w10<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w11<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w12<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
}

```

```
#####
```

```
#####
# Two stage Cluster Random Sampling from a clustered finite population
#####
```

```
n2 = rep(20,50) # select 20 units from 50 clusters
```

```
for (i in 1:nsamp) {
```

```

sample = mstage(popx, stage=list("cluster","cluster"), varnames=list("cl_id", "unit_id"),size=list(50,n2), method=list("srswor","srswor"))
#unique(sample[[1]]$cl_id)
#unique(sample[[2]]$unit_id) #double checking

```

```
s=getdata(popx,sample)[[2]]
```

```
#####
# Naive weights
#####

mean_naive[i]=mean(s$x)
variance_naive[i]=var(s$x)
cv_naive[i]=abs(sqrt(variance_naive[i])/mean_naive[i])
skewness_naive[i]<-skewness(s$x)
kurtosis_naive[i]<-kurtosis(s$x)

#####
# Design conscious weights
#####

plan=svydesign(data=s,ids=~cl_id+unit_id, nest=TRUE)
mean_st=svymean(~x,plan)
mean_st=as.data.frame(mean_st)

means[i]<-mean_st[1,1]          #sample mean for ith sample
variance[i]<-mean_st[1,2]^2     #sample variance, design consistent variance
cv[i]<-abs(sqrt(variance[i])/means[i]) #coefficient of variation

# skewness[i]<-skewness(s[,i])
moments = svymean(~l(x^4) + l(x^3) +
                 l(x^2) + l(x), plan)

skew=data.frame(svycontrast(moments,
                             quote(`l(x^3)` - 3*`l(x^2)`*`l(x)` +
                                     3*`l(x)`*`l(x)^2` - `l(x)^3`)/
                             (`l(x^2)` - `l(x)^2`^1.5)))
skewness[i]<-skew[1,1]

#kurtosis[i]<-kurtosis(s[,i])

kurt=data.frame(svycontrast(moments,
                             quote(`l(x^4)` -
                                     4*`l(x^3)`*`l(x)` +
                                     6*`l(x^2)`*`l(x)^2` -
                                     4*`l(x)`*`l(x)^3` +
                                     `l(x)^4`/
                                     (`l(x^2)` -
                                      `l(x)^2`^2)))
kurtosis[i]<-kurt[1,1]

meff.1 <- data.frame(svymean(s$x,plan,deff=TRUE))
meff_var[i]=variance[i]/((1-(n/N))*(var(s$x)/n))

#####Weights calculation#####

w1[i]=n          #inverse of the sample size

w2[i]<-1/(variance[i]) #inverse of the design variance of the estimator of the mean
w3[i]<-1/(variance_naive[i]) #inverse of the naive variance of the estimator of the mean

w4[i]=1          #simple average weights

w5[i]=1/cv[i]    #inverse of cv
w6[i]=1/cv_naive[i] #inverse of naive cv

w7[i]=abs(1/skewness[i]) #inverse skewness
w8[i]=abs(1/skewness_naive[i]) #inverse skewness

w9[i]=1/kurtosis[i] #inverse kurtosis
w10[i]=1/kurtosis_naive[i] #inverse kurtosis

w11[i]=meff_var[i]/variance[i] #in srs deff=1 so coincides with inverse of variance
w12[i]=1/meff_var[i]
```

```

combmeans1[i]<-round(sum(means[1:i]*w1[1:i])/sum(w1[1:i]),digits=9) #combine means using inverse sample size
combmeans2[i]<-round(sum(means[1:i]*w2[1:i])/sum(w2[1:i]),digits=9) # combine using inverse variance
combmeans3[i]<-round(sum(means[1:i]*w3[1:i])/sum(w3[1:i]),digits=9) # combine using inverse variance naive
combmeans4[i]<-round(sum(means[1:i]*w4[1:i])/sum(w4[1:i]),digits=9) # combine using a weight of 1
combmeans5[i]<-round(sum(means[1:i]*w5[1:i])/sum(w5[1:i]),digits=9) # combine using a weight of inverse CV
combmeans6[i]<-round(sum(means[1:i]*w6[1:i])/sum(w6[1:i]),digits=9) # combine using a weight of inverse CV naive
combmeans7[i]<-round(sum(means[1:i]*w7[1:i])/sum(w7[1:i]),digits=9) # combine using a weight of inverse skewness
combmeans8[i]<-round(sum(means[1:i]*w8[1:i])/sum(w8[1:i]),digits=9) #combine using a weight of inverse skewness naive
combmeans9[i]<-round(sum(means[1:i]*w9[1:i])/sum(w9[1:i]),digits=9) # combine using inverse kurtosis
combmeans10[i]<-round(sum(means[1:i]*w10[1:i])/sum(w10[1:i]),digits=9) # combine using inverse kurtosis naive
combmeans11[i]<-round(sum(means[1:i]*w11[1:i])/sum(w11[1:i]),digits=9) # combine using a weight of meff/var
combmeans12[i]<-round(sum(means[1:i]*w12[1:i])/sum(w12[1:i]),digits=9) # combine using a weight of 1/meff

```

```

}

```

```

sampmean[k]=means[1] # capturing the combined mean, where nsamp reps are combined, all weighting strategies

```

```

combmeans11[k]=combmeans1[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans21[k]=combmeans2[2]
combmeans31[k]=combmeans3[2]
combmeans41[k]=combmeans4[2]
combmeans51[k]=combmeans5[2]
combmeans61[k]=combmeans6[2]
combmeans71[k]=combmeans7[2]
combmeans81[k]=combmeans8[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans91[k]=combmeans9[2]
combmeans101[k]=combmeans10[2]
combmeans111[k]=combmeans11[2]
combmeans121[k]=combmeans12[2]

```

```

combmeans12[k]=combmeans1[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans22[k]=combmeans2[3]
combmeans32[k]=combmeans3[3]
combmeans42[k]=combmeans4[3]
combmeans52[k]=combmeans5[3]
combmeans62[k]=combmeans6[3]
combmeans72[k]=combmeans7[3]
combmeans82[k]=combmeans8[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans92[k]=combmeans9[3]
combmeans102[k]=combmeans10[3]
combmeans112[k]=combmeans11[3]
combmeans122[k]=combmeans12[3]

```

```

combmeans13[k]=combmeans1[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans23[k]=combmeans2[5]
combmeans33[k]=combmeans3[5]
combmeans43[k]=combmeans4[5]
combmeans53[k]=combmeans5[5]
combmeans63[k]=combmeans6[5]
combmeans73[k]=combmeans7[5]
combmeans83[k]=combmeans8[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans93[k]=combmeans9[5]
combmeans103[k]=combmeans10[5]
combmeans113[k]=combmeans11[5]
combmeans123[k]=combmeans12[5]

```

```

combmeans14[k]=combmeans1[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans24[k]=combmeans2[10]
combmeans34[k]=combmeans3[10]
combmeans44[k]=combmeans4[10]
combmeans54[k]=combmeans5[10]
combmeans64[k]=combmeans6[10]
combmeans74[k]=combmeans7[10]
combmeans84[k]=combmeans8[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies

```

```

combmeans94[k]=combmeans9[10]
combmeans104[k]=combmeans10[10]
combmeans114[k]=combmeans11[10]
combmeans124[k]=combmeans12[10]

combmeans15[k]=combmeans1[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans25[k]=combmeans2[15]
combmeans35[k]=combmeans3[15]
combmeans45[k]=combmeans4[15]
combmeans55[k]=combmeans5[15]
combmeans65[k]=combmeans6[15]
combmeans75[k]=combmeans7[15]
combmeans85[k]=combmeans8[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans95[k]=combmeans9[15]
combmeans105[k]=combmeans10[15]
combmeans115[k]=combmeans11[15]
combmeans125[k]=combmeans12[15]

combmeans16[k]=combmeans1[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans26[k]=combmeans2[20]
combmeans36[k]=combmeans3[20]
combmeans46[k]=combmeans4[20]
combmeans56[k]=combmeans5[20]
combmeans66[k]=combmeans6[20]
combmeans76[k]=combmeans7[20]
combmeans86[k]=combmeans8[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans96[k]=combmeans9[20]
combmeans106[k]=combmeans10[20]
combmeans116[k]=combmeans11[20]
combmeans126[k]=combmeans12[20]

}
expected1=mean(sampmean)      #expected mean for one sample, not combining

expected=matrix(0,12,6)

expected[1,1]=mean(combmeans11) #strategy 1 nsamp=2
expected[2,1]=mean(combmeans21)
expected[3,1]=mean(combmeans31)
expected[4,1]=mean(combmeans41)
expected[5,1]=mean(combmeans51)
expected[6,1]=mean(combmeans61)
expected[7,1]=mean(combmeans71)
expected[8,1]=mean(combmeans81)
expected[9,1]=mean(combmeans91)
expected[10,1]=mean(combmeans101)
expected[11,1]=mean(combmeans111)
expected[12,1]=mean(combmeans121)

expected[1,2]=mean(combmeans12) #strategy 1 nsamp=3
expected[2,2]=mean(combmeans22)
expected[3,2]=mean(combmeans32)
expected[4,2]=mean(combmeans42)
expected[5,2]=mean(combmeans52)
expected[6,2]=mean(combmeans62)
expected[7,2]=mean(combmeans72)
expected[8,2]=mean(combmeans82)
expected[9,2]=mean(combmeans92)
expected[10,2]=mean(combmeans102)
expected[11,2]=mean(combmeans112)
expected[12,2]=mean(combmeans122)

expected[1,3]=mean(combmeans13) #strategy 1 nsamp=5
expected[2,3]=mean(combmeans23)
expected[3,3]=mean(combmeans33)
expected[4,3]=mean(combmeans43)
expected[5,3]=mean(combmeans53)
expected[6,3]=mean(combmeans63)
expected[7,3]=mean(combmeans73)

```

