QATAR UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

TDEVELOPING A TWO-PHASE POST-STRATIFIED INVERSE SAMPLING TO REDUCE THE NONRESPONSE BIAS. CASE STUDY: STUDENTS' SATISFACTION SURVEY IN QATAR UNIVERSITY

BY

Mahmood Ahmed Hassan Ahmed

A Thesis Submitted to

the College of Arts and Sciences

in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Applied Statistics

June 2019

COMMITTEE PAGE

The members of the Committee approve the Thesis of
Mahmood Ahmed Hassan Ahmed defended on 13/05/2019.

Prof. Mohammad Mahmoud Salehi
Thesis/Dissertation Supervisor

Prof. Ayman Suleiman Bakleezi
Committee Member

Dr. Faiz Ahmed Mohamed Elfaki
Committee Member

Approved:

Ibrahim Alkaabi, Dean, College of Arts and Sciences

ABSTRACT

HASAN, MAHMOOD, AHMED., Masters : June : [2019]

Master of Science in Applied Statistics

Title: TDEVELOPING A TWO-PHASE POST-STRATIFIED INVERSE SAMPLING TO REDUCE THE NONRESPONSE BIAS. CASE STUDY: STUDENTS' SATISFACTION SURVEY IN QATAR UNIVERSITY

Supervisor of Thesis: Mohamed Mahmoud Salehi.

This research aims to develop a Two-phase Post-stratified Inverse Sampling (TPIS) to avoid or reduce the bias of nonresponse for students' satisfaction survey of Qatar university. In the first phase, the students are partitioned into two groups of response and non-response. The second phase method of data collection is an intensified method to get information from nonresponse. An unbiased estimator of the mean of overall satisfaction and its variance estimator are developed. The efficiency of the estimator is assessed using a simulation study. Moreover, the efficiency of (TPIS) is assessed. The cased study results showed that the mean estimate of overall satisfaction was higher for the nonresponse groups than the response groups but the different was not statistically significant at level of $\alpha = 0.05$ but significant at level of $\alpha = 0.10$. However, the overall satisfactions were significantly different at $\alpha = 0.05$ for 13 majors out of 51 (25%). This results indicate that the non-response bias is exists. In addition, simulation study results show that the TPIS is more efficient than a SRS with equal effective sample size for TPIS. Based on the results some recommendation and suggestion are provided.

ACKNOWLEDGMENTS

First of all, I would like to provide my gratefulness to my supervisor Professor Dr. Mohammed Mahmoud Salehi for his support, encouragement during the journey of my master study and for his supervision, help and guidance during the preparation of this thesis.

I would like to provide my thanks to the employees (faculty and staff) at College of Arts and Sciences, Qatar University for their support during this research work. In addition, thanking administrative staff of the Graduate Studies for their guidance and research workshops provided for graduate students. My grateful thanks goes to my friends and colleagues at QU for all the help, support provided during writing this Thesis.

I gained great experience during my studying at Mathematics, Physics and Statistics Department, thank you for all faculty members in this department, especially those who taught me the courses. Special thanks goes to Prof. Mouhammed Mosead Naser, for his encouragement and support during my master study. Special thanks also goes to Dr. Ahmed Al-Abbasi and Mr. Ozel A. Hurmuzlu for their editing and proofreading of this thesis. Thank you for my parents, relatives and friends, for their prayers for me to complete my study.

My thanks goes to my previous colleague at Institution Survey Research at QU Mrs. Wafaa Belhady for her support, encouragement, and help.

My highly appreciation goes to my wives, for their encouragement, patient, taking extra duties during my master study. Last but not the last, my gratefulness goes to my children for their understanding especially when I was not able to give my time to them in order

to complete this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: LITERATURE REVIEW

1.1Thesis Overview

The outlines of this thesis are as follows: *Chapter one* includes a review of the current literature on decreasing response rates, factors affecting the sample member's response behavior, the correlation between response rates and nonresponse bias; and it identifies nonresponse bias and its adjustment methods. Moreover, this chapter also focuses on the general inverse sampling design including general inverse of post-stratified sampling design (GIPSD), Research problems, objectives, and research significance. *Chapter two* describes the design used with the theoretical parts of the mean and variance estimators based on combining data collected from census followed up by inverse sampling for post-stratification. *Chapter three* evaluates the bias of the old methods used by Institutional Survey Research Section (ISRS), using simulation to estimate the mean and the variance of general inverse for post-stratified sampling and compare it with estimated mean and variance for simple random sample. *Chapter four* compares the bias estimator of mean and variance of general inverse of post- stratified sampling (GIPSS) and sample random sample (SRS) based on the simulation; in addition, the chapter summarizes the findings of this study, presents general remarks, and points to future research.

1.2 Introduction

Due to advanced technologies, different online survey software have been developed for collecting survey data for their advantages (Ronald D., Fricker .J, 2008, Lee M. H, 2011). The advantages of online questionnaire are mentioned in details in Duffy et al. (2005) and Bethlehem (2010). The speed and low cost are the key advantages. The first advantage is that distributing questionnaires using online software requires low cost compared to hardcopy questionnaire and no cost required in mailing, data entry and printing. Moreover, there is no need for a field researcher and that the influence of the field researchers is avoided. The second advantage of online surveys is that the researchers get the required data from respondents very quickly. In addition, using online surveys enable us to use sound, animation. Ronald and Fricker (2008) claimed that, the cost of collecting data using an online survey is approximately zero.

However, online surveys have disadvantages especially on sampling issues under coverage and self-selection. Researchers such as (Duffy et al. 2005, Bethlehem, 2010 and Lee 2011) mentioned that the online surveys lead to low response rates, which results in higher nonresponse rates increasing the chance of non-response bias in survey research. Grande, G., & Todd, C. (2002, p.171) states that "*the expressions for nonresponse bias in survey estimates come in early survey researchers used approaches that assumed nonresponse was affixed*"

Ye Cong (2012) answered a rational question, "are there any advantages for online surveys if the non-response rate is high? The answer of this question depends on the type of sampling design and efforts to overcome the factors affecting the decision of respondents. These factors include survey owner, advance notice, follow-up efforts, incentives and interest in the topic. Many researchers studied the results for each of these variables as follows:

1) The likelihood of response generally increases if they receive prior notification (Sarndal, C., Swensson, B., Wretman, J., 1992). 2) Respondents are more likely to collaborate when the survey owner is a government sector than private sector (Groves and Peytcheva, 2008). 3) Barclay et al. (2002) showed that response rates are strongly correlated with the interest in the topic of the survey. 4) Response rates increase when the number of attempts increases. 5) Providing different types of incentives is effective positively in increasing response rate (Barcly.C, 2002).

Similarly, sending a personal email, privacy concerns, and language problem may effect on the response rates (Dillman 1991).

From the researcher's experience at QU, another factor may affect the response rate; that is the number of survey received by respondents during the limited period of time, which leads to survey fatigue. Even though, the survey section at QU exerts tireless efforts to overcome the previous factor such as design a surveys with Arabic and English languages, reviewing surveys by experts, using incentives, sending the survey from the vice president office, the response rate still small. For example, the response rate of student satisfaction survey for four cycles not exceeding in average the 20% which means that the non-response rate is approximately 80%. The question arises: is the 20% of response rate adequate?

Previous studies (Lohr, S. , 2010; Robert M. Groves ,2006) provided a criteria for response rate required for analysis and report, these criteria are: at least 50 % response rate is appropriate; 60% response rate is good; a 70% response rate is very good. Collecting data using interview surveys, 85 % response rate is adequate; less than 70 %, the non-response bias may have a chance to exist. Based on these criteria, the 20% response rate of Student Satisfaction is considered not appropriate and may lead to bias estimator, so this study needs to evaluate this bias.

1.3 Nonresponse Error

The main purpose of a questionnaire is to collect information about a population from a sample or from a census. The results driven from the data may be affected by several sources of error. Some errors arise from the survey itself which is called survey errors in addition to some other errors (Ronald D., Fricker, J, 2008). Groves (2006) mentioned several sources of survey error such as: coverage, sampling, nonresponse and measurement errors.

Coverage error occurs when some persons in the population do not have chance to be selected or some of population parts are not included in the sample. In addition, cover error exists when the population frame is not fully complete. To reduce the coverage error, we should obtain the sampling frame as fully as possible, stratify the sample and use the weight of each strata to match the population of inference on the important characteristics observed.

The sampling error arises when different samples from same population produce different survey data and is reduced by increasing sample size. Nonresponse error is the failure to collect data on all units in the sample 'and it happens when data is not collected on either individual responses (nonresponse unit) or individual survey item (Groves, 2006). The response rate helps in calculating the sampling error and gives the researcher an idea about confidence on generalizing the result to all population. Higher response rates indicate a lower likelihood of nonresponse bias.

Non-response bias is one of the most types of errors that a survey suffers. It is a systematic error that occurs because individuals differ in their accessibility and tendency to participate in the questionnaire according to their characteristics as well as the characteristics of the survey itself. (Vandenplas C. et al, 2015).

Ronald D., Fricker, J, (2008) defined the measurement error as an accuracy of responses recorded on the toll of collecting data. It arises when there is a difference between respondents' responses to the real responses. For example, respondents may not answer certain questions honestly for different reasons, or they may answer some questions incorrectly. Therefore, to minimize the measurement error, a researcher should first ensure that the survey is valid and reliable before disseminating the survey to the sample. In other words, the survey should be reviewed, free of sensitive questions, and written by clear language to be clearly understood to the respondents (Ronald D., Fricke J., 2008).

Bethelehem (2010, p172) *"provided a taxonomy of survey errors as shown in the figure*

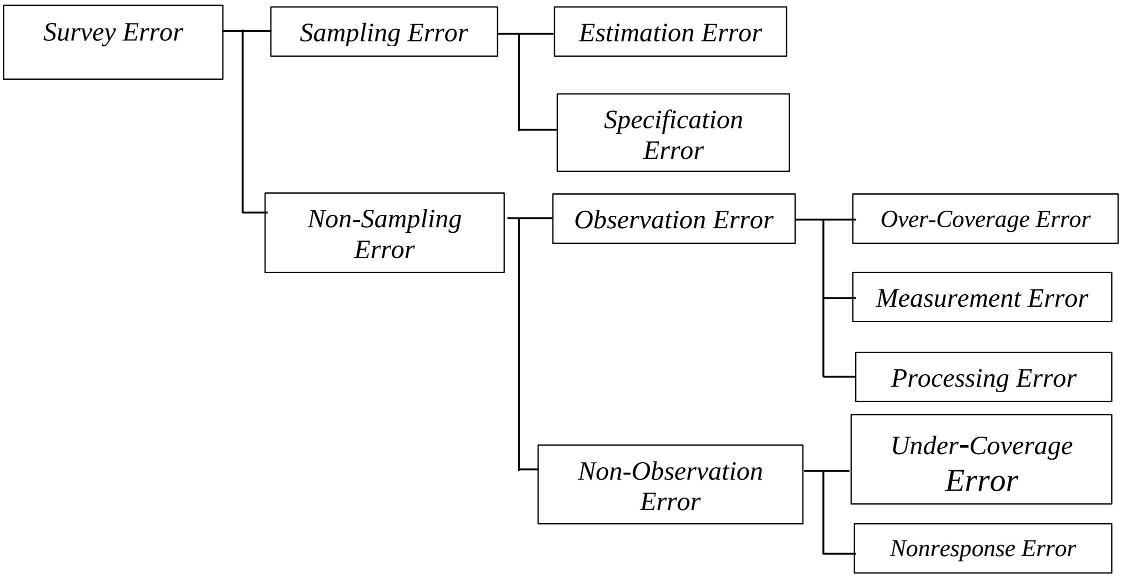

Figure 1.1 Type of survey errors, taken from Bethelehem (2010, p.172)

The difference between the estimator calculated from a sample and the corresponding property in the population is called total survey error. This can be classified into sampling error and non-sampling error. The sampling error happens when some parts of population are not represented in the sample, sampling error disappears when the whole population is observed.

Sampling error is divided into two parts: estimation errors and specification errors. Estimation error occurs because new samples lead to different estimates. It is inevitable, but it can be measured by applying probability theory. Some Researchers use online survey to collect data from non-probability sample drawn by non-probability selection (self-selection).

Specification error occurs if the true selection probabilities differs from the selection probabilities specified in the sampling design. As stated before, in the case of online surveys, Researchers cannot control the specification error because selection probability is unknown.

Non-sampling error occurs due to multiple reasons such as unreliable and invalid questionnaires; low response rate; error in data entry and so on. This type of error can be classified into two parts: observation error and non-observation error.

Observation error happens during collecting and recording responses. This can be classified into three types of error: measurement error, processing error and over-coverage error. Measurement error occurs when survey response differs from the real value. It happens as results of respondents' misinterpretation the survey items; respondents not tell the truth as a result of sensitive questions, mistakes happen by reviewers and so on. Processing error defined as an entry error. An over-coverage error occurs when some individuals are not in the target population included in the sample or because of duplication of selection, some individuals from the population frame.

Non-observation error is the error resulting from the deletion of the intended measurements. This type of errors classified into under-coverage error and non-response error. Under-coverage error arises when some individuals in the population cannot be in the sampling frame or cannot be invited or connected to participate in a survey. Non-response error occurs when selected people refuse to answer the survey questions. All types of error mentioned above are common errors in the online survey; however, self-selection error (under-coverage error and sampling error) may be the most one occurs among all types of errors.

1.4 Nonresponse Rates and Nonresponse Bias

Nonresponse is the failure to find a response from the same individuals in a sample. It is of concern to survey methodologists and practitioners because complete response is assumed by the randomization or design-based theory that allows inference from a sample to the target population. Nonresponse has ability to introduce bias into survey estimates and reduce the precision of survey estimates. As a result, survey practitioners make efforts to minimize nonresponse and its effects on inferences from sample surveys. However, even with the best efforts, there will be nonresponse; hence, it is essential to understand its potential effects and methods that can be used for limiting these effects (Brick JM, Montaquila JM, 2009).

When the difference between the answer of respondents and non-respondents are significantly different, the non-response may be considered as a real problem. In other words, non-response bias depends on both nonresponse rate and the difference between respondents and non-respondents. Non-response happens when a respondent does not answer certain items in the questionnaire or when a person does not accept to participate in a survey. (Ronad R. et al., 2015).

Regarding how many response rate we need, there is no magic response rate number. The magic answer is: the survey results are representative to the population when the response rate is high; however, attaining these high response rates can be very costly. On the other hand, low response rates increase the likelihood that survey results will not be representative to the population as a whole (Ronald R. et al., 2015).

Elizabeth D. McInnis, BA, MA (2006) defined nonresponse bias as a bias that happens in survey results when the answer of respondents to a survey are different from the answer of those who did not respond in terms of some variable (eg. demographic variable or other variables measured by the survey. According to DDS Research, Inc., (2004) the existence of nonresponse bias is a threat to the external validity of research results. A general formula for measuring bias which present by Elizabeth D. McInnis, BA, MA (2006) is:

$$\text{Bias}\ (\bar{Y}_r) = \frac{M}{N}\ (\bar{Y}_r - \bar{Y}_m), \qquad (1.1)$$

Bias ($\bar{Y}_r$) : The nonresponse bias of respondent mean; $\bar{Y}_r$: the mean of respondents in the target population; $\bar{Y}_m$: the mean of the no respondents in the target population, M: the number of no respondents in the target population, and: N: the total number in the target population.

If we select a sample Bias ($\bar{Y}_r$) = P ($\acute{y}_r - \acute{y}_{\mathrm{m}}$) , where P: non-response rate , y is the answer based on observed responses; $P = ((n - r)/n)$

Since (n) is the sample size, (r) is the number of respondents. ($\acute{Y}r$) is the respondent mean in the sample. ($\acute{Y}$m) is the non-respondent mean in the sample.

Based on equation (1.1) which is provided by Elizabeth D. McInnis, BA, MA (2006), the increase in either non-response rate or in the difference between the observed and non-respondent answers will lead to an increase in bias. According to Robert M. Groves (2006), this expression indirectly assumed all other sources of bias are fixed.

There is a growth in nonresponse rate as a result of using online survey to collect data, especially in social science, which has led to increased anxiety about the risk of non-response bias. Raphael Nishimura1 et al, (2016) provided in their simulation study a set of indicators of when non-response bias is possible to occur and examine how each of these indicators performs in a variety of situations.

Since non-response bias has an effect on generalizing sample results to the population, it is very important to discuss the identifying unit non-response bias in the next section.

1.5 Identifying Non-Response Bias

Related to Question mentioned in section 1.3 which is “How do we know if there is nonresponse bias?” Several studies have focused on answering this question (e.g. Sinibaldi, Trappmann, & Kreuter, 2014, Roberts, Vandenplas, & Ernst Stähli, 2014). Some methods used to identify non-response bias are such as; conducting a nonresponse follow-up survey, comparing Initial and Late Respondents, comparing Survey results to known population parameters and using known database variables (Vandenplas et al. 2015).

In some studies, (e.g., Sinibaldi, Trappmann, & Kreuter, 2014, Roberts, Vandenplas, & Ernst Stähli, 2014), non-response follow-up surveys (NRFS) are considered as a possible source of information about non-respondents. The advantage of NRFS is that the variables included in the survey can be chosen to be associated with response tendency and interest variables. However, it loses the benefit of having information about the entire sample, where there may still be a group of non-respondents that participate neither in the main survey nor in NRFS. According to Roberts, Vandenplas, & Ernst Stähli (2014), the extent of this issue depends on how far these respondents to the NRFS differ from those non-respondents to the NRFS. This difference can be assessed with para data.

To increase the response rate for non-response, these surveys should be short while still gathering useful information. It is therefore important to identify the most relevant questions to be a part of the survey

1.6 Adjusting Non-Response Bias

The methods of identifying non-response bias, which mentioned in the previous section, give us information about whether the bias exists or not. However, it does not give us any information about how we deal with it. In literature, several studies tried to reduce the non-response bias using different Adjustment methods (e.g. Robert M. Groves ,2006; Bethlehem, 2010; Ye Cong, 2012; Vera Toepoel & Matthias Schonlau ,2017; Rueegg, et al. ,2017; Tianji Cai and Hongyu Wang, 2018). According to (Vera Toepoel & Matthias Schonlau , 2017) The purpose of the non-response adjustment is to remove or avoid non-response biases while maintaining the accuracy of the estimate.

The following procedures are used for adjusting non-response bias.

1.6. 1 Weighting Adjustment

Weighting adjustment is a method that attempts to improve the precision of survey estimates using auxiliary variable (Bethlehem, 2010). Assume that additional variables are available (e.g. demographic data) for all individuals of the population. Based on this available variable respondents and non-respondents sub-population are created. Weights are then calculated based on the proportions in each sub-population and applied to the respondents to reflect the total population. Comparisons on the variables of interest are made between adjusted and unadjusted weighted respondents. If there is a significant difference, then non-response bias is supposed to be existing and the weighting adjustments are the alternative way to reduce the bias of results.

1.6.2 Stratified and Post- Stratified Sampling

Stratification is another technique similar to weighting adjustment, except that the procedure uses population size instead of the total sample size (Lin and Schaeffer 1995; Curtin, Presser, and Singer 2005) . Based on Robert M. Groves (2006, P. 11), comparison between respondents selected by any sampling technique in stage one and respondents selected by stratification in stage 2 is made.

To draw a sample using stratified sampling design, the first step is to divide the population of size into groups of $N_1, N_2, \ldots., N_L$ units. These groups are called strata. The summation of these strata must be equal to the whole population, so that $N_1 + N_2 + \cdots + N_L = N$. In stratified sampling, each stratum has a degree of homogeneity between its units and the researcher should also ensure that there is no overlaying among these strata. For the full benefit from stratified sampling design, the size of population (N), should be known. When the strata have been determined, a sample is drawn from each, the selection being made independently in different strata. The sample sizes within the strata are indicated by $n_1, n_2, \dots, n_l$, respectively. If unites are selected from each stratum by simple random sample design, the whole process is defined as stratified random sampling (Ozturk O. ,2014).

If the stratified sampling happened in the second stage of sampling, it is called post-stratified sampling, which is considered as a calibration estimation technique to reduce the estimator variance and to reduce the bias that occurs as a result of non-coverage and/or non-response (Cervantes et al., 2009). In a survey research, it is important to have a representative sample of the population. However, the absence of such a sample often occurs intentionally or unintentionally.

For example, the distribution of a particular property such as demographic variables (age, education level, sex, nationality) may differ in the sample from the distribution of the population. When responses are linked to demographic variables, this leads to increase of the probability of bias, for the reason that statistical procedures will give greater weight to the variables that have high response arte. Post-stratification technique provides a solution to overcome this problem by stratifying the sample into a number of cells, based on important features of the population, and then more weight is given to respondents with low response rate and less weight is given to those with high response arte.

Doss, et al. (1979) provided an equation for the post-stratification weights which can be used for one or more demographic variables based on the percentage of those variables for stratum in both the sample and the target population. The weight W_i for an component i in stratum h is equal to: $W_i = \frac{\left(\frac{N_h}{N}\right)}{\left(\frac{n_h}{n}\right)}$ Where N is the population size, N_h is the sub- population size for stratum h, n is the sample size, and n_h is the sample size in stratum h. Specific weight is given to all elements in the same stratum. Note that this equation cannot be applied if the population proportions in each stratum are not available

A judgment post-stratification sample technique, discussed by Ozturk (2014), begins with a simple random sample and uses further experimental units to produce post-strata among the observations that are already measured. The creation of this sampling technique requires selecting a sample size n and set size z. One then selects a simple random sample, X_i ; $i = 1, \dots . n$, of size n and measures all of them. For the i^{th} measured unit, a JPS sample needs $z - 1$ additional units to form a set of size z. The units in this set are ranked from smallest to largest without a measurement, and the rank of the measured unit, on which Xi is already measured, is recorded. The full judgment post-stratification sample data then consists of n measured values and n ranks associated with these measured values, $(Xi, Ri), i = 1, \dots, n$, where R_i is the rank of X_i.

Singh. S (2003) in his book (Advanced Sampling Theory with Application) defended post-stratification as a sampling design whereby a simple random sample is selected with or without replacement and he provided equations for estimate the mean and the variance based on this sampling design. These equations will be mentioned in chapter two when we describe the new sampling design (see also appendix A and B). Finally, although there are many advantages for this approach, there are some weakness such as the differences between respondents and non- respondents are assumed to be captured in the subgroups, and that there is no rule of thumb for comparing adjustments to determine which to use. In addition, some stratum include small number of response, so that if the sample unite from any stratum = 0, the value of n_h in the denominator will be zero and then the results will be infinite. One of the solution of this issue is using inverse sampling for post-stratification design. This design suggested by Seber G.A.F., Salehi M.M. (2012) and it will be discussed in detailed in section 1.7 .

1.6.3 Other Adjustment Techniques

There are several other techniques to control non-response bias such as propensity models, which require some information (such as demographics) to be known for the whole population or any calibration methods, which make the use of auxiliary variable data similar to the data from census (Rueegg .C, et al. ,2017).

In the Swiss Childhood Cancer Survivor Study which was conducted by Rueegg C., et al. (2017), a comparison was made on the features of those who responded without any reminder, late respondents responded with one reminder and those who did not respond at all. In addition, based on the information available from the Swiss Childhood Cancer Registry for all population, they compared observed prevalence of typical outcomes in responders to the expected prevalence in a complete representative population. The complete population is generated using inverse probability of participation weights in order to estimate the impact of nonresponse bias. Results of comparison show that, nonresponse bias has only a small impact on the results of childhood cancer survivor study, so that the researcher has confident to generalize the results of this study to the whole population .

Ye Cong (2012, p .7) mentioned in his Ph. D Dissertation that, the high non-response rate leads to low confidence in the results. Many researchers focused on the impact of nonresponse, which is considered as a big problem, because the true values for the non-respondents are mostly unknown. It is not easy to estimate the amount of non-response bias or to avoid it. So many researchers focused their effort on developing effective adjustment sampling design to eliminate or at least reduce the impact of nonresponse bias.

Ye Cong (2012) used two set of data containing records for each individual in the list to assess the effectiveness of adjustment methods that developed to avoid nonresponse bias; to examine sample quality indicators, and to study the relative amount of non-response bias under different techniques such response propensity weighting (RPW) and generalized regression weighting (GRW).

The results show that both RPW and GRW are not effective in reducing nonresponse bias present in the study data. There are also some declines in error, but these declines are limited. the result shows that there is a small difference on same set of auxiliary variables between these two techniques. Finally, the researcher found that comparing the bias of non-response to other types of errors indicated that bias of non-response in these two sets of data may be greater than the bias of sampling and coverage errors, but measurement bias can be the largest, especially when sensitive questions are not answered.

Another study by Tianji Cai and Hongyu Wang (2018) studied four familiar weighting adjustments: 1) logistic regression model, 2) response propensity stratification (RPS) method, 3) generalized exponential model (GEM), and 4) the random forests model (RFM) for longitudinal non-response.

The results indicated a similarity in the results of the logistic regression model, the RPS and GEM method, while there was a slight difference in the results of GEM and RPS methods in many cases, whereas the results obtained by RFM were not as reliable as those of other methods. For categorical variables, there were significant differences in cross-category bias while the total difference compared to the baseline was not significant. In addition, relative bias and mean square error changes if correlated to the specification of non-response model, the baseline weight, as well as the intervals between the baseline, and the wave for weighting adjustment.

In conclusion, any Researcher using sampling survey needs to feel confident that the results from analyses are speaking the truth by following some steps to ensure the data meet certain standards. Among these steps: screener survey targets the specific audience he /she wants, takes many steps to ensure the samples contain the target we need, weights respondents to match specific population demographics, etc. However, one of the most over-looked problems is that of non-response bias.

In practice, any Researcher using online survey cannot fix all the mentioned factors that affect the response rate, and then nonresponse bias exists. Based on Rueegg C, et al. (2017) this nonresponse bias affects inferences drawn from online surveys across different populations. Many studies such as Lin and Schaeffer (1995), Groves (2006) investigated biased estimators as the result of non- response rate using different sampling methods. However, this study attempted to evaluate the non-response bias using a new technique named, Tow-phase Post-stratified Inverse Sampling Design.

1.7 Follow up Post Stratified Inverse Sampling

1.7.1 Post Stratification Sampling Design

Post-stratification is a sampling technique, which is used in the case whereby researchers would like to stratify on a key variable, but they are unable to find the required units from some strata after the sample is selected. Personal characteristics such as marital status, income, nationality and educational level are common examples of a key variable. The procedure consists of the following three steps: Take a simple random sample; classify the above sample into strata; and use the classified data to estimate the unknown population parameter by the usual method of stratified random sampling (Hang K., CLiu J., Chien-Pai H ,1998).

The post-stratification technique is somehow precise compared to simple random sample, if each stratum weight is known, and the sample is reasonably large (Cochran, 1977 & Scheaffer et al. 1990). However if the sample is small, the post-stratification may raise the following question : What can a researcher do if there is no response from one or more stratum? Doss et al. (1979) have answered this question.

To discuss their answer, the notations used in their paper will first be introduced as follows:

N: population size	; N_h: h^{th} stratum size
n: sample size ;	; n_h sample size in hth stratum
μ: population mean	;μ_h: population mean of the h^{th} stratum
σ^2: population variance	σ_h^2 variance of the h^{th} stratum
Y_i: value obtained for the ith unit in the sample	; Y_{hj}: value of the j^{th} unit in the subsample
L: number of strata	W_h: h^{th} stratum weight
y: sample mean	$\bar{y}_h$: sample mean of h^{th} stratum
s^2: sample variance	
$\bar{y}_p$: estimator of y by post stratification	s_h^2 : sample variance of h^{th} stratum

Based on the aforementioned notations, the Doss et al. (1979) made use of the following formulae:

$$\mu = \sum_{h=1}^{l} W_h \ \mu_h, W_h = \frac{N_h}{N}, \bar{y}_h = \frac{1}{n_h}\sum_{j=1}^{n_h} y_{hj} \ , \ h = 1,2, \dots \dots L,$$

$$\bar{y} = \frac{1}{n}\sum_{h=1}^{n} y_i \ , \bar{y}_p = \ W_h \bar{y}_h,$$

Where $\bar{y}$ and $\bar{y}_b$ are estimators of μ obtained from simple random sampling and post-stratification ($n_h \geq 1$, h = 1,2 .... , L), respectively. It is worth noting that the sample size n of the first stage (simple random sample) is a fixed number, while the sample size n_h taken in the second stage is considered as random variables. Let $\bar{y}_c$ (l) be the estimator of μ under this procedure. Then, when L = 2, we have

$$\bar{y}_c(2) = \begin{cases} \bar{y}_p \text{, if } n_{1,} n_2 \neq 0 \\ \bar{y}, \quad \text{otherwise.} \end{cases} \qquad (1.1)$$

When L $\geq$ 3, they assume the following prior information throughout their paper so that it is easy to define 'neighboring strata'.

(X1) $\mu_1 \leq \mu_2 \leq \cdots .. \leq \mu_l$. (X2) $\mu_h - \mu_{h-1} \leq \mu_{h+1} - \mu_h$; $h = 2, \dots, L-1$

The above assumptions are realistic in practical survey sampling. For example, to estimate the population income mean, it can be classified into three strata: lowest, medium and highest; and the differences among lower income levels are usually less than those among higher income levels. Under assumptions (X1) and (X2), the h^{th} stratum is the 'neighboring stratum of the (h + 1)th stratum', $h = 1, \dots, L - 1$. (Chang et al. ,1998).