```

expected[8,3]=mean(combmeans83)
expected[9,3]=mean(combmeans93)
expected[10,3]=mean(combmeans103)
expected[11,3]=mean(combmeans113)
expected[12,3]=mean(combmeans123)

expected[1,4]=mean(combmeans14) #strategy 1 nsamp=10
expected[2,4]=mean(combmeans24)
expected[3,4]=mean(combmeans34)
expected[4,4]=mean(combmeans44)
expected[5,4]=mean(combmeans54)
expected[6,4]=mean(combmeans64)
expected[7,4]=mean(combmeans74)
expected[8,4]=mean(combmeans84)
expected[9,4]=mean(combmeans94)
expected[10,4]=mean(combmeans104)
expected[11,4]=mean(combmeans114)
expected[12,4]=mean(combmeans124)

expected[1,5]=mean(combmeans15) #strategy 1 nsamp=15
expected[2,5]=mean(combmeans25)
expected[3,5]=mean(combmeans35)
expected[4,5]=mean(combmeans45)
expected[5,5]=mean(combmeans55)
expected[6,5]=mean(combmeans65)
expected[7,5]=mean(combmeans75)
expected[8,5]=mean(combmeans85)
expected[9,5]=mean(combmeans95)
expected[10,5]=mean(combmeans105)
expected[11,5]=mean(combmeans115)
expected[12,5]=mean(combmeans125)

expected[1,6]=mean(combmeans16) #strategy 1 nsamp=20
expected[2,6]=mean(combmeans26)
expected[3,6]=mean(combmeans36)
expected[4,6]=mean(combmeans46)
expected[5,6]=mean(combmeans56)
expected[6,6]=mean(combmeans66)
expected[7,6]=mean(combmeans76)
expected[8,6]=mean(combmeans86)
expected[9,6]=mean(combmeans96)
expected[10,6]=mean(combmeans106)
expected[11,6]=mean(combmeans116)
expected[12,6]=mean(combmeans126)

expectedcol=c(expected1,expected[,1],expected[,2],expected[,3],expected[,4],expected[,5],expected[,6])

means=cbind(combmeans11,combmeans21,
combmeans31,combmeans41,combmeans51,combmeans61,combmeans71,combmeans81,combmeans91,
combmeans101,combmeans111,combmeans121,
combmeans12,combmeans22, combmeans32,combmeans42,combmeans52,combmeans62,combmeans72,combmeans82,combmeans92,
combmeans102,combmeans112,combmeans122,
combmeans13,combmeans23, combmeans33,combmeans43,combmeans53,combmeans63,combmeans73,combmeans83,combmeans93,
combmeans103,combmeans113,combmeans123,
combmeans14,combmeans24, combmeans34,combmeans44,combmeans54,combmeans64,combmeans74,combmeans84,combmeans94,
combmeans104,combmeans114,combmeans124,
combmeans15,combmeans25, combmeans35,combmeans45,combmeans55,combmeans65,combmeans75,combmeans85,combmeans95,
combmeans105,combmeans115,combmeans125,
combmeans16,combmeans26, combmeans36,combmeans46,combmeans56,combmeans66,combmeans76,combmeans86,combmeans96,
combmeans106,combmeans116,combmeans126)

variance1=(1/nvar)*sum((sampmean[,1]-expected1)**2)

variance11=(1/nvar)*sum((means[,1]-expected[1,1])**2)
variance21=(1/nvar)*sum((means[,2]-expected[2,1])**2)
variance31=(1/nvar)*sum((means[,3]-expected[3,1])**2)
variance41=(1/nvar)*sum((means[,4]-expected[4,1])**2)
variance51=(1/nvar)*sum((means[,5]-expected[5,1])**2)
variance61=(1/nvar)*sum((means[,6]-expected[6,1])**2)
variance71=(1/nvar)*sum((means[,7]-expected[7,1])**2)

```

variance81=(1/nvar)*sum((means[,8]-expected[8,1])**2)
variance91=(1/nvar)*sum((means[,9]-expected[9,1])**2)
variance101=(1/nvar)*sum((means[,10]-expected[10,1])**2)
variance111=(1/nvar)*sum((means[,11]-expected[11,1])**2)
variance121=(1/nvar)*sum((means[,12]-expected[12,1])**2)

variance12=(1/nvar)*sum((means[,13]-expected[1,2])**2)
variance22=(1/nvar)*sum((means[,14]-expected[2,2])**2)
variance32=(1/nvar)*sum((means[,15]-expected[3,2])**2)
variance42=(1/nvar)*sum((means[,16]-expected[4,2])**2)
variance52=(1/nvar)*sum((means[,17]-expected[5,2])**2)
variance62=(1/nvar)*sum((means[,18]-expected[6,2])**2)
variance72=(1/nvar)*sum((means[,19]-expected[7,2])**2)
variance82=(1/nvar)*sum((means[,20]-expected[8,2])**2)
variance92=(1/nvar)*sum((means[,21]-expected[9,2])**2)
variance102=(1/nvar)*sum((means[,22]-expected[10,2])**2)
variance112=(1/nvar)*sum((means[,23]-expected[11,2])**2)
variance122=(1/nvar)*sum((means[,24]-expected[12,2])**2)

variance13=(1/nvar)*sum((means[,25]-expected[1,3])**2)
variance23=(1/nvar)*sum((means[,26]-expected[2,3])**2)
variance33=(1/nvar)*sum((means[,27]-expected[3,3])**2)
variance43=(1/nvar)*sum((means[,28]-expected[4,3])**2)
variance53=(1/nvar)*sum((means[,29]-expected[5,3])**2)
variance63=(1/nvar)*sum((means[,30]-expected[6,3])**2)
variance73=(1/nvar)*sum((means[,31]-expected[7,3])**2)
variance83=(1/nvar)*sum((means[,32]-expected[8,3])**2)
variance93=(1/nvar)*sum((means[,33]-expected[9,3])**2)
variance103=(1/nvar)*sum((means[,34]-expected[10,3])**2)
variance113=(1/nvar)*sum((means[,35]-expected[11,3])**2)
variance123=(1/nvar)*sum((means[,36]-expected[12,3])**2)

variance14=(1/nvar)*sum((means[,37]-expected[1,4])**2)
variance24=(1/nvar)*sum((means[,38]-expected[2,4])**2)
variance34=(1/nvar)*sum((means[,39]-expected[3,4])**2)
variance44=(1/nvar)*sum((means[,40]-expected[4,4])**2)
variance54=(1/nvar)*sum((means[,41]-expected[5,4])**2)
variance64=(1/nvar)*sum((means[,42]-expected[6,4])**2)
variance74=(1/nvar)*sum((means[,43]-expected[7,4])**2)
variance84=(1/nvar)*sum((means[,44]-expected[8,4])**2)
variance94=(1/nvar)*sum((means[,45]-expected[9,4])**2)
variance104=(1/nvar)*sum((means[,46]-expected[10,4])**2)
variance114=(1/nvar)*sum((means[,47]-expected[11,4])**2)
variance124=(1/nvar)*sum((means[,48]-expected[12,4])**2)

variance15=(1/nvar)*sum((means[,49]-expected[1,5])**2)
variance25=(1/nvar)*sum((means[,50]-expected[2,5])**2)
variance35=(1/nvar)*sum((means[,51]-expected[3,5])**2)
variance45=(1/nvar)*sum((means[,52]-expected[4,5])**2)
variance55=(1/nvar)*sum((means[,53]-expected[5,5])**2)
variance65=(1/nvar)*sum((means[,54]-expected[6,5])**2)
variance75=(1/nvar)*sum((means[,55]-expected[7,5])**2)
variance85=(1/nvar)*sum((means[,56]-expected[8,5])**2)
variance95=(1/nvar)*sum((means[,57]-expected[9,5])**2)
variance105=(1/nvar)*sum((means[,58]-expected[10,5])**2)
variance115=(1/nvar)*sum((means[,59]-expected[11,5])**2)
variance125=(1/nvar)*sum((means[,60]-expected[12,5])**2)

variance16=(1/nvar)*sum((means[,61]-expected[1,6])**2)
variance26=(1/nvar)*sum((means[,62]-expected[2,6])**2)
variance36=(1/nvar)*sum((means[,63]-expected[3,6])**2)
variance46=(1/nvar)*sum((means[,64]-expected[4,6])**2)
variance56=(1/nvar)*sum((means[,65]-expected[5,6])**2)
variance66=(1/nvar)*sum((means[,66]-expected[6,6])**2)
variance76=(1/nvar)*sum((means[,67]-expected[7,6])**2)
variance86=(1/nvar)*sum((means[,68]-expected[8,6])**2)
variance96=(1/nvar)*sum((means[,69]-expected[9,6])**2)
variance106=(1/nvar)*sum((means[,70]-expected[10,6])**2)
variance116=(1/nvar)*sum((means[,71]-expected[11,6])**2)
variance126=(1/nvar)*sum((means[,72]-expected[12,6])**2)

```
variance=c(variance1,variance11,variance21,variance31,variance41,variance51,
variance61,variance71,variance81,variance91,variance101,variance111,variance121,
variance12,variance22,variance32,variance42,variance52,
variance62,variance72,variance82,variance92,variance102,variance112,variance122,
variance13,variance23,variance33,variance43,variance53,
variance63,variance73,variance83,variance93,variance103,variance113,variance123,
variance14,variance24,variance34,variance44,variance54,
variance64,variance74,variance84,variance94,variance104,variance114,variance124,
variance15,variance25,variance35,variance45,variance55,
variance65,variance75,variance85,variance95,variance105,variance115,variance125,
variance16,variance26,variance36,variance46,variance56,
variance66,variance76,variance86,variance96,variance106,variance116,variance126)
```

```
bias= expectedcol-popmean
```

```
relbias=(bias/expectedcol)*100
```

```
mse= variance + bias**2
```

```
cv= (sqrt(variance)/expectedcol)*100
```

```
s.e= (sqrt(variance)/sqrt(nvar))
```

```
LB= expectedcol - 1.96*s.e
```

```
UB= expectedcol + 1.96*s.e
```

```
result=cbind(popmean,expectedcol,relbias,s.e,cv,LB,UB,mse)
```

```
write.table(result,"result_t3_scenario1_k100.txt",sep="\t")
```

```
write.table(meff_var,"meff_t3_scenario1_k100.txt",sep="\t")
```

H3 Skewed Normal Distribution

```
#####  
# Variance Estimation: Normal distribution  
# author: "Loveness Dzikiti"  
# date: "20 April 2017"  
#####  
  
#####  
#Scenario 1: N=1000000 SN(100,25) n=1000 D=nsamp=20  
#####  
#####  
  
#####  
  
setwd("C:\\Users\\User\\Dropbox\\Review\\simulations_2019\\cluster\\anormal")  
setwd("C:\\Users\\Loveness\\Dropbox\\Review\\simulations_2019\\cluster\\anormal")  
  
library(moments)  
library(sampling)  
library(survey)  
library(sn)  
  
N=1000000      #finite population size  
nsamp=20      #number of samples or replicates  
k=nvar=100  
  
n=1000  
clsiz=250  
  
set.seed(123456)  
  
mean_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample means calculated as srs  
variance_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample variance calculated as srs  
cv_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample coefficient of variation calculated as srs  
skewness_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample skewness calculated as srs  
kurtosis_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample kurtosis calculated as srs  
  
means <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample means  
variance <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample variance  
cv <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample coefficient of variation  
skewness <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample skewness  
kurtosis <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample kurtosis  
meff_var<-matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample design effect  
  
combmeans1 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 1  
combmeans2 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 2  
combmeans3 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 3  
combmeans4 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 4  
combmeans5 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 5  
combmeans6 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 6  
combmeans7 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 7  
combmeans8 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 8  
combmeans9 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 9  
combmeans10<- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 10  
combmeans11<- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 11  
combmeans12<- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 12  
  
sampmean <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store mean of 1 sample, no combining  
  
combmeans11 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1  
combmeans21 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2  
combmeans31 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3  
combmeans41 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4  
combmeans51 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5  
combmeans61 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 6  
combmeans71 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 7  
combmeans81 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1  
combmeans91 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2
```



```

combmeans116 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
combmeans126 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5

```

```
#####
```

```
### Generating a clustered population#####
```

```

xi=-1.165          #
omega=1.53         #
alpha=3           # asymmetry parameter for strong asymmetry
nclus=4000

```

```

x1 <- rsn(N, xi, omega, alpha)
x2 <- rsn(nclus, xi, omega, alpha)

```

```

sigma1=sqrt(20)
mu1=0

```

```
e_ij=(x1*sigma1)+ mu1          #individual characteristic
```

```

sigma2=sqrt(5)
mu2=0

```

```
cl = (x2*sigma2) + mu2          #cluster characteristic
```

```
mu <- matrix(c(rep(100,N)), nrow = N, ncol = 1) #store superpopulation mean
```

```

v_j <-
c(rep(cl[1],clsize),rep(cl[2],clsize),rep(cl[3],clsize),rep(cl[4],clsize),rep(cl[5],clsize),rep(cl[6],clsize),rep(cl[7],clsize),rep(cl[8],clsize),rep(cl[9],cl
size),rep(cl[10],clsize),
rep(cl[11],clsize),rep(cl[12],clsize),rep(cl[13],clsize),rep(cl[14],clsize),rep(cl[15],clsize),rep(cl[16],clsize),rep(cl[17],clsize),rep(cl[18],clsize),re
p(cl[19],clsize),rep(cl[20],clsize),
rep(cl[21],clsize),rep(cl[22],clsize),rep(cl[23],clsize),rep(cl[24],clsize),rep(cl[25],clsize),rep(cl[26],clsize),rep(cl[27],clsize),rep(cl[28],clsize),re
p(cl[29],clsize),rep(cl[30],clsize),
rep(cl[31],clsize),rep(cl[32],clsize),rep(cl[33],clsize),rep(cl[34],clsize),rep(cl[35],clsize),rep(cl[36],clsize),rep(cl[37],clsize),rep(cl[38],clsize),re
p(cl[39],clsize),rep(cl[40],clsize),
rep(cl[41],clsize),rep(cl[42],clsize),rep(cl[43],clsize),rep(cl[44],clsize),rep(cl[45],clsize),rep(cl[46],clsize),rep(cl[47],clsize),rep(cl[48],clsize),re
p(cl[49],clsize),rep(cl[50],clsize),
rep(cl[51],clsize),rep(cl[52],clsize),rep(cl[53],clsize),rep(cl[54],clsize),rep(cl[55],clsize),rep(cl[56],clsize),rep(cl[57],clsize),rep(cl[58],clsize),re
p(cl[59],clsize),rep(cl[60],clsize),
rep(cl[61],clsize),rep(cl[62],clsize),rep(cl[63],clsize),rep(cl[64],clsize),rep(cl[65],clsize),rep(cl[66],clsize),rep(cl[67],clsize),rep(cl[68],clsize),re
p(cl[69],clsize),rep(cl[70],clsize),
rep(cl[71],clsize),rep(cl[72],clsize),rep(cl[73],clsize),rep(cl[74],clsize),rep(cl[75],clsize),rep(cl[76],clsize),rep(cl[77],clsize),rep(cl[78],clsize),re
p(cl[79],clsize),rep(cl[80],clsize),
rep(cl[81],clsize),rep(cl[82],clsize),rep(cl[83],clsize),rep(cl[84],clsize),rep(cl[85],clsize),rep(cl[86],clsize),rep(cl[87],clsize),rep(cl[88],clsize),re
p(cl[89],clsize),rep(cl[90],clsize),
rep(cl[91],clsize),rep(cl[92],clsize),rep(cl[93],clsize),rep(cl[94],clsize),rep(cl[95],clsize),rep(cl[96],clsize),rep(cl[97],clsize),rep(cl[98],clsize),re
p(cl[99],clsize),rep(cl[100],clsize),
rep(cl[101],clsize),rep(cl[102],clsize),rep(cl[103],clsize),rep(cl[104],clsize),rep(cl[105],clsize),rep(cl[106],clsize),rep(cl[107],clsize),rep(cl[108
],clsize),rep(cl[109],clsize),rep(cl[110],clsize),
rep(cl[111],clsize),rep(cl[112],clsize),rep(cl[113],clsize),rep(cl[114],clsize),rep(cl[115],clsize),rep(cl[116],clsize),rep(cl[117],clsize),rep(cl[118
],clsize),rep(cl[119],clsize),rep(cl[120],clsize),
rep(cl[121],clsize),rep(cl[122],clsize),rep(cl[123],clsize),rep(cl[124],clsize),rep(cl[125],clsize),rep(cl[126],clsize),rep(cl[127],clsize),rep(cl[128
],clsize),rep(cl[129],clsize),rep(cl[130],clsize),
rep(cl[131],clsize),rep(cl[132],clsize),rep(cl[133],clsize),rep(cl[134],clsize),rep(cl[135],clsize),rep(cl[136],clsize),rep(cl[137],clsize),rep(cl[138
],clsize),rep(cl[139],clsize),rep(cl[140],clsize),
rep(cl[141],clsize),rep(cl[142],clsize),rep(cl[143],clsize),rep(cl[144],clsize),rep(cl[145],clsize),rep(cl[146],clsize),rep(cl[147],clsize),rep(cl[148
],clsize),rep(cl[149],clsize),rep(cl[150],clsize),
rep(cl[151],clsize),rep(cl[152],clsize),rep(cl[153],clsize),rep(cl[154],clsize),rep(cl[155],clsize),rep(cl[156],clsize),rep(cl[157],clsize),rep(cl[158
],clsize),rep(cl[159],clsize),rep(cl[160],clsize),
rep(cl[161],clsize),rep(cl[162],clsize),rep(cl[163],clsize),rep(cl[164],clsize),rep(cl[165],clsize),rep(cl[166],clsize),rep(cl[167],clsize),rep(cl[168
],clsize),rep(cl[169],clsize),rep(cl[170],clsize),
rep(cl[171],clsize),rep(cl[172],clsize),rep(cl[173],clsize),rep(cl[174],clsize),rep(cl[175],clsize),rep(cl[176],clsize),rep(cl[177],clsize),rep(cl[178
],clsize),rep(cl[179],clsize),rep(cl[180],clsize),
rep(cl[181],clsize),rep(cl[182],clsize),rep(cl[183],clsize),rep(cl[184],clsize),rep(cl[185],clsize),rep(cl[186],clsize),rep(cl[187],clsize),rep(cl[188
],clsize),rep(cl[189],clsize),rep(cl[190],clsize),
rep(cl[191],clsize),rep(cl[192],clsize),rep(cl[193],clsize),rep(cl[194],clsize),rep(cl[195],clsize),rep(cl[196],clsize),rep(cl[197],clsize),rep(cl[198
],clsize),rep(cl[199],clsize),rep(cl[200],clsize),
rep(cl[201],clsize),rep(cl[202],clsize),rep(cl[203],clsize),rep(cl[204],clsize),rep(cl[205],clsize),rep(cl[206],clsize),rep(cl[207],clsize),rep(cl[208
],clsize),rep(cl[209],clsize),rep(cl[210],clsize),

```



```

rep('3921',clsize),rep('3922',clsize),rep('3923',clsize),rep('3924',clsize),rep('3925',clsize),rep('3926',clsize),rep('3927',clsize),rep('3928',clsize
),rep('3929',clsize),rep('3930',clsize),
rep('3931',clsize),rep('3932',clsize),rep('3933',clsize),rep('3934',clsize),rep('3935',clsize),rep('3936',clsize),rep('3937',clsize),rep('3938',clsize
),rep('3939',clsize),rep('3940',clsize),
rep('3941',clsize),rep('3942',clsize),rep('3943',clsize),rep('3944',clsize),rep('3945',clsize),rep('3946',clsize),rep('3947',clsize),rep('3948',clsize
),rep('3949',clsize),rep('3950',clsize),
rep('3951',clsize),rep('3952',clsize),rep('3953',clsize),rep('3954',clsize),rep('3955',clsize),rep('3956',clsize),rep('3957',clsize),rep('3958',clsize
),rep('3959',clsize),rep('3960',clsize),
rep('3961',clsize),rep('3962',clsize),rep('3963',clsize),rep('3964',clsize),rep('3965',clsize),rep('3966',clsize),rep('3967',clsize),rep('3968',clsize
),rep('3969',clsize),rep('3970',clsize),
rep('3971',clsize),rep('3972',clsize),rep('3973',clsize),rep('3974',clsize),rep('3975',clsize),rep('3976',clsize),rep('3977',clsize),rep('3978',clsize
),rep('3979',clsize),rep('3980',clsize),
rep('3981',clsize),rep('3982',clsize),rep('3983',clsize),rep('3984',clsize),rep('3985',clsize),rep('3986',clsize),rep('3987',clsize),rep('3988',clsize
),rep('3989',clsize),rep('3990',clsize),
rep('3991',clsize),rep('3992',clsize),rep('3993',clsize),rep('3994',clsize),rep('3995',clsize),rep('3996',clsize),rep('3997',clsize),rep('3998',clsize
),rep('3999',clsize),rep('4000',clsize)

```

```
)
```

```
popx1<-data.frame(cl_id,x)
popx1$unit<-1:nrow(popx1)
```

```
popx <- data.frame(popx1$unit,cl_id,mu,v_j,e_ij,x)
colnames(popx)<-c('unit_id','cl_id','mu','v_j','e_ij','x')
```