1.7.2 Inverse Sampling Design

Inverse sampling is an adaptive method used where the sampling is continued until a trigger number of units is attained (Salehi M. and Seber G.,2004, Enrico F., Tomasz Z. , 2018).

Seber G., Salehi M. (2013) focused on more traditional inverse sampling where the sampling is ongoing until a fixed number of individuals has been obtained. They introduced Murthy's estimator and gave a proof using the Rao-Blackwell theorem that allows the estimator to be used for more general sampling schemes.

When Seber G., Salehi M. (2013, pp. 50, 51) providing the new prod of Murthy's Estimator, they began by assuming that the sample size may be random and they define v to be the number of distinct units in the sample so that $sR = \{i1 , i2 , \dots, iv \}$, the unordered distinct units. Let J_i be an indicator variable that takes the value 1 (with probability p_i) when the i^{th} unit is selected as the first unit, and 0 otherwise. As $[J_i] = p_i$, a trivial unbiased estimator of μ is given by"

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N}\frac{y_i}{p_i}J_i$$

Since D_R , the random variable with value $D_R = \{(i, yi) : i \in sR \}$ is sufficient statistics for $\boldsymbol{\theta} = (y1 , y2 ,\dots, y_N)$

by Theorem provided in Thompson and Seber (1996) which is:

"*Consider an adaptive or conventional design in which the selection probability of the sample does not depend on any of the y-values outside the sample. (The probability may depend on y-values within the sample and may depend on the order of selection.) Then DR is a minimal sufficient statistic for **θ)"*** **(**in Seber G., Salehi M, 2013, p. 28).

Seber G. and Salehi M. (2013) also use the Rao-Blackwell theorem to obtain the unbiased estimator and as shown below. They approved that:

$$\hat{\mu}_{RB} = \hat{\mu}_M$$

As $P(S\_R \mid i\ ) = 0$ if the unit i is not in S_R . We see then that $\hat{\mu}_{RB}$ is Murthy's estimate (Murthy 1957 in Seber G.A.F., Salehi M.M. (2012, p. 4)

$$\hat{\mu}_M = \frac{1}{N} \sum_{i=1}^{n} y_i \ \frac{P(S_R|i)}{P(S_R)} ,$$

$$Va\hat{r}[\hat{\mu}_M] = \frac{1}{N^2} \sum_{i=1}^{n} \sum_{j<i}^{n} \left(\frac{P(S_r|i,j)}{P(S_R)} - \frac{P(S_R|i)P(S_R|j)}{[P(S_R)]^2} \right) \left(\frac{y_i}{p_i} - \frac{y_j}{P_j} \right)^2 P_i p_j ,$$

Based on Lavrakas, P. J. (2008) and Tang M.L., Ng H.K.T. (2011), inverse sampling is taken from a negative binomial distribution in that a series of Bernoulli experiments are carried out until a predetermined number of successful cases occurs. Under this design, the total sample size is considered as a random variable. Based on the binomial distribution, the traditional estimates of the attribute's probability occurrence are biased. However, it can be shown that if the total size of the sample is X, the unbiased amount of the minimum uniform variance is:

$$\hat{P} = \frac{r-1}{X-1}$$

Tang M.L., Ng H.K.T. (2011) provided a Statistical Model and Inference for study using inverse sampling whereby they continue to select individuals until the pre-determined number *r* (≥1) of certain characteristics of interest are obtained. They supposed that *Y* be the number of individuals without the characteristics of interest.

Finally accumulated in the sample before we obtain the first *r* individuals which are in order. They denote the probability of choosing the individuals with the characteristics of interest randomly as p, where $0 < p < 1$; according to them the random variable Y follows a negative binomial distribution with parameters r and p with probability mass function.

$$f(y|p) = \Pr(Y = y|p) = \binom{r+y-1}{y} p^r (1-p)^y \ , y = 0,1,2, \ldots$$

In practice, they used the maximum likelihood estimator (MLE) and the uniformly minimum variance unbiased estimator (UMVUE) to estimate p.

The MLE of p is given by:

$$p^{\hat{}} = \frac{r}{N}$$

Where $N = r + Y$ is the total number of trails required to obtain the predetermined number r.

Tang M.L., Ng H.K.T. (2011) Shaw that the variance of $p^{\hat{}}$ is

$$Var(\hat{p}) = \frac{P^2(1-P)}{r}$$

and noticed that MLE is actually a biased estimator of p and that UMVUE of p can be obtained by:

$$\hat{p} = \frac{r-1}{N-1}$$

$$var(\hat{p}) = (r-1)(1-P)\left[\sum_{k=1}^{r-1} \frac{(-p)^k}{(1-p)^k(r-k)} - \left(\frac{-P}{1-p}\right)^r \log(p)\right] - P^2$$

And an unbiased estimator of $var(\hat{P}) for\ r > 2$ is given by:

$$V\hat{a}r(\hat{P}) = \frac{\hat{P}\ (1-\hat{P}}{N-2} \qquad (ibid)$$

In the case of small samples, it is possible to attain zero sample sizes in some strata, which makes sample not representing the population effectively. Thus, Seber G., Salehi M (2012), tried to find a solution to this problem by a multiple inverse sampling scheme to ensure that each stratum has specified number of observations. Simulation study was conducted to compare the estimator obtained from the multiple inverse sampling with some other estimators obtained from other sampling design. The multiple inverse sampling is more efficient compared to simple random sample in identifying unbiased estimator.

Under a simple random sampling design, the ratio estimator would be undefined if all observations of the denominator variable are zero. Moradi M., Salehi M.(2007) considered the inverse sampling design as a natural solution to this problem whereby the researcher continues to take a sample until at least a predetermined number of nonzero values is observed for the denominator variable.

Salehi and Seber (2004) controlled the problem of an undefined ratio estimator by proposing general inverse sampling, which is a more realistic version of inverse sampling. They derived asymptotic unbiased estimator of the ratio and an approximate variance estimator for a general inverse sampling design using Taylor expansion. Based on a real population, simulation study was used to evaluate the efficiency of the developed estimator compered to SRS. .

Moradi M. Salehi M, (2007) claimed that when the population is divided into two subpopulations, one of which contains only a few units, the appropriate technique to be used is an inverse sampling design. Using this design, they derive the Horvitz-Thompson estimator for the population mean, where subpopulation size is known. They then introduced an alternative unbiased estimator, corresponding to post-stratification design. Using a simulation study, they found that the alternative estimator is an efficient estimator compared to the Horvitz-Thompson estimator (in Salehi M., 2012, p. 333).

General inverse sampling is generally a more appropriate sampling design than SRS when the event of interest is rare, and when estimator of the parameter of interest is likely to be undefined (Salehi M., 2012, p. 333) . In the past decade it has received considerable attention by Changet al (1998, 1999), Christman and Lan (2001) and Salehi & Seber (2001) to name a few among many others. One deficiency, however, is that the final sample size is not fixed, which makes it difficult to plan budgets and survey logistics. As a result, surveys having an inverse sampling design are rarely used in practice. To deal with this problem, Salehi and Saber (2004) proposed the following design:

Suppose that we can select at least n_0 and at most, n_1 units based on a minimum and a maximum budget. We first take an initial sample of size n_0. If we have the pre-determined number of events in the sample we would stop sampling. Otherwise we would keep sampling until we either achieve the pre-determined number of events or reach the sample size n_1. This sample design is called general inverse sampling (GIS). Salehi and Saber (2004) used Taylor series approximation to derive for any sampling design an estimator of the ratio as well as its variance and its variance estimator based on the Murthy estimator. To accomplish this, they used the Murthy estimator for deriving the variance estimator. Also they derived a formulation for general inverse sampling (in Moradi M., Salehi M. ,2007, pp.137 -138).

For traditional inverse sampling, population is partitioned into two subpopulations with unknown sizes and usually one of the subpopulations is rare (Haldane, 1945; Christman and Lan 2001). Chang et al. (1998) introduced a multiple inverse sampling (MIS) and supposed that population can be partitioned into subpopulations (post-strata) with known sizes. He claimed that MIS may avoid the undesirable events of obtaining no sample unit or very small sampled units from some post-strata in post-stratified sampling design.

Salehi & Chang (2005) used a Truncated Multiple Inverse Sampling (TMIS) and developed estimators and their variance estimators for the proportions of subpopulations under this sampling design. Using a simulation study, they found out that the TMIS estimator is reasonably efficient for estimating the proportions of rare subpopulations.

To ensure that a reasonable number of cases are obtained, the researcher considers the use of two-phase post-stratified inverse sampling (TPIS) design. This sampling design to be used will be discussed clearly in chapter two, using simulation to evaluate the developed estimator, based on a real population from Student Satisfaction Survey applied at Qatar University in 2017-18. Moreover, we compute its relative efficiency over the usual SRS estimator.

1.8 Problem Statement and Research Questions:

The Survey Section at Qatar University is responsible for conducting several types of surveys to evaluate QU services along with Key Performance Indicators (KPI). One of these surveys is Student Satisfaction Survey which is well-developed to assess the satisfaction rate among the undergraduate student at QU. However, the estimation of satisfaction rates of this survey is bias due to non-response rate. Thus, this study focuses on finding solutions to how to decrease or perhaps eliminate the bias of the satisfaction rate estimator and its variance estimator, which results from non-response rate by using a new techniques named, Follow Up Inverse Sampling for Post Stratified Sampling Design to Deal with Non –Response Bias and comparing this technique with simple random sample.

This study tries to answer the following two questions:

Q1: Are there any significant differences between Response and Non-response student on the satisfaction rate estimator?

Q2- By running a simulation study, what is the effectiveness of Two-phase Post-

Stratified Inverse Sampling (TPIS) Design compared to simple random sample?

1.9 Research Objectives

This research aims to:

- Assess Non-response bias in student satisfaction at Qatar University.
- Develop a new sampling technique through post stratified with different methods of data collection in order to reduce or eliminate Non-response bias effects.
- Investigate the effectiveness of this sampling design by running a simulation study and comparing this technique a simple random sample.

1.10 Research Significance:

This research is important for the following reasons:

- The present research study may provide a new sampling design, which can be a benchmark for all Institutions.
- The results of this research may supply institutional research at QU a new sampling method which will aid in decision-making.
- The results of this research will provide the institutional research with recommendations and suggestions that may increase efficiency of implementation of this sampling technique in large institutions.
- This research is considered an extension for other research works in the field of implementation of sampling techniques to collect accurate data.

CHAPTER 2: A NEW SAMPLING DESIGN

2.1 Describing the Old Method Used by Institutional Survey Research Section at QU

Institutional Survey Research Section (ISRS) at QU is responsible for conducting Institutional survey research to measure QU KPIs and to provide the necessary indicators and evidences to decision makers to tackle areas of concerns. A summary of the procedures is described below:

For data collection, Institutional Survey review committee reviews and approves the survey. Then the survey is designed on the Qualtrics (online survey software). A list of population is taken from Data warehouse, then the link of the survey is sent to all in the list. Several reminders with incentive is done to increase response rate. Then the online survey is closed.

For Data Analysis in the old method, The data was download from the Qualtrics (an online software used by QU) and it was cleaned to be ready for Analysis. Response rate is calculated which has been found to average 20% . Then we estimate the satisfaction survey university wide by calculating the weighted mean for all items. Moreover, the estimated satisfaction survey is calculated also for gender, nationality, colleges, and programs from those who respond to the survey.

This method has been criticized by many owners as a result of low response rate especially for some programs which leads to high nonresponse bias. So that this research focuses on finding a new method to deal with non-response bias by finding a good estimator for Student Satisfaction at QU

2.2 New Methods Used to Deal with Non-Response Bias.

2.2.1 Phase 1: Census

The survey questionnaire is sent to all population members, say$N$. Assume that N_r members respond to the questionnaires after several reminders. Thus, the population is partitioned into two subpopulations of N_r responses and $N_n = N - N_r$ nonresponses.

We therefore have all information about response subpopulation and we should try to get some information about the nonresponse subpopulation. We assume that the entire population can also be partitioned in to L strata and we would like to estimate the mean population for these L strata on top of estimating the mean population.

2.2.2 Phase 2: Inverse Sampling for Post-Stratified Design

In order to collect some information from the nonresponse subpopulation, we change the data collection method to a more intensive method (e.g. from email to phone) to get response from the nonresponse subpopulation in the phase one. A Simple Random Sample (SRS) of size n'_n is taken from the nonresponse subpopulation. We now post-stratify the sample into L strata and let n'_{nri} be the responses from stratum i, for $i = 1, \cdots L$. Let n'_{nni}be nonresponses from stratum i. We therefore have $n'_n = \sum_{i=1}^{L}(n'_{nri} + n'_{nni})$. Let $n_{nr} = \sum_{i=1}^{L} n'_{nri}$ be the number of responses from the nonresponses in the first phase. At this stage, we assume that the response rate at this phase is reasonably high but not perfect such that the nonresponse bias is ignorable. This can be achieved by more follow ups. Since n'_{nri} is random and can be very small such that the sample size creates problem to analysis the data at each stratum level we keep sampling from those strata for which n'_{nri} is smaller than 10% of N_{ni}, the population size for stratum i for nonresponse in the first phase, until we reach the triggered size of response $n_{nri} = 0.1N_{ni}$.

The first objective is to derive an estimator for the mean population, say μ which can be written as,

$$\mu = \frac{N_r\mu_r + N_n\mu_n}{N},$$

where μ_r will be known after survey and we should estimate μ_n.

We now introduce the notations used in this chapter:

N : Number of population size;

N: Population size; Nr: Responding sub-population size;

N_n: non- respondents sup-population;

N_{ni} :the i^{th} stratum size in the non-response sup-population;

L: Number of strata (Academic Programs)

n'_{sr} : Size of simple random sample selected from the non-response sup-population N_n .

n'_{ni} : sample size in i^{th} stratum collected by SRS; $i = 1,2,\ldots,L$

n_i : The required sample size in i^{th} stratum collected by follow-up inverse sampling.

n'_{nr}: Initial sample size selected by SRS, n^+ “the predetermined minimum sample size from each program that we call it trigger.

n_{nr}: The effective sample size size from all strata in non-response sup-population under TPIS

$n_r = Nr$: Number of respondents from the first stage

y_{rij} : The overall satisfaction of j^{th} *unit* in i^{th} stratum in the responding subpopulation (N_r)

y_{nij} : The overall satisfaction of j^{th} unit in i^{th} stratum in the nonresponding sup-population (N_n)

μ_r : Respondents sup population mean ; $\hat{\mu}_n$: The estimated mean for non-respondents sup-population

y_{ij} : Value obtained for the i^{th} unit in the i^{th}startum; $\bar{y}_i$: The mean of i^{th} stratum ;

μ : Population mean ; σ : Population variance

σ_1: Variance of respondents population; σ_2 : Variance of Non-respondents population

S_i^2 : Variance of i^{th} stratum; W_i : Weight of the i^{th}stratum

According post-stratify theorem, an unbiased estimator is

$$\hat{\mu}_n = \sum_{i=1}^{L} \frac{N_{ni}}{N_n} \hat{\mu}_{ni},$$

where, $\hat{\mu}_{ni}$ is an unbiased estimator of μ_{ni} the population mean for stratum i of the first phase (Appendix A).