```
pop.des=svydesign(data=popx,ids=~cl_id+unit_id, nest=TRUE)
pop.mean=data.frame(svymean(x,pop.des))
popmean <- pop.mean[1,1]#finite population mean
```

```
##### Weights storage #####
```

```
for (k in 1:nvar){
```

```
for (j in 1:nsamp){
```

```

w1<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1) #store weights
w2<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w3<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w4<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w5<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w6<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w7<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w8<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1) #store weights
w9<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w10<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w11<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
w12<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
}

```

```
#####
#####
# Two stage Cluster Random Sampling from a clustered finite population
#####
```

```
n2 = rep(20,50) # select 20 units from 50 clusters
```

```
for (i in 1:nsamp) {
```

```

sample = mstage(popx, stage=list("cluster","cluster"), varnames=list("cl_id", "unit_id"),size=list(50,n2), method=list("srswor","srswor"))
#unique(sample[[1]]$cl_id)
#unique(sample[[2]]$unit_id) #double checking

```

```
s=getdata(popx,sample)[[2]]
```

```
#####
# Naive weights
#####
```

```
mean_naive[i]=mean(s$X)
```

```

variance_naive[i]=var(s$x)
cv_naive[i]=abs(sqrt(variance_naive[i])/mean_naive[i])
skewness_naive[i]<-skewness(s$x)
kurtosis_naive[i]<-kurtosis(s$x)

#####
# Design conscious weights
#####

plan=svydesign(data=s,ids=~cl_id+unit_id, nest=TRUE)
mean_st=svymean(~x,plan)
mean_st=as.data.frame(mean_st)

means[i]<-mean_st[1,1]          #sample mean for ith sample
variance[i]<-mean_st[1,2]^2    #sample variance, design consistent variance
cv[i]<-abs(sqrt(variance[i])/means[i])    #coefficient of variation

# skewness[i]<-skewness(s[,i])
moments = svymean(~l(x^4) + l(x^3) +
                 l(x^2) + l(x), plan)

skew=data.frame(svycontrast(moments,
                             quote(`l(x^3)` - 3*`l(x^2)`*`l(x)` +
                                     3*`l(x)`*`l(x)^2` - `l(x)^3`)/
                             (`l(x^2)` - `l(x)^2`^1.5)))
skewness[i]<-skew[1,1]

#kurtosis[i]<-kurtosis(s[,i])

kurt=data.frame(svycontrast(moments,
                             quote(`l(x^4)` -
                                     4*`l(x^3)`*`l(x)` +
                                     6*`l(x^2)`*`l(x)^2` -
                                     4*`l(x)`*`l(x)^3` +
                                     `l(x)^4`/
                                     (`l(x^2)` -
                                     `l(x)^2`^2)))
kurtosis[i]<-kurt[1,1]

meff.1 <- data.frame(svymean(s$x,plan,deff=TRUE))
meff_var[i]=variance[i]/((1-(n/N))*(var(s$x)/n))

#####Weights calculation#####

w1[i]=n          #inverse of the sample size

w2[i]<-1/(variance[i]) #inverse of the design variance of the estimator of the mean
w3[i]<-1/(variance_naive[i]) #inverse of the naive variance of the estimator of the mean

w4[i]=1          #simple average weights

w5[i]=1/cv[i]    #inverse of cv
w6[i]=1/cv_naive[i]    #inverse of naive cv

w7[i]=abs(1/skewness[i]) #inverse skewness
w8[i]=abs(1/skewness_naive[i]) #inverse skewness

w9[i]=1/kurtosis[i] #inverse kurtosis
w10[i]=1/kurtosis_naive[i] #inverse kurtosis

w11[i]=meff_var[i]/variance[i] #in srs deff=1 so coincides with inverse of variance
w12[i]=1/meff_var[i]

combmeans1[i]<-round(sum(means[1:i]*w1[1:i])/sum(w1[1:i]),digits=9) #combine means using inverse sample size
combmeans2[i]<-round(sum(means[1:i]*w2[1:i])/sum(w2[1:i]),digits=9) # combine using inverse variance
combmeans3[i]<-round(sum(means[1:i]*w3[1:i])/sum(w3[1:i]),digits=9) # combine using inverse variance naive
combmeans4[i]<-round(sum(means[1:i]*w4[1:i])/sum(w4[1:i]),digits=9) # combine using a weight of 1

```

```

combmeans5[i]<-round(sum(means[1:i]*w5[1:i])/sum(w5[1:i]),digits=9) # combine using a weight of inverse CV
combmeans6[i]<-round(sum(means[1:i]*w6[1:i])/sum(w6[1:i]),digits=9) # combine using a weight of inverse CV naive
combmeans7[i]<-round(sum(means[1:i]*w7[1:i])/sum(w7[1:i]),digits=9) # combine using a weight of inverse skewness
combmeans8[i]<-round(sum(means[1:i]*w8[1:i])/sum(w8[1:i]),digits=9) #combine using a weight of inverse skewness naive
combmeans9[i]<-round(sum(means[1:i]*w9[1:i])/sum(w9[1:i]),digits=9) # combine using inverse kurtosis
combmeans10[i]<-round(sum(means[1:i]*w10[1:i])/sum(w10[1:i]),digits=9) # combine using inverse kurtosis naive
combmeans11[i]<-round(sum(means[1:i]*w11[1:i])/sum(w11[1:i]),digits=9) # combine using a weight of meff/var
combmeans12[i]<-round(sum(means[1:i]*w12[1:i])/sum(w12[1:i]),digits=9) # combine using a weight of 1/meff

```

```

}

```

```

sampmean[k]=means[1] # capturing the combined mean, where nsamp reps are combined, all weighting strategies

```

```

combmeans11[k]=combmeans1[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies

```

```

combmeans21[k]=combmeans2[2]

```

```

combmeans31[k]=combmeans3[2]

```

```

combmeans41[k]=combmeans4[2]

```

```

combmeans51[k]=combmeans5[2]

```

```

combmeans61[k]=combmeans6[2]

```

```

combmeans71[k]=combmeans7[2]

```

```

combmeans81[k]=combmeans8[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies

```

```

combmeans91[k]=combmeans9[2]

```

```

combmeans101[k]=combmeans10[2]

```

```

combmeans111[k]=combmeans11[2]

```

```

combmeans121[k]=combmeans12[2]

```

```

combmeans12[k]=combmeans1[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies

```

```

combmeans22[k]=combmeans2[3]

```

```

combmeans32[k]=combmeans3[3]

```

```

combmeans42[k]=combmeans4[3]

```

```

combmeans52[k]=combmeans5[3]

```

```

combmeans62[k]=combmeans6[3]

```

```

combmeans72[k]=combmeans7[3]

```

```

combmeans82[k]=combmeans8[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies

```

```

combmeans92[k]=combmeans9[3]

```

```

combmeans102[k]=combmeans10[3]

```

```

combmeans112[k]=combmeans11[3]

```

```

combmeans122[k]=combmeans12[3]

```

```

combmeans13[k]=combmeans1[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies

```

```

combmeans23[k]=combmeans2[5]

```

```

combmeans33[k]=combmeans3[5]

```

```

combmeans43[k]=combmeans4[5]

```

```

combmeans53[k]=combmeans5[5]

```

```

combmeans63[k]=combmeans6[5]

```

```

combmeans73[k]=combmeans7[5]

```

```

combmeans83[k]=combmeans8[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies

```

```

combmeans93[k]=combmeans9[5]

```

```

combmeans103[k]=combmeans10[5]

```

```

combmeans113[k]=combmeans11[5]

```

```

combmeans123[k]=combmeans12[5]

```

```

combmeans14[k]=combmeans1[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies

```

```

combmeans24[k]=combmeans2[10]

```

```

combmeans34[k]=combmeans3[10]

```

```

combmeans44[k]=combmeans4[10]

```

```

combmeans54[k]=combmeans5[10]

```

```

combmeans64[k]=combmeans6[10]

```

```

combmeans74[k]=combmeans7[10]

```

```

combmeans84[k]=combmeans8[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies

```

```

combmeans94[k]=combmeans9[10]

```

```

combmeans104[k]=combmeans10[10]

```

```

combmeans114[k]=combmeans11[10]

```

```

combmeans124[k]=combmeans12[10]

```

```

combmeans15[k]=combmeans1[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans25[k]=combmeans2[15]
combmeans35[k]=combmeans3[15]
combmeans45[k]=combmeans4[15]
combmeans55[k]=combmeans5[15]
combmeans65[k]=combmeans6[15]
combmeans75[k]=combmeans7[15]
combmeans85[k]=combmeans8[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans95[k]=combmeans9[15]
combmeans105[k]=combmeans10[15]
combmeans115[k]=combmeans11[15]
combmeans125[k]=combmeans12[15]

combmeans16[k]=combmeans1[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans26[k]=combmeans2[20]
combmeans36[k]=combmeans3[20]
combmeans46[k]=combmeans4[20]
combmeans56[k]=combmeans5[20]
combmeans66[k]=combmeans6[20]
combmeans76[k]=combmeans7[20]
combmeans86[k]=combmeans8[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans96[k]=combmeans9[20]
combmeans106[k]=combmeans10[20]
combmeans116[k]=combmeans11[20]
combmeans126[k]=combmeans12[20]

}
expected1=mean(sampmean) #expected mean for one sample, not combining

expected=matrix(0,12,6)

expected[1,1]=mean(combmeans11) #strategy 1 nsamp=2
expected[2,1]=mean(combmeans21)
expected[3,1]=mean(combmeans31)
expected[4,1]=mean(combmeans41)
expected[5,1]=mean(combmeans51)
expected[6,1]=mean(combmeans61)
expected[7,1]=mean(combmeans71)
expected[8,1]=mean(combmeans81)
expected[9,1]=mean(combmeans91)
expected[10,1]=mean(combmeans101)
expected[11,1]=mean(combmeans111)
expected[12,1]=mean(combmeans121)

expected[1,2]=mean(combmeans12) #strategy 1 nsamp=3
expected[2,2]=mean(combmeans22)
expected[3,2]=mean(combmeans32)
expected[4,2]=mean(combmeans42)
expected[5,2]=mean(combmeans52)
expected[6,2]=mean(combmeans62)
expected[7,2]=mean(combmeans72)
expected[8,2]=mean(combmeans82)
expected[9,2]=mean(combmeans92)
expected[10,2]=mean(combmeans102)
expected[11,2]=mean(combmeans112)
expected[12,2]=mean(combmeans122)

expected[1,3]=mean(combmeans13) #strategy 1 nsamp=5
expected[2,3]=mean(combmeans23)
expected[3,3]=mean(combmeans33)
expected[4,3]=mean(combmeans43)
expected[5,3]=mean(combmeans53)
expected[6,3]=mean(combmeans63)
expected[7,3]=mean(combmeans73)
expected[8,3]=mean(combmeans83)
expected[9,3]=mean(combmeans93)
expected[10,3]=mean(combmeans103)
expected[11,3]=mean(combmeans113)

```