For those strata that sample sizes are greater than or equal to $0.1N_{ni}$, we do not take extra samples so that $n_{nri} = n'_{nri}$ and with assumption of the nonresponse being ignorable an approximate unbiased estimator of μ_{ni} can be written as:

$$\hat{\mu}_{ni} = \bar{y}_{nri} = \sum_{j=1}^{n'_{nri}} \frac{y_{nij}}{n'_{nri}}$$

where, y_{nij} is the variable of interest for unit j in stratum i for nonresponse subpopulation in the first phase. For those strata that response sizes are smaller than $0.1N_{ni}$, we keep sample to reach the triggered response in the second phase size $n_{nri} = 0.1N_{ni}$.

From (Salehi and Seber; 2004) an unbiased estimator of μ_{in} is

$$\hat{\mu}_{ni} = \hat{p}_i \bar{y}_{nri} + (1 - \hat{p}_i)\bar{y}_{nni} \text{ (1)}$$

where $\hat{p}_i = n_{nri}/n^*_{ni}$, where n^*_{ni} is the total sample size from stratum i until we get n_{nri} responses, and $\bar{y}_{nri}$and $\bar{y}_{nni}$ the sample means of responses and nonresponses from the second phase respectively. Assuming the second phase nonresponse is ignorable, which can be achieved by more follow ups , $\hat{p}_i \approx 1$ and an approximate unbiased estimator for these strata are also the sample mean.

$$\hat{\mu}_{ni} = \bar{y}_{nri} = \sum_{j=1}^{n_{nri}} \frac{y_{nij}}{n_{nri}}$$

Therefore,

$$\hat{\mu}_n = \sum_{i=1}^{L} \frac{N_{ni}}{N_n} \bar{y}_{nri}$$

Consequently, an approximate unbiased estimator is

$$\hat{\mu} = \frac{N_r}{N}\mu_r + \frac{N_n}{N}\hat{\mu}_n \qquad (2.1)$$

Its approximate variance is,

$$Var(\hat{\mu}) = Var\left(\frac{N_r}{N}\mu_r + \frac{N_n}{N}\hat{\mu}_n\right) = Var\left(\frac{N_n}{N}\hat{\mu}_n\right) \quad = \left(\frac{N_n}{N}\right)^2 Var\,(\hat{\mu}_n) \qquad (2.2)$$

To find $Var(\hat{\mu}_n), we$ use the terminology of Sarjinder Singh (2003), for the post stratifications used in phase 2 [See appendix (B)] who proved that:

$$Var(\hat{\mu}_n) = \left(\frac{N_n - n_{nr}}{n_{nr} * N_n}\right) \sum_{i=1}^{L}\left(\frac{N_{ni}}{N_n}\right)\sigma_{ni}^2$$

$$+ \frac{1}{n_{nr}^2}\left(\frac{N_n - n_{nr}}{N_n - 1}\right)\sum_{i=1}^{L}\left(\frac{N_n - N_{ni}}{N_n}\right)\sigma_{ni}^2 \qquad (2.3)$$

On substituting (2.3) in (2.2),

$$Var\,(\hat{\mu}) = \left(\frac{N_n}{N}\right)^2 \left[\left(\frac{N_n - n_{nr}}{n_{nr} * N_n}\right) \quad \sum_{i=1}^{L}\left(\frac{N_{ni}}{N_n}\right)\sigma_{ni}^2 + \frac{1}{n_{nr}^2}\left(\frac{N_n - n_{nr}}{N_n - 1}\right)\sum_{i=1}^{L}\left(\frac{N_n - N_{ni}}{N_n}\right)\sigma_{ni}^2\right]$$

An approximate variance estimator is:

$$\widehat{Var}(\hat{\mu}) = \left(\frac{N_n}{N}\right)^2 \left[\left(\frac{N_n - n_{nr}}{n_{nr} * N_n}\right) \sum_{i=1}^{L}\left(\frac{N_{ni}}{N_n}\right)s_{ni}^2 + \frac{1}{n_{nr}^2}\left(\frac{N_n - n_{nr}}{N_n - 1}\right)\sum_{i=1}^{L}\left(\frac{N_n - N_{ni}}{N_n}\right)s_{ni}^2\right] \quad (2.2)$$

Where

$$s_{ni}^2 = \left(\frac{1}{n_{nri-1}}\right)\sum_{i=1}^{L}(\,y_{nij} - \bar{y}_{nri})^2$$

A pictorial representation of two-phase post-stratified inverse sampling scheme is given in Figure 2.1.

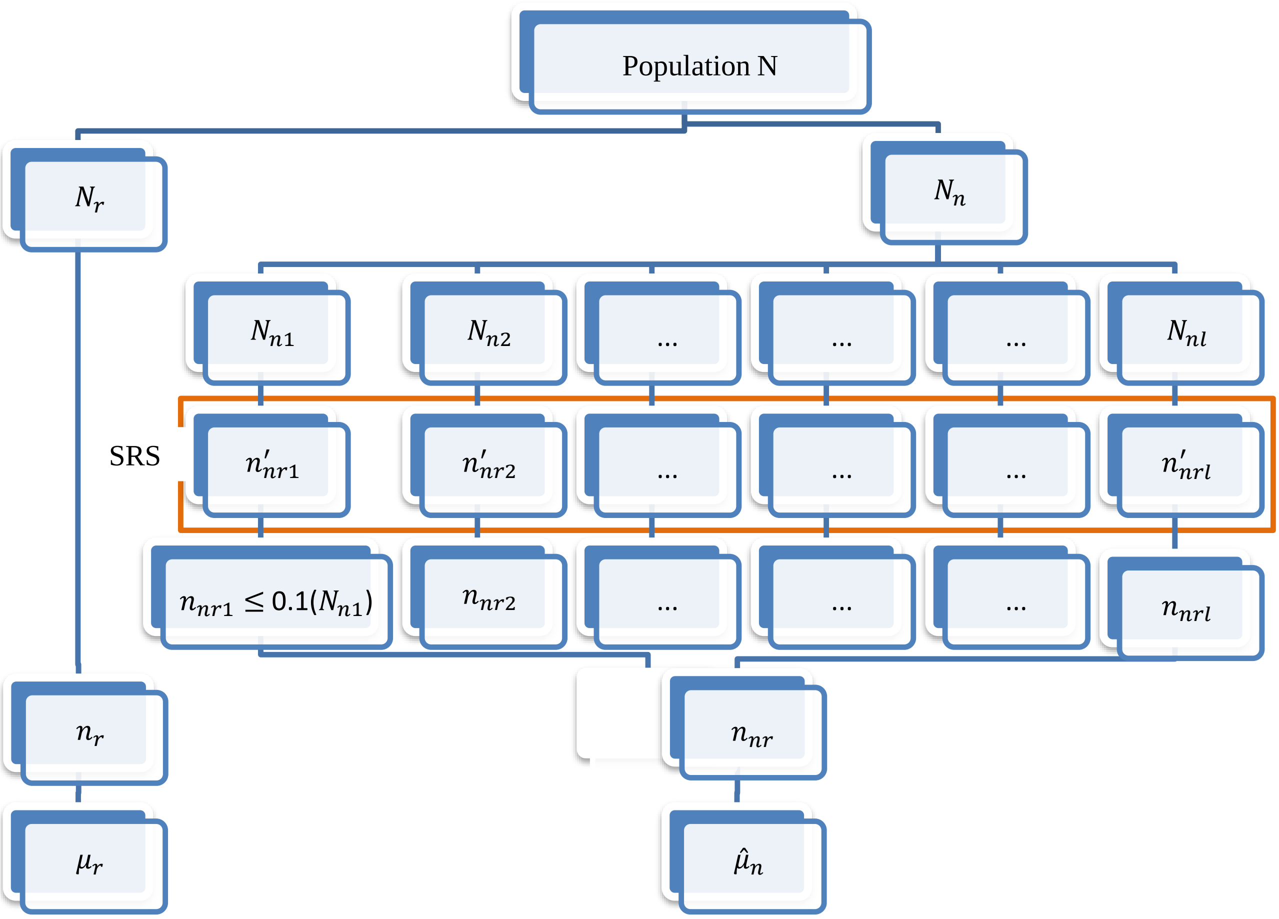

Figure 2.1 Two-phase post -stratified inverse sampling design

CHAPTER 3: CASE STUDY

This chapter is dedicated to a case study to generate the population data that will be used for simulation studies in Chapter 4; in addition , the data collected from case study is also used for evaluating the new design(TPIS) using the real data.

3.1 Estimation of Student Satisfaction by the New Design Using Case Study

3.1.1 Collecting data using Student Satisfaction Survey

The first (and possibly most important) step in reducing non-response bias is to create a properly designed survey. Whether it is online or by phone, the design of the survey can have a large impact on whether a respondent chooses to participate in the survey, and to what extent they complete the survey (Fuchs M, Bossert D, Stukowski S. ,2013). In this research, the researcher collects the data using the Student Satisfaction Survey for 2017 academic year which is used by Institutional Research and Analytics Department at Office of Chief Strategy & Development- Qatar University . This survey aims to estimate the students' satisfaction on services offered by Qatar University.

For each item in the survey, the students were asked to response a 4-point Likert-type scale with the following weighs: (1) strongly dissatisfied, (2) somewhat dissatisfied, (3) somewhat satisfied, (4) strongly satisfied. The survey validity and reliability are tested, achieved and presented in Appendix (C). Having ensured that the survey is valid and reliable, the researcher started collecting the data by applying the new sampling Design [Two-Phase Post-stratified inverse sampling (TPIS)] explained in chapter two. In the following section 3.2 the researcher explained this new sampling design.

3.2 Applying the New Design: (TPIS)

3.2.1 Phase 1: Collecting Data from Census (Old method)

The Student Satisfaction Survey was sent to all student (population $N = 16979$). Some students responded to the online survey after several reminders (five reminders), the number of students responded to the survey is ($N_r = 3278$). The response rate from this stage is 19%, which is approximately similar to the previous years. Figure (3.1) represents the cumulative response rate based on number of reminders and time. As shown in the figure (3.1) sending a reminder has been stopped at the fifth reminder because at the last reminders the changes in response rate is very small.

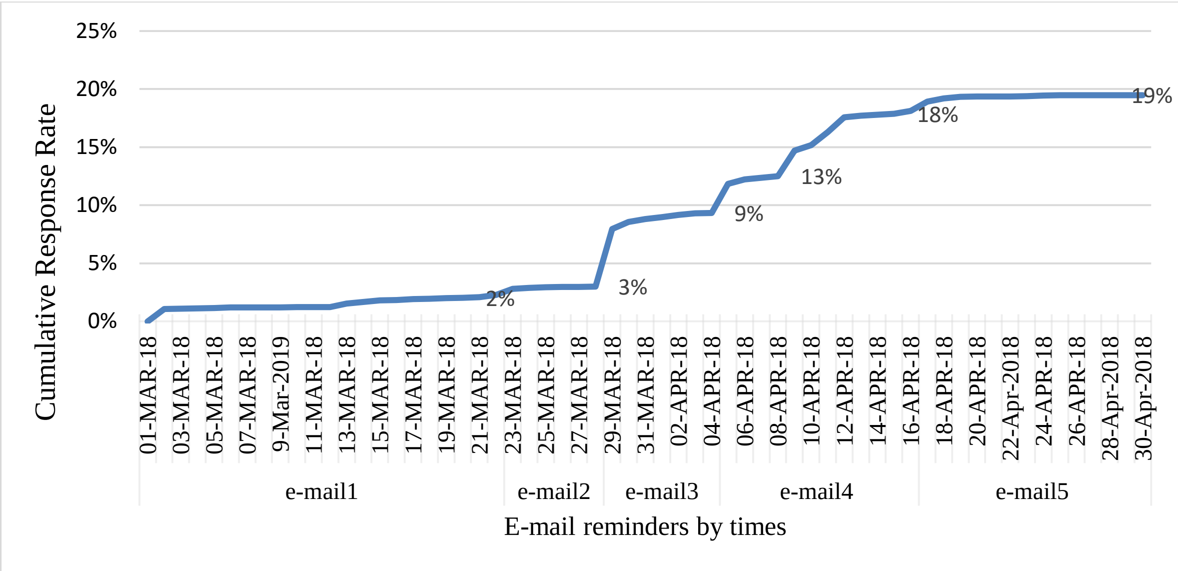

Figure 3.1 Cumulative response rate based on number of reminders and time

This step is similar to the old method used in the previous years to collect data through survey. Now, when the researcher stopped collecting data through e-mail, the population is partitioned into two subpopulations of $N_r = 3278$ responses and $N_n = N - N_r$ = 13701 nonresponses. The researcher therefore has all information about response subpopulation Nr and an attempt was made to get some information about the nonresponse subpopulation of size Nn. The entire population can also be partitioned into programs as strata (L= 51) and we would like to estimate the mean population for these L programs on top of estimating the mean population.

3.2.2 Phase 2: Post-stratified Inverse Sampling (PIS)

In order to collect some information from the nonresponse subpopulation $(Nn), we$ change the data collection method to a more intensive method (e.g. from email to phone) to get response from the nonresponse subpopulation in the phase one. A Simple Random Sample (SRS) of size $n'_n = 1000$ is taken from the nonresponse subpopulation, and then we send the survey for them by e-mail followed by phone calling including several incentives. We now post-stratify the respondents from the sample into L program= 51 and let n'_{nri} be the responses from program $i,$ for $i = 1, \cdots 51$. Let n'_{nni}be nonresponses from program i. We therefore have $n'_n = \sum_{i=1}^{L}(n'_{nri}$ + $n'_{nni})$. Let $n_{nr} = \sum_{i=1}^{L} n'_{nri}$ be the number of responses from the nonresponses in the first phase.

At this stage, the response rate a proximately 50% (492/ 1000) which is reasonably high compared to the first stage (19%) but not perfect such that the nonresponse bias is ignorable. This can be achieved by more follow-ups. Since n'_{nri} is random and can be very small such that the sample size creates problem to analysis the data at each program level we keep sampling from those programs for which n'_{nri} is smaller than 7 of N_{ni}, the population size for program i for nonresponse in the first phase, until we reach the triggered size of response $n_{nri} = 7$.

To estimate the mean of student satisfaction for the population (μ), we calculate the μ_r from first phase, then we calculate the estimated men $\hat{\mu}_n$ for non-response students. According post-stratify theorem, an unbiased estimator is

$$\hat{\mu}_n = \sum_{i=1}^{L} \frac{N_{ni}}{N_n} \hat{\mu}_{ni},$$

Using the equation (2.1),

$$\mu = (N_r\mu_r + N_n \frac{\mu_n}{N}$$

where μ_r will be known after survey and we should estimate μ_n.

For those programs that sample sizes are greater than or equal to 7, we do not take extra samples so that $n_{nri} = n'_{nri}$ and with assumption of the nonresponse being ignorable an approximate unbiased estimator of μ_{ni} can be written as

$$\hat{\mu}_{ni} = \bar{y}_{nri} = \sum_{j=1}^{n'_{nri}} \frac{y_{nij}}{n'_{nri}}$$

where, y_{nij} is the satisfaction rate of student j in program i for nonresponse subpopulation in the first phase. For those program that response sizes are smaller than 7, we keep sample to reach the triggered response in the second phase size $n_{nri} = 7$.

an unbiased estimator of μ_{ni} is:

$$\hat{\mu}_{ni} = \bar{y}_{nri} = \sum_{j=1}^{n_{nri}} \frac{y_{nij}}{n_{nri}}$$

Therefore,

$$\hat{\mu}_n = \sum\nolimits_{i=1}^{L} \frac{N_{ni}}{N_n} \bar{y}_{nri}$$

Consequently, an approximate unbiased estimator is

$$\hat{\mu} = \tfrac{N_r}{N}\mu_r + \tfrac{N_n}{N}\hat{\mu}_n.$$

To calculate the approximate variance estimator for population $\hat{V}ar(\hat{\mu})$, we using equation (2.2)

$$\widehat{Var}(\hat{\mu}) = \left(\frac{N_n}{N}\right)^2 \left[\left(\frac{N_n - n_{nr}}{n_{nr} * N_n}\right) \sum_{i=1}^{L} \left(\frac{N_{ni}}{N_n}\right) s_{ni}^2 + \frac{1}{n_{nr}^2}\left(\frac{N_n - n_{nr}}{N_n - 1}\right) \sum_{i=1}^{L} \left(\frac{N_n - N_{ni}}{N_n}\right) s_{ni}^2\right]$$

where

$$s_{ni}^2 = \left(\frac{1}{n_{nri}}\right)\sum_{i=1}^{L}(y_{nij} - \bar{y}_{nri})^2$$

The results of this case study will be presented in the next section.

3.3 Results of CAS Study

3.3.1 Estimating the mean of Student Satisfaction using TPIS design

The survey was sent to all population, then $N = 16979$

From phase 1: we found that, , $N_r = 3278$, and $\mu_r = \sum_1^{3278} y_i$ = 3.43300

From phase 2: we found that:

$N_r = 3278,\ N_n = 13701, n'_{n(srs)} = 1000, n'_{nr} = 492\ , n'_{nr} = 492\ , n_{nr} = 560\ ,$

To estimate the $\hat{\mu}$ *and* $\hat{V}(\hat{\mu})$, we recorded and organized the data needed and collected from phase 2 in table 3.1 below

Table 3.1 Data collected from nonresponse subpopulation by program.

Program	N_{ni}	n'_{nri}	n'_{nni}	n_{nri}	$\bar{y}_{nri}$	$w * \bar{y}$	s_{nri}^2	$w * s_{nri}^2$	$\frac{Nn - Nni}{Nn} * s_{nri}^2$
1	395	14	381	14	3.5714	0.1030	0.2637	0.0076	0.2561
2	83	12	71	12	3.7500	0.0227	0.3864	0.0023	0.3840
3	115	9	106	9	3.3333	0.0280	0.5000	0.0042	0.4958
4	78	5	73	12	3.8333	0.0218	0.1515	0.0009	0.1507
5	87	13	74	13	3.5385	0.0225	0.2692	0.0017	0.2675
6	253	11	242	11	3.6364	0.0671	0.2545	0.0047	0.2498
7	109	14	95	14	3.6429	0.0290	0.2473	0.0020	0.2453
8	184	6	178	11	3.0000	0.0403	0.6000	0.0081	0.5919
9	202	12	190	12	3.0833	0.0455	0.6288	0.0093	0.6195
10	142	14	128	14	3.0000	0.0311	0.6154	0.0064	0.6090
11	177	6	171	9	3.2222	0.0416	0.4444	0.0057	0.4387
12	95	8	87	8	3.5000	0.0243	0.2857	0.0020	0.2837
13	18	5	13	10	3.5000	0.0046	0.5000	0.0007	0.4993
14	269	15	254	15	3.1333	0.0615	0.4095	0.0080	0.4015
15	213	9	204	9	3.5556	0.0553	0.5278	0.0082	0.5196
16	187	16	171	16	3.2500	0.0444	0.3333	0.0045	0.3288
17	106	8	98	8	3.1250	0.0242	0.4107	0.0032	0.4075
18	228	15	213	15	3.4667	0.0577	0.2667	0.0044	0.2622
19	1521	11	1510	11	3.6364	0.4037	0.4545	0.0505	0.4041

Program	N_{ni}	n'_{nri}	n'_{nni}	n_{nri}	$\bar{y}_{nri}$	$w * \bar{y}$	s^2_{nri}	$w * s^2_{nri}$	$\frac{Nn - Nni}{Nn} * s^2_{nri}$
20	1821	12	1809	12	3.5833	0.4763	0.4470	0.0594	0.3876
21	614	10	604	10	3.5000	0.1568	0.2778	0.0124	0.2653
22	55	5	50	11	3.2727	0.0131	0.2182	0.0009	0.2173
23	85	6	79	7	3.1429	0.0195	0.4762	0.0030	0.4732
24	35	4	31	12	3.3333	0.0085	0.2424	0.0006	0.2418
25	228	12	216	12	3.6667	0.0610	0.2424	0.0040	0.2384
26	354	14	340	14	3.5000	0.0904	0.4231	0.0109	0.4121
27	68	4	64	7	3.4286	0.0170	0.2857	0.0014	0.2843
28	134	12	122	12	3.5833	0.0350	0.4470	0.0044	0.4426
29	89	8	81	8	3.2500	0.0211	0.5000	0.0032	0.4968
30	237	9	228	9	3.0000	0.0519	0.5000	0.0086	0.4914
31	356	10	346	10	3.4000	0.0883	0.9333	0.0243	0.9091
32	140	11	129	11	3.6364	0.0372	0.2545	0.0026	0.2519
33	1115	11	1104	11	3.5455	0.2885	0.4727	0.0385	0.4343
34	379	15	364	15	3.7333	0.1033	0.3524	0.0097	0.3426
35	52	5	47	10	3.5000	0.0133	0.2778	0.0011	0.2767
36	336	13	323	13	3.1538	0.0773	0.9744	0.0239	0.9505
37	467	11	456	11	3.4545	0.1177	0.4727	0.0161	0.4566
38	221	11	210	11	3.2727	0.0528	0.2182	0.0035	0.2147
39	103	9	94	9	3.0000	0.0226	0.5000	0.0038	0.4962
40	73	5	68	11	3.7273	0.0199	0.2182	0.0012	0.2170
41	383	9	374	9	3.3333	0.0932	0.7500	0.0210	0.7290
42	430	12	418	12	3.5000	0.1098	0.2727	0.0086	0.2642
43	167	11	156	11	3.2727	0.0399	0.6182	0.0075	0.6106
44	73	5	68	10	3.5000	0.0186	0.2778	0.0015	0.2763
45	136	9	127	9	3.5556	0.0353	0.2778	0.0028	0.2750
46	337	11	326	11	3.6364	0.0894	0.2545	0.0063	0.2483
47	211	10	201	10	3.4000	0.0524	0.2667	0.0041	0.2626
48	331	11	320	11	3.6364	0.0879	0.4545	0.0110	0.4436
49	61	4	57	11	3.5455	0.0158	0.2727	0.0012	0.2715
50	85	6	79	8	3.2500	0.0202	0.7857	0.0049	0.7808
51	63	4	59	9	3.3333	0.0153	0.5000	0.0023	0.4977
Total	13701	492	13209	560	3.4776		0.4171	0.4389	20.575

$$\hat{\mu}_n = \sum_{i=1}^{L} \frac{N_{ni}}{N_n} \bar{y}_{nri} = 3.47757$$

$$\hat{\mu} = \frac{N_r}{N}\mu_r + \frac{N_n}{N}\hat{\mu}_n = \frac{3278}{16979} * 3.43300 + \frac{13701}{16979} * 3.47757 = 3.468968$$

$$95\% \text{ CI} = \hat{\mu} \pm z * \sqrt{\left(\frac{Var}{n}\right)} = 3.468968 \pm 1.96 * \sqrt{\left(\frac{0.4389}{560}\right)} = 3.468968 \pm 0.054871$$

$$3.523839 \geq \hat{\mu} \leq 3.414097$$

We found that the $\widehat{\mu_n} = 3.447757$ is greater than the $\mu_r = 3.43300$, which mean that those who did not respond in the first phase has higher satisfaction than those who responded in the first phase. Our interpretation to this difference, those who response in first phase may have some issues about the services provided by Qatar University and they need to rise their sound to the QU leaders.

3.3.2 Estimating of the Overall variance of Student Satisfaction using invers sampling for post stratified design

To estimate the variance of student satisfaction based on the new design (TPIS), we used the data collected from Phase 1 and phase 2 and the formula 2.2 which is:

$$\widehat{Var}(\hat{\mu}) = \left(\frac{N_n}{N}\right)^2 \left[\left(\frac{N_n - n_{nr}}{n_{nr} * N_n}\right) \sum_{i=1}^{L} \left(\frac{N_{ni}}{N_n}\right) s_{ni}^2 + \frac{1}{n_{nr}^2}\left(\frac{N_n - n_{nr}}{N_n - 1}\right) \sum_{i=1}^{L} \left(\frac{N_n - N_{ni}}{N_n}\right) s_{ni}^2\right]$$

All components of the variance equation above are calculated based on the data collected in Phase one and Phase 2 which presented in the previous section (Table 3.1). Summary results for the variance equation components derived from table 3.1 are presented in table 3.2 .

Table 3.2 Summary results for the variance equation components derived from case study.

$N=$	16979	$(N_n - n_{nr})/(n_{nr} * N_n) =$	0.0017
$N_n =$	13701	$(N_n - n_{nr})/(N_n - 1) =$	0.9592
nnr	560	$sum(\left(\frac{N_{ni}}{N}\right) S_{ni}^2 =$	0.43890
$\left(\frac{1}{n_{nr}}\right)^2$	0.32000	$\frac{N_n - N_{ni}}{N_n}\ S_{ni}^2$	20.57524
$\left(\frac{N_n}{N}\right)^2$	0.65115	$\widehat{Var}(\hat{\mu})$	0.000530

By substituting the values in table 3.2 in variance equation above,

We found that : $\widehat{Var}(\hat{\mu}) = 0.000530$. This value is very small compared to the variance calculated from Phase one alone (old methods) $Var\ (\mu_r) = 0.444$.

This mean that the new design is beater for estimating the student satisfaction compared to the old methods.

3.3.3. Evaluating the Bias of Non-respondents on Student Satisfaction

In a survey research, there are several ways to evaluate the non-response bias as mentioned in the literature review in chapter 1(section 1.4.2).One way is following up on non-respondents which is considered an excellent way to reduce the non-response. Because late-respondents, or those that respond after several attempts, are theorized to have some similarities with non-respondents. One approach to evaluate whether there is non-response bias or not, the population may be divided into two sup-populations (Response and non-response), then the comparison between the estimated variable from response and the estimated variable from nonresponse subpopulations is made. If there is a significant difference between two estimated of the variable, this means that, the nonresponse bias is existing, otherwise not.

In our case study, we compare the mean of student satisfaction for those who response in the first phase (response sup-population) with the mean of student satisfaction estimated in Phase two (nonresponse subpopulation). In other words, we try to evaluate the non-response bias by testing the following null hypothesis: $H_0 : \mu_r - \mu_n = 0$, The alternative hypothesis is $H_1 : \mu_r - \mu_n \neq 0$

Z test was used to test the null hypothesis. Using the following equation:

$$Z = \frac{\hat{\mu}_{nr-}\mu_r}{\sqrt{V\hat{}ar(\hat{\mu})}} \qquad (3.1)$$

By substituting the values in table 3.2 in equation 3.1

$$Z = \frac{\hat{\mu}_{nr} - \mu_r -}{\sqrt{V\hat{}ar(\hat{\mu})}} = \frac{3.43300 - 3.477573}{\sqrt{0.0005}} = \quad 1.940237209$$

The value of Z shows that the estimated mean of overall satisfaction was higher for the nonresponse groups than the response groups but the difference was not statistically significant at level of $\alpha = 0.05$ but significant at level of $\alpha = 0.10$.

The non-response bias of estimator in terms of programs are also evaluated by comparing the satisfaction mean of students in each program who responded in the first phase and those who responded in the second phase. The results show that there is a significant difference between the two means of respondents in the first Phase and non-respondents in the second for 13 programs out of 51 (25%). The significant differences mean that our estimation of the student satisfaction is biased for 25% of program. This means that we reject the null hypothesis: for 13 programs and accept the alternative hypothesis

Table 3.3 presents the results of programs that have z values that have significant differences between two means of respondents and non-respondents and the z test results for all programs are posted in the appendix C.

Table 1.3 Evaluating non-response bias by program using z-test

$Programs$	N_{ri}	n_{nri}	(μ_{ri})	μ_{nri}	$Va\hat{r}(\hat{\mu})$	z	$p-value$
Computer Science	51	14	3.4510	3.0000	0.0048	6.5401	0.0000
General Pharmacy	25	12	3.5600	3.3333	0.0014	6.1355	0.0000
General Health Sciences	25	11	3.5200	3.2727	0.0021	5.3963	0.0000
Social Work	19	10	3.7368	3.4000	0.0053	4.6144	0.0000
Health Sciences Foundation	20	7	3.8000	3.4286	0.0084	4.0419	0.0001
Industrial and Systems Eng.	45	9	3.3556	3.0000	0.0114	3.3294	0.0009
Computer Engineering	87	12	3.3448	3.0833	0.0067	3.1945	0.0014
Dawa	35	9	3.5143	3.2222	0.0096	2.9741	0.0029
Psychology	50	11	3.5000	3.2727	0.0083	2.4970	0.0125
Civil Engineering	136	11	3.1765	3.0000	0.0061	2.2575	0.0240
Education Foundation	8	10	3.6250	3.5000	0.0033	2.1779	0.0294
English Literature	64	16	3.3438	3.2500	0.0021	2.0234	0.0430
General Education	148	10	3.6554	3.5000	0.0062	1.9758	0.0482

Since the non-response bias exists for several programs, this means that the old method explained in chapter two needs to be changed. This research attempts to find a solution to reduce the bias existing in the old method (only collecting data by e-mail as first phase) by creating a new design called TPIS and we consider it as an important contribution of this research. So that, we need to evaluate the bias estimators based on this new design by generating data and simulations. Section 3.4 will focus on the generating population by linking survey respondents to non-respondents while chapter four will focus on the simulation to evaluate the bias estimator for mean and variance. In addition, it will evaluate the effectiveness of this design compered to SRS. Chapter 4 will focus on this simulation studies.