```

expected[12,3]=mean(combmeans123)

expected[1,4]=mean(combmeans14) #strategy 1 nsamp=10
expected[2,4]=mean(combmeans24)
expected[3,4]=mean(combmeans34)
expected[4,4]=mean(combmeans44)
expected[5,4]=mean(combmeans54)
expected[6,4]=mean(combmeans64)
expected[7,4]=mean(combmeans74)
expected[8,4]=mean(combmeans84)
expected[9,4]=mean(combmeans94)
expected[10,4]=mean(combmeans104)
expected[11,4]=mean(combmeans114)
expected[12,4]=mean(combmeans124)

expected[1,5]=mean(combmeans15) #strategy 1 nsamp=15
expected[2,5]=mean(combmeans25)
expected[3,5]=mean(combmeans35)
expected[4,5]=mean(combmeans45)
expected[5,5]=mean(combmeans55)
expected[6,5]=mean(combmeans65)
expected[7,5]=mean(combmeans75)
expected[8,5]=mean(combmeans85)
expected[9,5]=mean(combmeans95)
expected[10,5]=mean(combmeans105)
expected[11,5]=mean(combmeans115)
expected[12,5]=mean(combmeans125)

expected[1,6]=mean(combmeans16) #strategy 1 nsamp=20
expected[2,6]=mean(combmeans26)
expected[3,6]=mean(combmeans36)
expected[4,6]=mean(combmeans46)
expected[5,6]=mean(combmeans56)
expected[6,6]=mean(combmeans66)
expected[7,6]=mean(combmeans76)
expected[8,6]=mean(combmeans86)
expected[9,6]=mean(combmeans96)
expected[10,6]=mean(combmeans106)
expected[11,6]=mean(combmeans116)
expected[12,6]=mean(combmeans126)

expectedcol=c(expected1,expected[,1],expected[,2],expected[,3],expected[,4],expected[,5],expected[,6])

means=cbind(combmeans11,combmeans21,
combmeans31,combmeans41,combmeans51,combmeans61,combmeans71,combmeans81,combmeans91,
combmeans101,combmeans111,combmeans121,
combmeans12,combmeans22, combmeans32,combmeans42,combmeans52,combmeans62,combmeans72,combmeans82,combmeans92,
combmeans102,combmeans112,combmeans122,
combmeans13,combmeans23, combmeans33,combmeans43,combmeans53,combmeans63,combmeans73,combmeans83,combmeans93,
combmeans103,combmeans113,combmeans123,
combmeans14,combmeans24, combmeans34,combmeans44,combmeans54,combmeans64,combmeans74,combmeans84,combmeans94,
combmeans104,combmeans114,combmeans124,
combmeans15,combmeans25, combmeans35,combmeans45,combmeans55,combmeans65,combmeans75,combmeans85,combmeans95,
combmeans105,combmeans115,combmeans125,
combmeans16,combmeans26, combmeans36,combmeans46,combmeans56,combmeans66,combmeans76,combmeans86,combmeans96,
combmeans106,combmeans116,combmeans126)

variance1=(1/nvar)*sum((sampmean[,1]-expected1)**2)

variance11=(1/nvar)*sum((means[,1]-expected[1,1])**2)
variance21=(1/nvar)*sum((means[,2]-expected[2,1])**2)
variance31=(1/nvar)*sum((means[,3]-expected[3,1])**2)
variance41=(1/nvar)*sum((means[,4]-expected[4,1])**2)
variance51=(1/nvar)*sum((means[,5]-expected[5,1])**2)
variance61=(1/nvar)*sum((means[,6]-expected[6,1])**2)
variance71=(1/nvar)*sum((means[,7]-expected[7,1])**2)
variance81=(1/nvar)*sum((means[,8]-expected[8,1])**2)
variance91=(1/nvar)*sum((means[,9]-expected[9,1])**2)
variance101=(1/nvar)*sum((means[,10]-expected[10,1])**2)
variance111=(1/nvar)*sum((means[,11]-expected[11,1])**2)

```



```
variance12,variance22,variance32,variance42,variance52,  
variance62,variance72,variance82,variance92,variance102,variance112,variance122,  
variance13,variance23,variance33,variance43,variance53,  
variance63,variance73,variance83,variance93,variance103,variance113,variance123,  
variance14,variance24,variance34,variance44,variance54,  
variance64,variance74,variance84,variance94,variance104,variance114,variance124,  
variance15,variance25,variance35,variance45,variance55,  
variance65,variance75,variance85,variance95,variance105,variance115,variance125,  
variance16,variance26,variance36,variance46,variance56,  
variance66,variance76,variance86,variance96,variance106,variance116,variance126)
```

```
bias= expectedcol-popmean
```

```
relbias=(bias/expectedcol)*100
```

```
mse= variance + bias**2
```

```
cv= (sqrt(variance)/expectedcol)*100
```

```
s.e= (sqrt(variance)/sqrt(nvar))
```

```
LB= expectedcol - 1.96*s.e
```

```
UB= expectedcol + 1.96*s.e
```

```
result=cbind(popmean,expectedcol,relbias,s.e,cv,LB,UB,mse)
```

```
write.table(result,"result_anormal_scenario1_100.txt",sep="\t")
```

```
write.table(meff_var,"meff_anormal_scenario1_100.txt",sep="\t")
```

H4 Skewed T Distribution, 3 degrees of freedom

```
#####  
# Variance Estimation: Normal distribution  
# author: "Loveness Dzikiti"  
# date: "20 April 2017"  
#####  
  
# Scenario 1: N=1 000 000 at3(100,25) n=1000 D=20  
#####  
# Generating a clustered population  
# X_ij= mu + v_j + e_ij  
  
setwd("C:\\Users\\User\\Dropbox\\Review\\simulations_2019\\cluster\\at3")  
  
library(moments)  
library(sampling)  
library(survey)  
library(sn)  
  
N=1000000      #finite population size  
n=1000        #sample size  
nsamp=20      #number of samples  
clsiz=250  
k=nvar=100  
set.seed(123456)  
  
#####storage matrices#####  
  
mean_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample means calculated as srs  
variance_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample variance calculated as srs  
cv_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample coefficient of variation calculated as srs  
skewness_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample skewness calculated as srs  
kurtosis_naive <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample kurtosis calculated as srs  
  
means <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample means  
variance <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample variance  
designvar <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store design consistent variance  
cv <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample coefficient of variation  
skewness <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample skewness  
kurtosis <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample kurtosis  
meff_var<-matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store sample design effect  
  
combmeans1 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 1  
combmeans2 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 2  
combmeans3 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 3  
combmeans4 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 4  
combmeans5 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 5  
combmeans6 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 6  
combmeans7 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 7  
combmeans8 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 8  
combmeans9 <- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 9  
combmeans10<- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 10  
combmeans11<- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 11  
combmeans12<- matrix(c(rep.int(NA,nsamp)),nrow=nsamp,ncol=1) #store means combined weighting strategy 12  
  
samppmean <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store mean of 1 sample, no combining  
  
combmeans11 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1  
combmeans21 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2  
combmeans31 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3  
combmeans41 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4  
combmeans51 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 5  
combmeans61 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 6  
combmeans71 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 7  
combmeans81 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 1  
combmeans91 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 2  
combmeans101 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 3  
combmeans111 <- matrix(c(rep.int(NA,k)),nrow=k,ncol=1) #store means combined weighting strategy 4
```


#####

Generating a clustered population#####

```
v=3
nclus=4000

xi=-0.7891
omega=0.746
alpha=3

mu1=0
sigma1=sqrt(20)

mu2=0
sigma2=sqrt(5)
```

```
x1<- rst(N, xi, omega, alpha, v)      #simulate data from a t distribution random mean and variance
e_ij=(x1*sigma1)+mu1
```

#####

```
x2<- rst(nclus, xi, omega, alpha, v)
cl=(x2*sigma2)+mu2
```

```
mu <- matrix(c(rep(100,N)), nrow = N, ncol = 1) #store superpopulation mean
```