3.4 Generating the Non –Response Population by linking Survey Respondents to Non-respondent.

Based on the sampling design explained in chapter 2, we collect data from Non-respondent sup population, by SRS followed by inverse sampling for post stratification as explained in section 3.2.

From the students list, we have $N = 16979,$ from case study $N_r = 3278, N_{nri} = 560$ then we need to generate:

$$N_{nn} = N - (N_r + N_{nr}) = 13142.$$

This number should be distributed by program, since we have the list of all students. The N_i is known for each program where N_i number of student in program$i$. From case study we know the N_{ri} *and* N_{nri} those who respond from program i in phase one and two. Then we determine the N_{nni} those who did not respond to the survey in phase one or two.

$$N_{nni} = N_i - (N_{ri} + N_{nri})$$

Now, to generate the $N_{nn}i$ non response with satisfaction, we use the data collected from student satisfaction survey in phase two for nonresponse subpopulation, we estimate the probability of each survey response [strongly satisfied (4), satisfied (3), dissatisfied (2) and strongly dissatisfied (1)]. Then the distribution of data is consider as a Multinomial Distribution by considering an experiment with n independent traiels. Each traile can result in any of r possible outcomes (1,2,3,4).

p_i denotes the propability of outcome $i, \sum_{i=1}^{4} p_i = 1$

$n_i denotes\ the\ number\ of\ trials\ resulting\ in\ outcomes\ i, \sum_{i=1}^{4} n_i = n$

$P = (p_1, \dots, p_r), n = (n_1, \dots, n_r)$ the maximum likelihood estimator for p_i is

$$l(p) = \prod_{i=1}^{4} p_i^{n_i}$$

By taking the first derivation and equaling it to zero , the ML is $\widehat{p_\iota} = \frac{n_i}{n}$

We calculate the $\widehat{p_\iota}$ for each program from data collected in the second phase. Table 3.4 shows the number of non-responses needed from each program with probability of responses for (1, 2,-3, 4). The overall non-response to the survey in phase one and phase two N_{nn} = 13141 are generated using R software. [R code is attached in the Appendix (D)]. The generated non- response was added to the non-respondents from calling ($n_{nn} = 560$). The total non-respondents $N_n = 13701$.

The total mean for all population is calculated. $\mu = 3.479800$

The estimated mean from case study = 3.468968

The relative bias of estimator mean (RBE) = $\frac{3.479800-3.468968}{3.479800}$ = 0.3%

The RBE is very small and approximately zero which mean that the estimated mean from case study is not bias. Generated population will be used for simulation study in chapter 4.

CHAPTER 4: SIMULATION STUDY

4.1 Simulation Using New Design (TPIS)

In this section, we evaluate the TPIS design by conducting simulation studies based on generated data explained in chapter three. The TPIS design is evaluated in terms of properties and efficiencies of the estimators$\hat{\mu}$; $\hat{V}ar(\hat{\mu})$, which were calculated, based on 10,000 samples (n'_{nn}) taken according to post-stratified inverse sampling design from the generated populations. The different samples taken are;

$n'_{nn} = 452, 550, 600, 700, 800, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700$)

With fixed predefined numbers (trigger) $n^{+} = 7$ and $n^{+} = 10$. Moreover, we tried to evaluate the new design when we fixed a sample size with different triggers $n^{+} = 5,7,10, 12,14, 16,17$.

More details about TPIS designed evaluation in terms of properties (bias) and efficiencies of the estimators $\hat{\mu}$, $\hat{V}ar(\hat{\mu})$ are explained in the following sections. Simulation R- code is attached in Appendix (E).

4.1.1 Evaluating the Bias of Estimator for Population Mean

Simulation study is used to evaluate the bias of estimator for population mean. The Mont carol mean and variance of estimator are computed as follows;

$$\hat{\mu} = \frac{1}{10000} \sum_{R=1}^{10000} \hat{\mu}_R$$

The relative bias of estimator (RBE) for population mean is calculated using the following formula:

$$RBE = (\hat{\mu} - \mu)/\ \hat{\mu} \ ,$$

where $\mu = 3.479800$ (for all population)

Table (4.1) shows the relative bias estimator of mean for different sample size n'_{nn} at trigger (n^{+} = 7; n^{+} = 10). Graph 4.1 illustrates the relationship between sample size and relative bias estimator of μ at the trigger n^{+} = 7, and 10, w $n_{nn} = n'_{nn} + n^{+}$

Table 4.1 RBE for population mean based on different sample size n_nn selected based on TPIS design

R=10000	Trigger (n+=7)			Trigger (n+=10)		
n'_{nn}	n_{nn}	$\hat{\mu}$	RBE For $\hat{\mu}$	n_{nn}	$\hat{\mu}$	RBE For $\hat{\mu}$
492	597	3.479687	-0.0032%	695	3.479725	-0.002%
550	644	3.479782	-0.0005%	735	3.479499	-0.009%
600	685	3.479583	-0.0062%	771	3.479889	0.003%
650	727	3.479996	0.0056%	808	3.479667	-0.004%
700	770	3.479362	-0.0126%	847	3.479781	-0.001%
800	859	3.479968	0.0048%	927	3.479659	0.013%
900	949	3.47937	-0.0124%	1010	3.480249	0.001%
1000	1041	3.479805	0.0001%	1097	3.479851	0.001%
1100	1135	3.479769	-0.0009%	1184	3.479851	0.001%
1200	1229	3.479773	-0.0008%	1365	3.479852	0.001%
1300	1324	3.479675	-0.0036%	1364	3.479852	0.001%
1400	1421	3.479931	0.0038%	1457	3.479878	0.002%
1500	1518	3.479809	0.0003%	1550	3.480009	0.006%
1600	1615	3.479583	-0.0062%	1643	3.479881	0.002%
1700	1713	3.479939	0.0040%	1738	3.479881	0.002%

From table (4.1) we found that the relative bias of estimator for population mean is very small (between -0.002% - 0.013%). This small percentage of RBE considered as a simulation error as we show the estimator is unbiased.

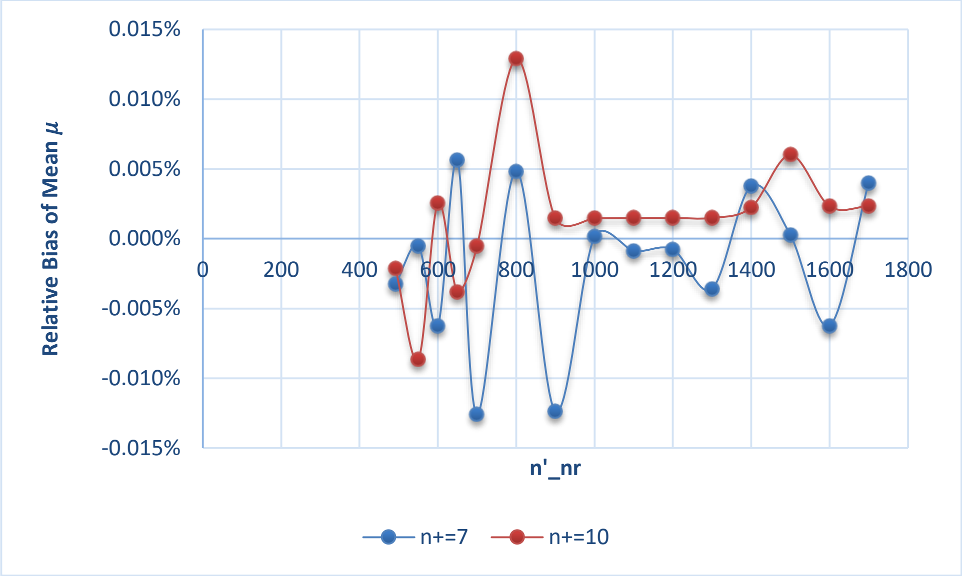

Figure 4.1 RBE for population mean based on different sample size n_{nn}

4.1.2 Evaluating the Variance of Estimators $(\hat{\mu})$

The simulation study explained in the previous section also used to evaluate the variance of estimator$\hat{V}ar\ (\hat{\mu})$. In the simulation study, an approximate variance estimator is calculated based on the following formal.

$$\widehat{Var}(\hat{\mu}) = \left(\frac{N_n}{N}\right)^2 \left[\left(\frac{N_n - n_{nr}}{n_{nr} * N_n}\right) \sum_{i=1}^{L} \left(\frac{N_{ni}}{N_n}\right) s_{ni}^2 + \frac{1}{n_{nr}^2} \left(\frac{N_n - n_{nr}}{N_n - 1}\right) \sum_{i=1}^{L} \left(\frac{N_n - N_{ni}}{N_n}\right) s_{ni}^2\right]$$

Where

$$s_{ni}^2 = \left(\frac{1}{n_{nri} - 1}\right) \sum_{i=1}^{L} (y_{nij} - \bar{y}_{nri})^2$$

Mean Square Error of the mean $MSE(\hat{\mu})$ is also calculated based on the following formula:

$$MSE(\hat{\mu}) = \frac{1}{9999} \sum_{R=1}^{10000} (\hat{\mu}_R - \mu)^2$$

Based on the results we got from the simulation, the relative bias of estimator (RBE) for population variance is calculated using the following formula:

$$RBE = \hat{V}ar(\hat{\mu}) - MSE\ (\hat{\mu})/\ \hat{V}ar(\hat{\mu})$$

Table (4.2) shows the relative bias estimator (RBE) of mean for different sample size n'_{nn} at trigger (n^+ = 7; n^+ = 10) . Figure 4.2 also demonstrate the relationship between sample size n'_{nn} and relative bias estimator of Varaince at the trigger n^+ = 7, and 10.

From table (4.2) and figures (4.2, 4.3) we found that: when the sample size increases the RBE decreases, which means that the precision of estimator increases. Moreover when the sample size is approximately > 900 with any trigger (n+ = 7 or 10), the RBE is less than 5% which means that the $MSE\ (\hat{\mu})$ is very closes to $\hat{V}ar\ (\hat{\mu})$ at $n'_{nn} \geq 900$. In other word the variance estimator is over estimating the variance of estimator which mean that the $\hat{\mu}$ has high precision to estimate mean of population especially at the sample size $n'_{nn} \geq 900$.

Table 4.2 RBE for population variance based on different sample size n'_{nn} at trigger (n+ = 7,n+=10)

	Replicates =10,000 , n+=7				Replicates =10,000 , n+=10			
n'_{nr}	n_{nr}	$E[\hat{V}ar(\hat{\mu})]$	$MSE\ (\hat{\mu}_{nr})$	RBE of $\hat{\mu}$	n_{nr}	$E[\hat{V}ar(\hat{\mu})]$	$MSE\ (\hat{\mu}_{nn})$	RBE of $\hat{\mu}$
492	597	0.000870777	0.000763246	12.3%	695	0.000870886	0.000718868	17.5%
550	644	0.000767828	0.00069763	9.1%	735	0.000768229	0.000644036	16.2%
600	685	0.000696998	0.000637049	8.6%	771	0.000696035	0.000593648	14.7%
650	727	0.000637307	0.000594523	6.7%	847	0.00058666	0.00051042	13.0%
700	770	0.000587086	0.000547659	6.7%	927	0.000637602	0.000561739	11.9%
800	859	0.000505533	0.000484624	4.1%	1010	0.000505759	0.000454254	10.2%
900	949	0.000443447	0.000424793	4.2%	1097	0.000442924	0.000419541	5.3%
1000	1041	0.000393483	0.000383481	2.5%	1184	0.000393662	0.000374292	4.9%
1100	1135	0.000356226	0.000345735	2.9%	1365	0.00035344	0.000338191	4.3%
1200	1229	0.000320471	0.00031013	3.2%	1364	0.000320064	0.000312176	2.5%
1300	1324	0.000292784	0.000287606	1.8%	1457	0.000292642	0.000283676	3.1%
1400	1421	0.000268856	0.000268398	0.2%	1550	0.000269003	0.00026324	2.1%
1500	1518	0.000248294	0.000246127	0.9%	1643	0.000248378	0.000241886	2.6%
1600	1615	0.000230607	0.000224139	2.8%	1738	0.000230456	0.000226051	1.9%
1700	1713	0.000214755	0.00021402	0.3%	1738	0.000214845	0.000213306	0.7%

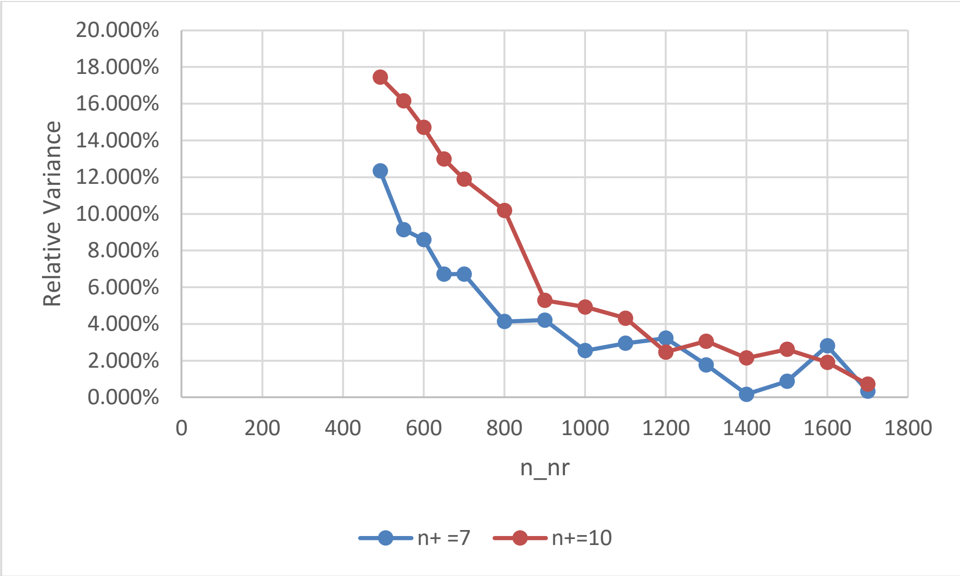

Figure 4.2 RBE for population variance based on different sample size n'_{nn}

Now let us look for the influence of the trigger n^+ when we fixed the sample size n'_{nn} at 900 at different triggers $n^+ = 5, 7,10, 12, 15, 16,17$.

Table 4.3 RBE for population variance based on different trigger (n+) at affixed sample size $n'_{nn} = 900$)

n'_{nr}	n_{nr}	$E[\hat{V}ar(\hat{\mu})]$	$MSE\ (\hat{\mu}_{nr})$	RBE of $\hat{\mu}$
5	921	0.000442933	0.00043062	2.8%
7	949	0.000443447	0.00042479	4.2%
10	1010	0.000442924	0.00041954	5.3%
12	1062	0.000428101	0.00040355	5.7%
15	1151	0.000404285	0.00037692	6.8%
16	1183	0.000441918	0.00040648	8.0%
17	1218	0.000442842	0.00038707	12.6%
18	1253	0.000428611	0.00037586	12.3%

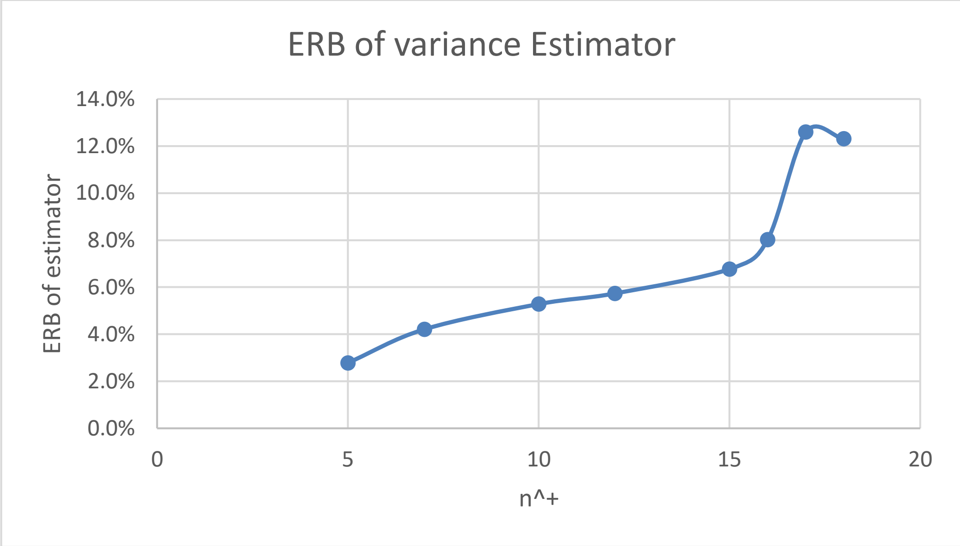

Figure 4.3 RBE for population variance based on different trigger (n+) at affixed sample size n'_{nn} =900)

The results in table 4.3 shows that the precision of estimator decreases if the trigger n^+ increases. The best precision is when the trigger $n^+ \leq 10$.

We conclude that the bias of estimator not only depend on a sample size n'_{nn} but also by the trigger n^+.

4.1.3 Evaluating the Effeminacy of New Design Compared to SRS

To estimate the gain in precision due to TPIS design or to evaluate its efficiency of this design, it is important to compare variance estimator of this design with any other design used in the survey research. In this section we will compare the $MSE\ (\hat{\mu}_{TPIS})$ under TPIS design with the $MSE\ (\hat{\mu}_{srs})$ under simple random sample (SRS) without stratification. This comparison gives us an idea about the gain in efficiency due to TPIS design.

Efficiency of new design (TPIS) is calculated using the formula:

$$Efficieancy\ of\ TPPIS = \frac{MSE(\hat{y}_{srs})}{MSE((\hat{\mu}_{TPIS}))} \quad (4.1)$$

$$MSE\ (\hat{\mu}) = \left(\frac{N_n}{N}\right)^2 * \hat{V}(\hat{\mu})\ \ (4.2\ );$$

sinse the $\hat{V}(\hat{\mu})$ is calculated by the simulation.

From the simulation, $MSE(\hat{y}_{srs})$ is calculated as below;

$$MSE(\hat{y}_{srs}) = \hat{V}ar\ (\hat{y}_{srs}) + (bias)^2\ \ (4.3)\ ,$$

Using the above equations and the known values for the below items

N_r= 3838, μ = 3.479800, $\hat{\mu}$= 3.463369

Variance= 0.444, Bias= 0.007431, Replicates= 10,000

The results of simulation are shown in tables 4.4; 4.5 and figure 4.3

Table 4.4 Efficiency of new design (TPIS) compared to SRS (n+=7)

n'_{nr}	n_{nr}	$E[\hat{V}ar(\hat{\mu})]$	$MSE(\hat{\mu}_{TPIS})$	$MSE(\hat{\mu}_{SRS})$	*Eff.*	Gain Eff.
492	597	0.00087077	0.0005216	0.000683253	131%	31%
550	644	0.00076783	0.00045994	0.000628976	137%	37%
600	685	0.000697	0.00041751	0.00058771	141%	41%
650	727	0.00063731	0.00038175	0.000550264	144%	44%
700	770	0.00058709	0.00035167	0.000516158	147%	47%
800	859	0.00050553	0.00030282	0.000456415	151%	51%
900	949	0.00044345	0.00026563	0.000407395	153%	53%
1000	1041	0.00039348	0.0002357	0.000366047	155%	55%
1100	1135	0.00035623	0.00021338	0.000330724	155%	55%
1200	1229	0.00031947	0.00019137	0.000300804	157%	57%
1300	1324	0.00029278	0.00017538	0.000274882	157%	57%
1400	1421	0.00026886	0.00016105	0.000251991	156%	56%
1500	1518	0.00024829	0.00014873	0.000232025	156%	56%
1600	1615	0.00023001	0.00013778	0.000214457	156%	56%
1700	1713	0.00021275	0.00012744	0.000198729	156%	56%

Table 4.5 Efficiency of new design (TPIS) compared to SRS (n+=10)

n'_{nr}	n_{nr}	$\hat{V}\overline{\hat{\mu}_{nr}}$	$MSE(\hat{\mu}_{TPIS})$	$MSE(\hat{\mu}_{SRS})$	Eff.	Gain Eff.
492	695	0.00087089	0.0005217	0.000578383	111%	10.9%
550	735	0.00076823	0.0004602	0.000543616	118%	18.1%
600	771	0.00069603	0.0004169	0.00051541	124%	23.6%
650	808	0.0006376	0.0003819	0.000489039	128%	28.0%
700	847	0.00058666	0.0003514	0.000463738	132%	32.0%
800	927	0.00050576	0.0003030	0.000418499	138%	38.1%
900	1010	0.00044292	0.0002653	0.000379138	143%	42.9%
1000	1097	0.00039366	0.0002358	0.000344275	146%	46.0%
1100	1184	0.00035344	0.0002117	0.000314535	149%	48.6%
1200	1255	0.000330064	0.0001977	0.000293319	148%	48.4%
1300	1364	0.00029264	0.0001753	0.000265048	151%	51.2%
1400	1457	0.00026900	0.0001611	0.00024427	152%	51.6%
1500	1550	0.00024838	0.0001488	0.000225986	152%	51.9%
1600	1643	0.00023046	0.0001380	0.000209772	152%	52.0%
1700	1738	0.00021385	0.0001281	0.000195001	152%	52.2%

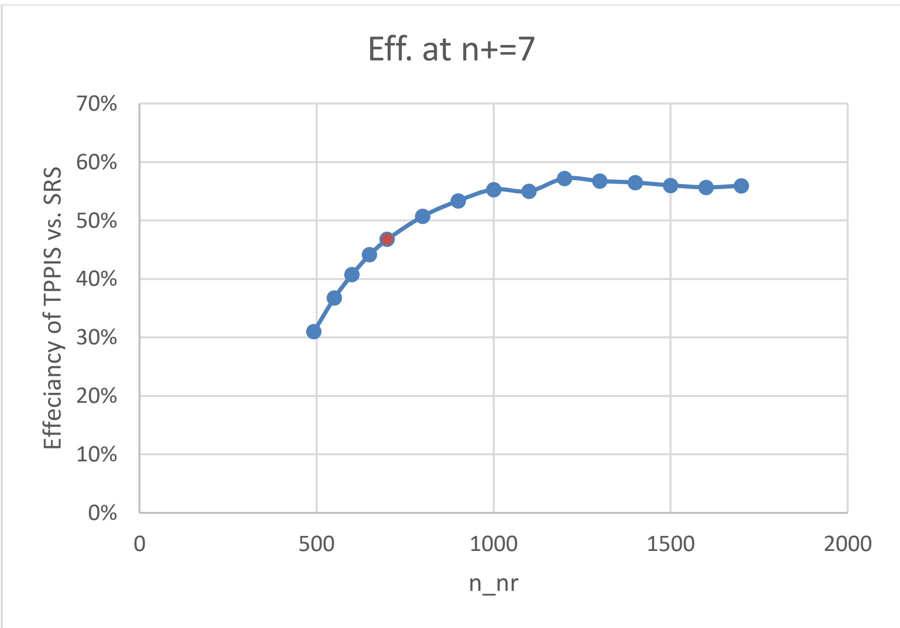

Figure 4.4 Efficiency of new design (TPIS) compared to SRS (n+=7)

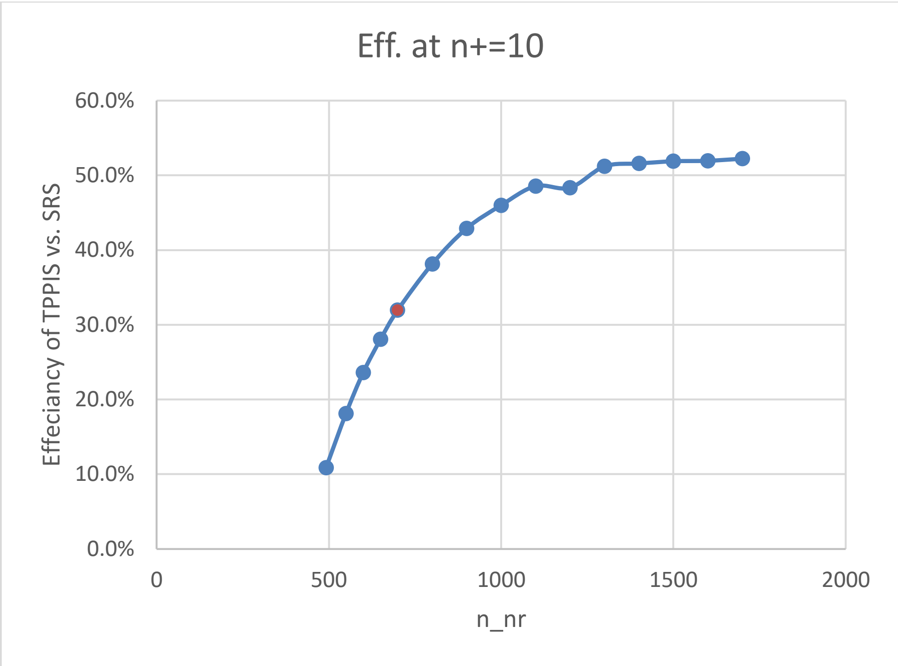

Figure 4.5 Efficiency of new design (TPIS) compared to SRS (n+=10)

From tables (4.4 & 4.5) and graphs (4.4 & 4.5), we found that efficiency of TPIS design increases to approximately 56% compared to SRS when sample size increases from $n_{nr} = 492$ to approximately 1300 with $n^+ = 7$. At $n_{nr} \geq\ 1300$; then the changes in the efficiency is still constant at 56%. From table 4.4 and Figure 4.4 we found that when we increase the trigger from $n^+ = 7$ to $n^+ = 10$ the efficiency approximately fixed at 52% for both. This means that the highest sample size that give more efficiency for the new design compared the SRS is approximately 1300 after that any increase in Sale size the efficiency will be fixed.

Now let us look for the effect of the triggered when we fix the sample size n'_{nn} at 900 and take different triggers at $n^+ = 5, 7{,}10, 12, 15, 16{,}17$.

Table 4.3: RBE for population variance based on different trigger (n^+) at affixed sample size n'_{nn} =900).

The simulation results in table 4.6 and figure 4.6 show that TPIS design has more efficiency in estimator the population mean and variance than SRS without stratification at any value of trigger n^+ whoever this efficiency increases when the trigger n^+ decrease. The highest efficiency of TPIS is when$(n^+ \leq 10)$.

Table 4.6 Efficiency of new Design (TPIS) compared to SRS with different trigger (n+) with fixed simple size n_nn'=900

n^+	n_{nr}	$\widehat{V}\overline{\hat{\mu}_{nr}}$	$MSE(\hat{\mu}_{TPIS})$	$MSE(\hat{\mu}_{SRS)})$	Eff.	Gain Eff.
5	921	0.000442933	0.000265321	0.000421619	159%	59%
7	949	0.000443447	0.000265628	0.000407395	153%	53%
10	1010	0.000442924	0.000265315	0.000379138	143%	43%
12	1062	0.000428101	0.000256436	0.000357614	139%	39%
15	1151	0.000404285	0.00024217	0.000325286	134%	34%
16	1183	0.000441918	0.000264712	0.000314851	119%	19%
17	1218	0.000442842	0.000265266	0.000304067	115%	15%
18	1253	0.000428611	0.000256741	0.000293884	114%	14%

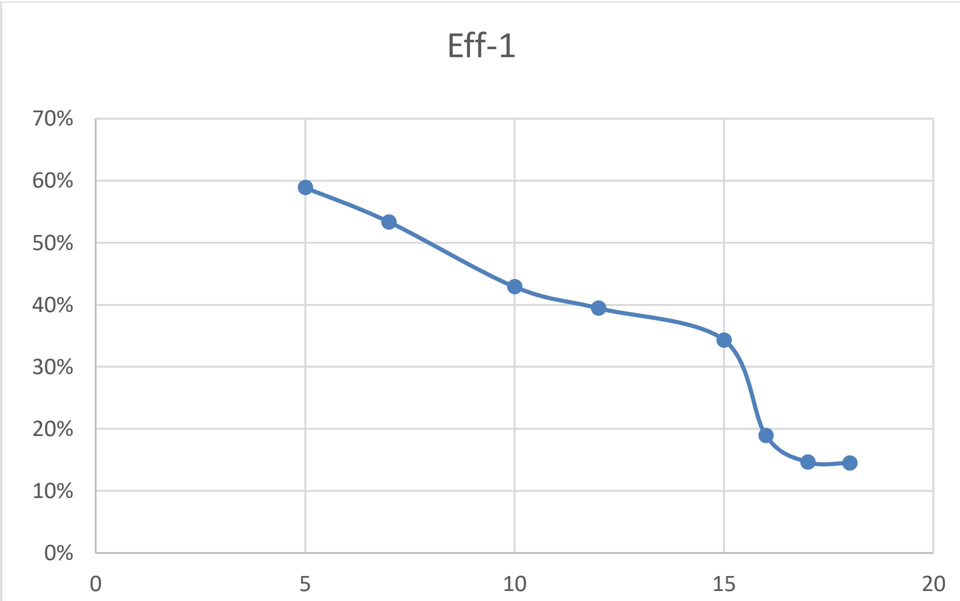

Figure 4.6 Efficiency of new design (TPIS) compared to SRS with different trigger (n+) with fixed simple size n'_{nn}=900

4.2 Conclusions

This research aims to 1) Develop a new sampling technique through post stratified with different methods of data collection in order to reduce or eliminate Non-response bias effects. 2) Investigate the effectiveness of this sampling design by running a simulation study and 3) compare this technique with a simple random sample (SRS). To achieve this objective, new sampling design was developed and case study followed by simulation were conducted.