```
v_j <-
c(rep(cl[1],clsize),rep(cl[2],clsize),rep(cl[3],clsize),rep(cl[4],clsize),rep(cl[5],clsize),rep(cl[6],clsize),rep(cl[7],clsize),rep(cl[8],clsize),rep(cl[9],cl
size),rep(cl[10],clsize),
rep(cl[11],clsize),rep(cl[12],clsize),rep(cl[13],clsize),rep(cl[14],clsize),rep(cl[15],clsize),rep(cl[16],clsize),rep(cl[17],clsize),rep(cl[18],clsize),re
p(cl[19],clsize),rep(cl[20],clsize),
rep(cl[21],clsize),rep(cl[22],clsize),rep(cl[23],clsize),rep(cl[24],clsize),rep(cl[25],clsize),rep(cl[26],clsize),rep(cl[27],clsize),rep(cl[28],clsize),re
p(cl[29],clsize),rep(cl[30],clsize),
rep(cl[31],clsize),rep(cl[32],clsize),rep(cl[33],clsize),rep(cl[34],clsize),rep(cl[35],clsize),rep(cl[36],clsize),rep(cl[37],clsize),rep(cl[38],clsize),re
p(cl[39],clsize),rep(cl[40],clsize),
rep(cl[41],clsize),rep(cl[42],clsize),rep(cl[43],clsize),rep(cl[44],clsize),rep(cl[45],clsize),rep(cl[46],clsize),rep(cl[47],clsize),rep(cl[48],clsize),re
p(cl[49],clsize),rep(cl[50],clsize),
rep(cl[51],clsize),rep(cl[52],clsize),rep(cl[53],clsize),rep(cl[54],clsize),rep(cl[55],clsize),rep(cl[56],clsize),rep(cl[57],clsize),rep(cl[58],clsize),re
p(cl[59],clsize),rep(cl[60],clsize),
rep(cl[61],clsize),rep(cl[62],clsize),rep(cl[63],clsize),rep(cl[64],clsize),rep(cl[65],clsize),rep(cl[66],clsize),rep(cl[67],clsize),rep(cl[68],clsize),re
p(cl[69],clsize),rep(cl[70],clsize),
rep(cl[71],clsize),rep(cl[72],clsize),rep(cl[73],clsize),rep(cl[74],clsize),rep(cl[75],clsize),rep(cl[76],clsize),rep(cl[77],clsize),rep(cl[78],clsize),re
p(cl[79],clsize),rep(cl[80],clsize),
rep(cl[81],clsize),rep(cl[82],clsize),rep(cl[83],clsize),rep(cl[84],clsize),rep(cl[85],clsize),rep(cl[86],clsize),rep(cl[87],clsize),rep(cl[88],clsize),re
p(cl[89],clsize),rep(cl[90],clsize),
rep(cl[91],clsize),rep(cl[92],clsize),rep(cl[93],clsize),rep(cl[94],clsize),rep(cl[95],clsize),rep(cl[96],clsize),rep(cl[97],clsize),rep(cl[98],clsize),re
p(cl[99],clsize),rep(cl[100],clsize),
rep(cl[101],clsize),rep(cl[102],clsize),rep(cl[103],clsize),rep(cl[104],clsize),rep(cl[105],clsize),rep(cl[106],clsize),rep(cl[107],clsize),rep(cl[108
],clsize),rep(cl[109],clsize),rep(cl[110],clsize),
rep(cl[111],clsize),rep(cl[112],clsize),rep(cl[113],clsize),rep(cl[114],clsize),rep(cl[115],clsize),rep(cl[116],clsize),rep(cl[117],clsize),rep(cl[118
],clsize),rep(cl[119],clsize),rep(cl[120],clsize),
rep(cl[121],clsize),rep(cl[122],clsize),rep(cl[123],clsize),rep(cl[124],clsize),rep(cl[125],clsize),rep(cl[126],clsize),rep(cl[127],clsize),rep(cl[128
],clsize),rep(cl[129],clsize),rep(cl[130],clsize),
rep(cl[131],clsize),rep(cl[132],clsize),rep(cl[133],clsize),rep(cl[134],clsize),rep(cl[135],clsize),rep(cl[136],clsize),rep(cl[137],clsize),rep(cl[138
],clsize),rep(cl[139],clsize),rep(cl[140],clsize),
rep(cl[141],clsize),rep(cl[142],clsize),rep(cl[143],clsize),rep(cl[144],clsize),rep(cl[145],clsize),rep(cl[146],clsize),rep(cl[147],clsize),rep(cl[148
],clsize),rep(cl[149],clsize),rep(cl[150],clsize),
rep(cl[151],clsize),rep(cl[152],clsize),rep(cl[153],clsize),rep(cl[154],clsize),rep(cl[155],clsize),rep(cl[156],clsize),rep(cl[157],clsize),rep(cl[158
],clsize),rep(cl[159],clsize),rep(cl[160],clsize),
rep(cl[161],clsize),rep(cl[162],clsize),rep(cl[163],clsize),rep(cl[164],clsize),rep(cl[165],clsize),rep(cl[166],clsize),rep(cl[167],clsize),rep(cl[168
],clsize),rep(cl[169],clsize),rep(cl[170],clsize),
rep(cl[171],clsize),rep(cl[172],clsize),rep(cl[173],clsize),rep(cl[174],clsize),rep(cl[175],clsize),rep(cl[176],clsize),rep(cl[177],clsize),rep(cl[178
],clsize),rep(cl[179],clsize),rep(cl[180],clsize),
rep(cl[181],clsize),rep(cl[182],clsize),rep(cl[183],clsize),rep(cl[184],clsize),rep(cl[185],clsize),rep(cl[186],clsize),rep(cl[187],clsize),rep(cl[188
],clsize),rep(cl[189],clsize),rep(cl[190],clsize),
rep(cl[191],clsize),rep(cl[192],clsize),rep(cl[193],clsize),rep(cl[194],clsize),rep(cl[195],clsize),rep(cl[196],clsize),rep(cl[197],clsize),rep(cl[198
],clsize),rep(cl[199],clsize),rep(cl[200],clsize),
```



```

rep('121',clsize),rep('122',clsize),rep('123',clsize),rep('124',clsize),rep('125',clsize),rep('126',clsize),rep('127',clsize),rep('128',clsize),rep('129',clsize),rep('130',clsize),
rep('131',clsize),rep('132',clsize),rep('133',clsize),rep('134',clsize),rep('135',clsize),rep('136',clsize),rep('137',clsize),rep('138',clsize),rep('139',clsize),rep('140',clsize),
rep('141',clsize),rep('142',clsize),rep('143',clsize),rep('144',clsize),rep('145',clsize),rep('146',clsize),rep('147',clsize),rep('148',clsize),rep('149',clsize),rep('150',clsize),
rep('151',clsize),rep('152',clsize),rep('153',clsize),rep('154',clsize),rep('155',clsize),rep('156',clsize),rep('157',clsize),rep('158',clsize),rep('159',clsize),rep('160',clsize),
rep('161',clsize),rep('162',clsize),rep('163',clsize),rep('164',clsize),rep('165',clsize),rep('166',clsize),rep('167',clsize),rep('168',clsize),rep('169',clsize),rep('170',clsize),
rep('171',clsize),rep('172',clsize),rep('173',clsize),rep('174',clsize),rep('175',clsize),rep('176',clsize),rep('177',clsize),rep('178',clsize),rep('179',clsize),rep('180',clsize),
rep('181',clsize),rep('182',clsize),rep('183',clsize),rep('184',clsize),rep('185',clsize),rep('186',clsize),rep('187',clsize),rep('188',clsize),rep('189',clsize),rep('190',clsize),
rep('191',clsize),rep('192',clsize),rep('193',clsize),rep('194',clsize),rep('195',clsize),rep('196',clsize),rep('197',clsize),rep('198',clsize),rep('199',clsize),rep('200',clsize),
rep('201',clsize),rep('202',clsize),rep('203',clsize),rep('204',clsize),rep('205',clsize),rep('206',clsize),rep('207',clsize),rep('208',clsize),rep('209',clsize),rep('210',clsize),
rep('211',clsize),rep('212',clsize),rep('213',clsize),rep('214',clsize),rep('215',clsize),rep('216',clsize),rep('217',clsize),rep('218',clsize),rep('219',clsize),rep('220',clsize),
rep('221',clsize),rep('222',clsize),rep('223',clsize),rep('224',clsize),rep('225',clsize),rep('226',clsize),rep('227',clsize),rep('228',clsize),rep('229',clsize),rep('230',clsize),
rep('231',clsize),rep('232',clsize),rep('233',clsize),rep('234',clsize),rep('235',clsize),rep('236',clsize),rep('237',clsize),rep('238',clsize),rep('239',clsize),rep('240',clsize),
rep('241',clsize),rep('242',clsize),rep('243',clsize),rep('244',clsize),rep('245',clsize),rep('246',clsize),rep('247',clsize),rep('248',clsize),rep('249',clsize),rep('250',clsize),
rep('251',clsize),rep('252',clsize),rep('253',clsize),rep('254',clsize),rep('255',clsize),rep('256',clsize),rep('257',clsize),rep('258',clsize),rep('259',clsize),rep('260',clsize),
rep('261',clsize),rep('262',clsize),rep('263',clsize),rep('264',clsize),rep('265',clsize),rep('266',clsize),rep('267',clsize),rep('268',clsize),rep('269',clsize),rep('270',clsize),
rep('271',clsize),rep('272',clsize),rep('273',clsize),rep('274',clsize),rep('275',clsize),rep('276',clsize),rep('277',clsize),rep('278',clsize),rep('279',clsize),rep('280',clsize),
rep('281',clsize),rep('282',clsize),rep('283',clsize),rep('284',clsize),rep('285',clsize),rep('286',clsize),rep('287',clsize),rep('288',clsize),rep('289',clsize),rep('290',clsize),
rep('291',clsize),rep('292',clsize),rep('293',clsize),rep('294',clsize),rep('295',clsize),rep('296',clsize),rep('297',clsize),rep('298',clsize),rep('299',clsize),rep('300',clsize),
rep('301',clsize),rep('302',clsize),rep('303',clsize),rep('304',clsize),rep('305',clsize),rep('306',clsize),rep('307',clsize),rep('308',clsize),rep('309',clsize),rep('310',clsize),
rep('311',clsize),rep('312',clsize),rep('313',clsize),rep('314',clsize),rep('315',clsize),rep('316',clsize),rep('317',clsize),rep('318',clsize),rep('319',clsize),rep('320',clsize),
rep('321',clsize),rep('322',clsize),rep('323',clsize),rep('324',clsize),rep('325',clsize),rep('326',clsize),rep('327',clsize),rep('328',clsize),rep('329',clsize),rep('330',clsize),
rep('331',clsize),rep('332',clsize),rep('333',clsize),rep('334',clsize),rep('335',clsize),rep('336',clsize),rep('337',clsize),rep('338',clsize),rep('339',clsize),rep('340',clsize),
rep('341',clsize),rep('342',clsize),rep('343',clsize),rep('344',clsize),rep('345',clsize),rep('346',clsize),rep('347',clsize),rep('348',clsize),rep('349',clsize),rep('350',clsize),
rep('351',clsize),rep('352',clsize),rep('353',clsize),rep('354',clsize),rep('355',clsize),rep('356',clsize),rep('357',clsize),rep('358',clsize),rep('359',clsize),rep('360',clsize),
rep('361',clsize),rep('362',clsize),rep('363',clsize),rep('364',clsize),rep('365',clsize),rep('366',clsize),rep('367',clsize),rep('368',clsize),rep('369',clsize),rep('370',clsize),
rep('371',clsize),rep('372',clsize),rep('373',clsize),rep('374',clsize),rep('375',clsize),rep('376',clsize),rep('377',clsize),rep('378',clsize),rep('379',clsize),rep('380',clsize),
rep('381',clsize),rep('382',clsize),rep('383',clsize),rep('384',clsize),rep('385',clsize),rep('386',clsize),rep('387',clsize),rep('388',clsize),rep('389',clsize),rep('390',clsize),
rep('391',clsize),rep('392',clsize),rep('393',clsize),rep('394',clsize),rep('395',clsize),rep('396',clsize),rep('397',clsize),rep('398',clsize),rep('399',clsize),rep('400',clsize)
)

```

```
)
```

```
popx1<-data.frame(cl_id,x)
popx1$unit<-1:nrow(popx1)
```

```
popx <- data.frame(popx1$unit,cl_id,mu,v_j,e_ij,x)
colnames(popx)<-c('unit_id','cl_id','mu','v_j','e_ij','x')
```

```
pop.des=svydesign(data=popx,ids=~cl_id+unit_id, nest=TRUE)
pop.mean=data.frame(svymean(x,pop.des))
popmean <- pop.mean[1,1]#finite population mean
```

```
for (k in 1:nvar) {
```

```

##### Weights storage #####
for (j in 1:nsamp){

  w1<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1) #store weights
  w2<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
  w3<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
  w4<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
  w5<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
  w6<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
  w7<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
  w8<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1) #store weights
  w9<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
  w10<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
  w11<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
  w12<-matrix(c(rep.int(NA,j)),nrow=j,ncol=1)
}

#####

#####
# Two stage Cluster Random Sampling from a clustered finite population
#####

n2 = rep(20,50)    # select 10 units from 50 clusters

for (i in 1:nsamp) {

sample = mstage(popx, stage=list("cluster","cluster"), varnames=list("cl_id", "unit_id"),size=list(50,n2), method=list("srswor","srswor"))
#unique(sample[[1]]$cl_id)
#unique(sample[[2]]$unit_id)  #double checking

s=getdata(popx,sample)[[2]]

#####
# Naive weights
#####

mean_naive[i]=mean(s$x)
variance_naive[i]=var(s$x)
cv_naive[i]=abs(sqrt(variance_naive[i])/mean_naive[i])
skewness_naive[i]<-skewness(s$x)
kurtosis_naive[i]<-kurtosis(s$x)

#####
# Design conscious weights
#####

plan=svydesign(data=s,ids=~cl_id+unit_id, nest=TRUE)
mean_st=svymean(~x,plan)
mean_st=as.data.frame(mean_st)

means[i]<-mean_st[1,1]          #sample mean for ith sample
variance[i]<-mean_st[1,2]^2    #sample variance, design consistent variance
cv[i]<-abs(sqrt(variance[i])/means[i])    #coefficient of variation

# skewness[i]<-skewness(s[,i])
moments = svymean(~l(x^4) + l(x^3) +
  l(x^2) + l(x), plan)

skew=data.frame(svycontrast(moments,
  quote(`l(x^3)` - 3*`l(x^2)`*`l(x)` +
    3*`l(x)`*`l(x)^2` - `l(x)^3`/
    (`l(x^2)` - `l(x)^2`^1.5)))
skewness[i]<-skew[1,1]

#kurtosis[i]<-kurtosis(s[,i])

kurt=data.frame(svycontrast(moments,

```

```

quote(`l(x^4)` -
      4*`l(x^3)`*`l(x)` +
      6*`l(x^2)`*`l(x)`^2 -
      4*`l(x)`*`l(x)`^3 +
      `l(x)`^4)/
(`l(x^2)` -
 `l(x)`^2)^2))

kurtosis[i]<-kurt[1,1]

meff.1 <- data.frame(svymean(s$x,plan,deff=TRUE))
meff_var[i]=variance[i]/((1-(n/N))*(var(s$x)/n))

#####Weights calculation#####

w1[i]=n      #inverse of the sample size

w2[i]<-1/(variance[i]) #inverse of the design variance of the estimator of the mean
w3[i]<-1/(variance_naive[i]) #inverse of the naive variance of the estimator of the mean

w4[i]=1      #simple average weights

w5[i]=1/cv[i] #inverse of cv
w6[i]=1/cv_naive[i] #inverse of naive cv

w7[i]=abs(1/skewness[i]) #inverse skewness
w8[i]=abs(1/skewness_naive[i]) #inverse skewness

w9[i]=1/kurtosis[i] #inverse kurtosis
w10[i]=1/kurtosis_naive[i] #inverse kurtosis

w11[i]=meff_var[i]/variance[i] #in srs deff=1 so coincides with inverse of variance
w12[i]=1/meff_var[i]

combmeans1[i]<-round(sum(means[1:i]*w1[1:i])/sum(w1[1:i]),digits=9) #combine means using inverse sample size
combmeans2[i]<-round(sum(means[1:i]*w2[1:i])/sum(w2[1:i]),digits=9) # combine using inverse variance
combmeans3[i]<-round(sum(means[1:i]*w3[1:i])/sum(w3[1:i]),digits=9) # combine using inverse variance naive
combmeans4[i]<-round(sum(means[1:i]*w4[1:i])/sum(w4[1:i]),digits=9) # combine using a weight of 1
combmeans5[i]<-round(sum(means[1:i]*w5[1:i])/sum(w5[1:i]),digits=9) # combine using a weight of inverse CV
combmeans6[i]<-round(sum(means[1:i]*w6[1:i])/sum(w6[1:i]),digits=9) # combine using a weight of inverse CV naive
combmeans7[i]<-round(sum(means[1:i]*w7[1:i])/sum(w7[1:i]),digits=9) # combine using a weight of inverse skewness
combmeans8[i]<-round(sum(means[1:i]*w8[1:i])/sum(w8[1:i]),digits=9) #combine using a weight of inverse skewness naive
combmeans9[i]<-round(sum(means[1:i]*w9[1:i])/sum(w9[1:i]),digits=9) # combine using inverse kurtosis
combmeans10[i]<-round(sum(means[1:i]*w10[1:i])/sum(w10[1:i]),digits=9) # combine using inverse kurtosis naive
combmeans11[i]<-round(sum(means[1:i]*w11[1:i])/sum(w11[1:i]),digits=9) # combine using a weight of meff/var
combmeans12[i]<-round(sum(means[1:i]*w12[1:i])/sum(w12[1:i]),digits=9) # combine using a weight of 1/meff

}

sampmean[k]=means[1] # capturing the combined mean, where nsamp reps are combined, all weighting strategies

combmeans11[k]=combmeans1[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans21[k]=combmeans2[2]
combmeans31[k]=combmeans3[2]
combmeans41[k]=combmeans4[2]
combmeans51[k]=combmeans5[2]
combmeans61[k]=combmeans6[2]
combmeans71[k]=combmeans7[2]
combmeans81[k]=combmeans8[2] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans91[k]=combmeans9[2]
combmeans101[k]=combmeans10[2]
combmeans111[k]=combmeans11[2]
combmeans121[k]=combmeans12[2]

combmeans12[k]=combmeans1[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans22[k]=combmeans2[3]

```

```

combmeans32[k]=combmeans3[3]
combmeans42[k]=combmeans4[3]
combmeans52[k]=combmeans5[3]
combmeans62[k]=combmeans6[3]
combmeans72[k]=combmeans7[3]
combmeans82[k]=combmeans8[3] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans92[k]=combmeans9[3]
combmeans102[k]=combmeans10[3]
combmeans112[k]=combmeans11[3]
combmeans122[k]=combmeans12[3]

combmeans13[k]=combmeans1[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans23[k]=combmeans2[5]
combmeans33[k]=combmeans3[5]
combmeans43[k]=combmeans4[5]
combmeans53[k]=combmeans5[5]
combmeans63[k]=combmeans6[5]
combmeans73[k]=combmeans7[5]
combmeans83[k]=combmeans8[5] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans93[k]=combmeans9[5]
combmeans103[k]=combmeans10[5]
combmeans113[k]=combmeans11[5]
combmeans123[k]=combmeans12[5]

combmeans14[k]=combmeans1[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans24[k]=combmeans2[10]
combmeans34[k]=combmeans3[10]
combmeans44[k]=combmeans4[10]
combmeans54[k]=combmeans5[10]
combmeans64[k]=combmeans6[10]
combmeans74[k]=combmeans7[10]
combmeans84[k]=combmeans8[10] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans94[k]=combmeans9[10]
combmeans104[k]=combmeans10[10]
combmeans114[k]=combmeans11[10]
combmeans124[k]=combmeans12[10]

combmeans15[k]=combmeans1[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans25[k]=combmeans2[15]
combmeans35[k]=combmeans3[15]
combmeans45[k]=combmeans4[15]
combmeans55[k]=combmeans5[15]
combmeans65[k]=combmeans6[15]
combmeans75[k]=combmeans7[15]
combmeans85[k]=combmeans8[15] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans95[k]=combmeans9[15]
combmeans105[k]=combmeans10[15]
combmeans115[k]=combmeans11[15]
combmeans125[k]=combmeans12[15]

combmeans16[k]=combmeans1[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans26[k]=combmeans2[20]
combmeans36[k]=combmeans3[20]
combmeans46[k]=combmeans4[20]
combmeans56[k]=combmeans5[20]
combmeans66[k]=combmeans6[20]
combmeans76[k]=combmeans7[20]
combmeans86[k]=combmeans8[20] # capturing the combined mean, where nsamp reps are combined, all weighting strategies
combmeans96[k]=combmeans9[20]
combmeans106[k]=combmeans10[20]
combmeans116[k]=combmeans11[20]
combmeans126[k]=combmeans12[20]
}
expected1=mean(sampmean) #expected mean for one sample, not combining

expected=matrix(0,12,6)

```

```
expected[1,1]=mean(combmeans11) #strategy 1 nsamp=2
expected[2,1]=mean(combmeans21)
expected[3,1]=mean(combmeans31)
expected[4,1]=mean(combmeans41)
expected[5,1]=mean(combmeans51)
expected[6,1]=mean(combmeans61)
expected[7,1]=mean(combmeans71)
expected[8,1]=mean(combmeans81)
expected[9,1]=mean(combmeans91)
expected[10,1]=mean(combmeans101)
expected[11,1]=mean(combmeans111)
expected[12,1]=mean(combmeans121)
```

```
expected[1,2]=mean(combmeans12) #strategy 1 nsamp=3
expected[2,2]=mean(combmeans22)
expected[3,2]=mean(combmeans32)
expected[4,2]=mean(combmeans42)
expected[5,2]=mean(combmeans52)
expected[6,2]=mean(combmeans62)
expected[7,2]=mean(combmeans72)
expected[8,2]=mean(combmeans82)
expected[9,2]=mean(combmeans92)
expected[10,2]=mean(combmeans102)
expected[11,2]=mean(combmeans112)
expected[12,2]=mean(combmeans122)
```

```
expected[1,3]=mean(combmeans13) #strategy 1 nsamp=5
expected[2,3]=mean(combmeans23)
expected[3,3]=mean(combmeans33)
expected[4,3]=mean(combmeans43)
expected[5,3]=mean(combmeans53)
expected[6,3]=mean(combmeans63)
expected[7,3]=mean(combmeans73)
expected[8,3]=mean(combmeans83)
expected[9,3]=mean(combmeans93)
expected[10,3]=mean(combmeans103)
expected[11,3]=mean(combmeans113)
expected[12,3]=mean(combmeans123)
```

```
expected[1,4]=mean(combmeans14) #strategy 1 nsamp=10
expected[2,4]=mean(combmeans24)
expected[3,4]=mean(combmeans34)
expected[4,4]=mean(combmeans44)
expected[5,4]=mean(combmeans54)
expected[6,4]=mean(combmeans64)
expected[7,4]=mean(combmeans74)
expected[8,4]=mean(combmeans84)
expected[9,4]=mean(combmeans94)
expected[10,4]=mean(combmeans104)
expected[11,4]=mean(combmeans114)
expected[12,4]=mean(combmeans124)
```

```
expected[1,5]=mean(combmeans15) #strategy 1 nsamp=15
expected[2,5]=mean(combmeans25)
expected[3,5]=mean(combmeans35)
expected[4,5]=mean(combmeans45)
expected[5,5]=mean(combmeans55)
expected[6,5]=mean(combmeans65)
expected[7,5]=mean(combmeans75)
expected[8,5]=mean(combmeans85)
expected[9,5]=mean(combmeans95)
expected[10,5]=mean(combmeans105)
expected[11,5]=mean(combmeans115)
expected[12,5]=mean(combmeans125)
```

```
expected[1,6]=mean(combmeans16) #strategy 1 nsamp=20
expected[2,6]=mean(combmeans26)
expected[3,6]=mean(combmeans36)
expected[4,6]=mean(combmeans46)
expected[5,6]=mean(combmeans56)
```

```

expected[6,6]=mean(combmeans66)
expected[7,6]=mean(combmeans76)
expected[8,6]=mean(combmeans86)
expected[9,6]=mean(combmeans96)
expected[10,6]=mean(combmeans106)
expected[11,6]=mean(combmeans116)
expected[12,6]=mean(combmeans126)

```

```

expectedcol=c(expected1,expected[,1],expected[,2],expected[,3],expected[,4],expected[,5],expected[,6])

```

```

means=cbind(combmeans11,combmeans21,
combmeans31,combmeans41,combmeans51,combmeans61,combmeans71,combmeans81,combmeans91,
combmeans101,combmeans111,combmeans121,
combmeans12,combmeans22,combmeans32,combmeans42,combmeans52,combmeans62,combmeans72,combmeans82,combmeans92,
combmeans102,combmeans112,combmeans122,
combmeans13,combmeans23,combmeans33,combmeans43,combmeans53,combmeans63,combmeans73,combmeans83,combmeans93,
combmeans103,combmeans113,combmeans123,
combmeans14,combmeans24,combmeans34,combmeans44,combmeans54,combmeans64,combmeans74,combmeans84,combmeans94,
combmeans104,combmeans114,combmeans124,
combmeans15,combmeans25,combmeans35,combmeans45,combmeans55,combmeans65,combmeans75,combmeans85,combmeans95,
combmeans105,combmeans115,combmeans125,
combmeans16,combmeans26,combmeans36,combmeans46,combmeans56,combmeans66,combmeans76,combmeans86,combmeans96,
combmeans106,combmeans116,combmeans126)

```

```

variance1=(1/nvar)*sum((sampmean[,1]-expected1)**2)

```

```

variance11=(1/nvar)*sum((means[,1]-expected[1,1])**2)
variance21=(1/nvar)*sum((means[,2]-expected[2,1])**2)
variance31=(1/nvar)*sum((means[,3]-expected[3,1])**2)
variance41=(1/nvar)*sum((means[,4]-expected[4,1])**2)
variance51=(1/nvar)*sum((means[,5]-expected[5,1])**2)
variance61=(1/nvar)*sum((means[,6]-expected[6,1])**2)
variance71=(1/nvar)*sum((means[,7]-expected[7,1])**2)
variance81=(1/nvar)*sum((means[,8]-expected[8,1])**2)
variance91=(1/nvar)*sum((means[,9]-expected[9,1])**2)
variance101=(1/nvar)*sum((means[,10]-expected[10,1])**2)
variance111=(1/nvar)*sum((means[,11]-expected[11,1])**2)
variance121=(1/nvar)*sum((means[,12]-expected[12,1])**2)

```

```

variance12=(1/nvar)*sum((means[,13]-expected[1,2])**2)
variance22=(1/nvar)*sum((means[,14]-expected[2,2])**2)
variance32=(1/nvar)*sum((means[,15]-expected[3,2])**2)
variance42=(1/nvar)*sum((means[,16]-expected[4,2])**2)
variance52=(1/nvar)*sum((means[,17]-expected[5,2])**2)
variance62=(1/nvar)*sum((means[,18]-expected[6,2])**2)
variance72=(1/nvar)*sum((means[,19]-expected[7,2])**2)
variance82=(1/nvar)*sum((means[,20]-expected[8,2])**2)
variance92=(1/nvar)*sum((means[,21]-expected[9,2])**2)
variance102=(1/nvar)*sum((means[,22]-expected[10,2])**2)
variance112=(1/nvar)*sum((means[,23]-expected[11,2])**2)
variance122=(1/nvar)*sum((means[,24]-expected[12,2])**2)

```

```

variance13=(1/nvar)*sum((means[,25]-expected[1,3])**2)
variance23=(1/nvar)*sum((means[,26]-expected[2,3])**2)
variance33=(1/nvar)*sum((means[,27]-expected[3,3])**2)
variance43=(1/nvar)*sum((means[,28]-expected[4,3])**2)
variance53=(1/nvar)*sum((means[,29]-expected[5,3])**2)
variance63=(1/nvar)*sum((means[,30]-expected[6,3])**2)
variance73=(1/nvar)*sum((means[,31]-expected[7,3])**2)
variance83=(1/nvar)*sum((means[,32]-expected[8,3])**2)
variance93=(1/nvar)*sum((means[,33]-expected[9,3])**2)
variance103=(1/nvar)*sum((means[,34]-expected[10,3])**2)
variance113=(1/nvar)*sum((means[,35]-expected[11,3])**2)
variance123=(1/nvar)*sum((means[,36]-expected[12,3])**2)

```

```

variance14=(1/nvar)*sum((means[,37]-expected[1,4])**2)
variance24=(1/nvar)*sum((means[,38]-expected[2,4])**2)
variance34=(1/nvar)*sum((means[,39]-expected[3,4])**2)
variance44=(1/nvar)*sum((means[,40]-expected[4,4])**2)
variance54=(1/nvar)*sum((means[,41]-expected[5,4])**2)

```

```

variance64=(1/nvar)*sum((means[,42]-expected[6,4])**2)
variance74=(1/nvar)*sum((means[,43]-expected[7,4])**2)
variance84=(1/nvar)*sum((means[,44]-expected[8,4])**2)
variance94=(1/nvar)*sum((means[,45]-expected[9,4])**2)
variance104=(1/nvar)*sum((means[,46]-expected[10,4])**2)
variance114=(1/nvar)*sum((means[,47]-expected[11,4])**2)
variance124=(1/nvar)*sum((means[,48]-expected[12,4])**2)

```

```

variance15=(1/nvar)*sum((means[,49]-expected[1,5])**2)
variance25=(1/nvar)*sum((means[,50]-expected[2,5])**2)
variance35=(1/nvar)*sum((means[,51]-expected[3,5])**2)
variance45=(1/nvar)*sum((means[,52]-expected[4,5])**2)
variance55=(1/nvar)*sum((means[,53]-expected[5,5])**2)
variance65=(1/nvar)*sum((means[,54]-expected[6,5])**2)
variance75=(1/nvar)*sum((means[,55]-expected[7,5])**2)
variance85=(1/nvar)*sum((means[,56]-expected[8,5])**2)
variance95=(1/nvar)*sum((means[,57]-expected[9,5])**2)
variance105=(1/nvar)*sum((means[,58]-expected[10,5])**2)
variance115=(1/nvar)*sum((means[,59]-expected[11,5])**2)
variance125=(1/nvar)*sum((means[,60]-expected[12,5])**2)

```

```

variance16=(1/nvar)*sum((means[,61]-expected[1,6])**2)
variance26=(1/nvar)*sum((means[,62]-expected[2,6])**2)
variance36=(1/nvar)*sum((means[,63]-expected[3,6])**2)
variance46=(1/nvar)*sum((means[,64]-expected[4,6])**2)
variance56=(1/nvar)*sum((means[,65]-expected[5,6])**2)
variance66=(1/nvar)*sum((means[,66]-expected[6,6])**2)
variance76=(1/nvar)*sum((means[,67]-expected[7,6])**2)
variance86=(1/nvar)*sum((means[,68]-expected[8,6])**2)
variance96=(1/nvar)*sum((means[,69]-expected[9,6])**2)
variance106=(1/nvar)*sum((means[,70]-expected[10,6])**2)
variance116=(1/nvar)*sum((means[,71]-expected[11,6])**2)
variance126=(1/nvar)*sum((means[,72]-expected[12,6])**2)

```

```

variance=c(variance1,variance11,variance21,variance31,variance41,variance51,
variance61,variance71,variance81,variance91,variance101,variance111,variance121,
variance12,variance22,variance32,variance42,variance52,
variance62,variance72,variance82,variance92,variance102,variance112,variance122,
variance13,variance23,variance33,variance43,variance53,
variance63,variance73,variance83,variance93,variance103,variance113,variance123,
variance14,variance24,variance34,variance44,variance54,
variance64,variance74,variance84,variance94,variance104,variance114,variance124,
variance15,variance25,variance35,variance45,variance55,
variance65,variance75,variance85,variance95,variance105,variance115,variance125,
variance16,variance26,variance36,variance46,variance56,
variance66,variance76,variance86,variance96,variance106,variance116,variance126)

```

```

bias= expectedcol-popmean

```

```

relbias=(bias/expectedcol)*100

```

```

mse= variance + bias**2

```

```

cv= (sqrt(variance)/expectedcol)*100

```

```

s.e= (sqrt(variance)/sqrt(nvar))

```

```

LB= expectedcol - 1.96*s.e

```

```

UB= expectedcol + 1.96*s.e

```

```

result=cbind(popmean,expectedcol,relbias,s.e,cv,LB,UB,mse)

```

```

write.table(result,"result_at3_scenario1.txt",sep="\t")

```

```

write.table(meff_var,"meff_at3_scenario1.txt",sep="\t")

```