The results from case study show that the mean estimate of overall satisfaction was higher for the nonresponse groups than the response groups but the difference was not statistically significant at level of $\alpha = 0.05$ but significant at level of$\alpha = 0.10$. However, the overall satisfactions were significantly different for 13 majors out of 51 (25%). Moreover, we introduce TPIS design to eradicate the bias for both cases. Under condition that the second phase response rate is 100% the nonresponse bias will be eradicated. In the case, the response rate is more than 50%, the bias will be reduced significantly.

The results from simulation study show that: 1) Estimator of mean ($\hat{\mu}$) is un-bias estimator of the mean μ for any sample size (n'_{nn}) under TPIS because the relative bias of estimator (RBE) for population mean is very small and it is considered as a simulation error. 2) Under TPIS, the precision of estimator increases when the sample size is approximately ≥ 900. In this case, the RBE is less than 5% and variance estimator is over estimating the variance of estimator. 3) TPIS design is more efficient in estimating the population mean and variance compared to SRS at equal sample size. The efficiency of TPIS increases to 56% when sample size increases to approximately 1300. After this value of sample size, the efficiency remains constant at approximately 56%. 4) By taking in our consideration the pre-determined number of sample from each program n^{+}, the precision of estimator and the efficiency of TPIS compared to SRS not only depends on the sample size but also depends on the trigger n^{+}. The highest precision and efficiency of TPIS happen when n^{+} is small (≤ 10). In this case, the RBE of population mean is less than 5%.

There are three advantages of the new design (TPPSI): 1) We reduce the bias of nonresponse in phase one, 2) We are able to analyses the data at program level using the TPIS design and 3) TPIS design is more efficient than SRS with the same sample size.

4.3 Recommendations and suggestions:

On the result of this research, we provide the following recommendations and suggestions:

- For Qatar University, we recommend Institutional Survey Research to use this new design (ITPS) which reduce the impact of non-response bias especially when the analysis needed by programs.
- For any institution has a frame of population, and needs to estimate any parameter by strata, it is recommended to use the TIPS, which reduce the bias of estimator and provide more precision estimators.
- In this research, the Two-phase post-stratified inverse Sampling Design (TPIS) used with subpopulation sizes known, we suggest to conduct this new design with subpopulation sizes unknown.
- For future research, this work may be extended to multi inverse post-stratified Sampling Design (MIPS) with subpopulation sizes known and with subpopulation sizes unknown.

REFERENCES

Al-Swidi A., Mohammed R., Sheikh and Hafeez, Muhammad H. and Mohd S. and Mohd N. (2015), The role of subjective norms in theory of planned behavior in the context of organic food consumption, British Food Journal, 116. 1561-1580 . doi = 10.1108/BFJ-05-2013-0105

Barclay S., Todd C., Finlay I., Grande G., and Wyatt P. (2002). Not another questionnaire! Maximizing the response rate, predicting non-response and assessing non-response bias in postal questionnaire studies of GPs. Family Practice; 19: 105–111

Bennett B. (1981) On the use of the negative binomial in epidemiology. Biometrical J 23:69–72

Bethlehem J. (2010). Selection bias in web surveys. International Statistical Review, 78(2), 161-188.

Bollmer J., Bethel J., Garrison-Mogren, R., & Brauen, M. (2007). Using the Risk Ratio to Assess Racial/Ethnic Disproportionality in Special Education at the School-District Level. The Journal of Special Education, 41(3), 186–198.

Brick JM, Montaquila JM (2009): Nonresponse and Weighting. Handbook of Statistics. 29 (Part A):163-185

Brick JM, Montaquila JM (2009): Nonresponse and Weighting. Handbook of Statistics. 29(Part A):163-185

BRICK, J. M.; MONTAQUILA, J. M. (2009).Nonresponse and Weighting. Handbook of Statistics, Chapter 8: [s. l.], v. 29, n. Part A, p. 163–185

Casella G, Berger R. (1990) Statistical inference. Duxbury, Belmont Cochran WG (1977) Sampling techniques, 3rd edn. Wiley, New York Haldane JBS (1945) On a method of estimating frequencies, Biometrika 33:222–225

Cervantes, I. F., Brick, M. J., and Jones, M. (2009). Efficiency of post-stratification in complex sampling designs. In JSM 2009 .

Chang, K. C., Liu, J. F., and Han, C. P. (1998). Multiple inverse sampling in post-stratification, J. Statist. Plan Inference,69, 209-227.

Chang, K. C., Han, C. P. and Hawkins, D. L. (1999). Truncated multiple inverse sampling in post-stratification, J. Statist. Plan. Inference,76, 215-234.

Chau, P.Y.K. and Hu, P.J.H. (2001), Information technology acceptance by individual professional: a model comparison approach", Decision Sciences, Vol. 32 No. 4, pp. 699-719.

Christman, M. C. and Lan, F. (2001). Inverse adaptive cluster sampling,Biometrics,57,1096-1105.

Cochran, W.G., 1977. Sampling Techniques, 3rd Edition. Wiley, New York.

Curtin, R., Presser, S., and Singer, E. (2005). Changes in telephone survey nonresponse over the past quarter century. Public Opinion Quarterly 69(1): 87–98. doi:10.1093/poq/nfi002.

DDS Research, Inc. (2004). Sampling error in survey research. Researcher's Resource

Dilman D (1991)The Design and administration of mail surveys, Annu Rev Sociol; 17:225-249

Doss, D.C., Hartley, H.O., Somayajulu, G.R., (1979). An exact small sample theory for post-stratification. J. Statist. Plann. Inference 3, 235-248.

Duffy, B., Smith, K., Terhanian, G. and Bremer, J. (2005), “Comparing data from online and face-to-face surveys”, International Journal of Market Research, Vol. 47, No. 6, pp. 615– 639.

Elizabeth D. McInnis, BA, MA (2006). Nonresponse Bias in Student Assessment Surveys: A Comparison of Respondents and Non-Respondents of the National Survey of Student Engagement at an Independent Comprehensive Catholic University. A Dissertation in Human Development, Marywood University , College Of Education And Human Development.

Enrico Fabrizi, Tomasz Żądło. (2018) Survey sampling and small-area estimation. Mathematical Population Studies 25:4, pages 181-183.

Fuchs M, Bossert D, Stukowski S. (2013). Response rate and nonresponse bias - impact of the number of contact attempts on data quality in the European Social Survey. BSM. 2013;117(1):26–45.

Grande, G., & Todd, C. (2002). Not another questionnaire! Maximizing the response rate, predicting non-response and assessing non-response bias in postal questionnaire studies of GPs. Family Practice, 19(1), 105-111

Groves, R. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. The Public Opinion Quarterly, 70(5), 646-675. Retrieved from http://www.jstor.org/stable/4124220

Groves, R., & Peytcheva, E. (2008). The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis. The Public Opinion Quarterly, 72(2), 167-189. Retrieved from http://www.jstor.org/stable/25167621

Groves, R.M. and Couper, M.P. (1998). Nonresponse in household interview surveys. New York: John Wiley and Sons. doi:10.1002/9781118490082

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E. and Tatham, R.L. (2010), Multivariate Data Analysis, Prentice Hall, Upper Saddle River, NJ.

HaldaneJ. B. S. (1945). A method for estimating frequencies. Biometrika 33(3),222–225.

Hang K., CLiu J., Chien-Pai H (1998).Multiple inverse sampling in post-stratification. Journal of Statistical Planning and Inference, 69 , 209–227.

Lavrakas, P. J. (2008). Encyclopedia of survey research methods Thousand Oaks, CA: Sage Publications, Inc. doi: 10.4135/9781412963947

Lee, M.H. (2011) Statistical Methods for Reducing Bias in Web Surveys. Master Project in statistics, Simon Fraser University

https://www.stat.sfu.ca/content/dam/sfu/stat/alumnitheses/2011/MyoungLee_

Lin, I.-F. and Schaeffer, N.C. (1995). Using survey participants to estimate the impact of nonparticipation. Public Opinion Quarterly 59(2): 236−258. doi:10.1086/ 26947 1.

Lohr, S. L. (2010). Sampling: Design and analysis. Boston: Brooks/Cole.

https://www.worldcat.org/title/sampling-design-and-analysis/oclc/730306813

Lombardi, J. A. (2015). Evaluation of sampling strategies for estimating area of landcover and land-cover change (Order No. 10013999). Available from ProQuest Dissertations & Theses Global. (1766589661).

Lui KJ (2004) Statistical estimation of epidemiological risk. Wiley, New York

Moradi M , Salehi M. (2007). Using General Inverse Sampling Design to Avoid Undefined Estimator. Journal of Probability and Statistical Science,5(2), 137-150.

Murthy, M.N. (1957) Ordered and unordered estimators in sampling without replacement. Sankhy , 18, 379–390.

Ozturk O (2014). Statistical inference for population quantiles and variance in judgment post-stratified samples. Computational Statistics and Data Analysis.77,188-205. doi:10.1016/j.csda.2014.02.021.

Raphael N, James W and Michael E (2016), International Statistical Review 84, 1, 43–62 doi:10.1111/insr.12100

Roberts, C., Vandenplas, C., & Ernst Stähli, M. (2014). Evaluating the impact of response enhancement methods on the risk of nonresponse bias and survey. Survey Research Methods, 8(2), 67–80.

Ronald D. Fricker, J. (2008). Sampling Methods for Web and E-mail Surveys, Fielding: Online Research Methods (Handbook) ,ch 11, pp 195–217

Ronald R. Rindfuss, Minja K. Choe, Noriko O. Tsuya, Larry L. Bumpass (2015). Do low Survey Response Rates Bias Results? Evidence From Japan , Demographic Research ,Volume 32, Article 26, Pages 797−828 . http://www.demographic-research.org/volumes/vol32/26/ doi: 10.4054/demres.2015.32.26

Ropert M Groves (2006): Non Response Rate and non response bias in house hold survey, Public Opinion Quarterly, Vol. 70, No. 5, Special Issue 2006, pp. 646-675

Rueegg, C. S., Gianinazzi, M. E., Michel, G., Zwahlen, M., von der Weid, N. X., Kuehni, C. E., & and the Swiss Paediatric Oncology Group (SPOG). (2017). No evidence of response bias in a population-based childhood cancer survivor questionnaire survey — Results from the Swiss Childhood Cancer Survivor Study. PLoS ONE, 12(5), e0176442. http://doi.org/10.1371/journal.pone.0176442

Salehi M. Moradi,M., AlKhayat, J. , Brown J. Yousef A. (2015). Inverse adaptive cluster sampling with unequal selection probabilities: case studies on crab holes and arsenic pollution, Australian and New Zealand Journal of Statistics. , 57: 189–201. doi: 10.1111/anzs.12118

Salehi M. (2012). HORVITZ-THOMPSON ESTIMATOR OF POPULATION, MEAN UNDER INVERSE SAMPLING DESIGNS,Bulletin of the Iranian Mathematical Society Vol. 38 No. 2 (2012), pp 333-347.

Salehi, M. M. and Seber G. A. F. (2004). A general inverse sampling scheme and its application to adaptive cluster sampling. Aust. N. Z. J. Stat., 46, 483–494.

Salehi, M. M. and Seber G. A. F. (2001). A new proof of Murthy’s estimator which appliesto sequential sampling, Australian & New Zealand Journal of Statistics46(2004)(3)483–494

Salehi,M.M. and Chang, K.-C. (2005) Multiple inverse sampling in post-stratification with sub-population sizes unknown: a solution for quota sampling. J. Stat. Plann. Inference, 131 (2), 379–392.

Sarndal, C. E., Swensson, B., and Wretman, J. (1992). Model Assisted Survey Sampling, New York, Springer Verlag.

ScheaGer, R.L., Mendenhall, W., Ott, L., (1990). Elementary Survey Sampling, 4th Edition. PWS-KENT, Boston.

Seber, G.A.F., Salehi M.M. (2012). Inverse Sampling Methods. In: Adaptive Sampling Designs. Springer Briefs in Statistics. Springer, Berlin, Heidelberg

-Seber GAF and Salehi M. (2013).Adaptive Sampling Designs: Inference for Sparse and Clustered Populations, Berlin: Springer-Verlag.

Singer, E. (2006). Special Issue: Nonresponse Bias in Household Surveys, Public Opinion Quarterly vol. 70.

Singh, Sarjinder. (2003). Advanced sampling theory with applications: How Michael “selected” Amy. 2 Vols. 10.1007/978-94-007-0789-4.

Sinibaldi, J., Trappmann, M., & Kreuter, F. (2014). Which is the better investment for nonresponse adjustment: purchasing commercial auxiliary data or collecting interviewer observations? doi: 10.1093/poq/nfu003.

Tang M.L., Ng H.K.T. (2011). Inverse Sampling. In: Lovric M. (eds) International Encyclopedia of Statistical Science. Springer, Berlin, Heidelberg

Tang ML, Liao Y, Ng HKT (2008). Testing rate ratio under inverse sampling. Biometrical J 49:551–564

Tang ML, Liao Y, Ng HKT, Chan PS (2007). On tests of rate ratio under standard inverse sampling. Comput Meth Prog Bio 89:261–268

Tang ML, Tian M (2009) asymptotic confidence interval construction for risk difference under inverse sampling. Compute Stat Data An 53:621–631

Tang ML, Tian M (2010). approximate confidence interval construction for risk difference under inverse sampling. Stat Compute 20:87–98

Thompson, S.K., Seber, G.A.F., (1996). Adaptive Sampling. Wiley, New York.

Tian M, Tang ML, Ng HKT, Chan PS (2008). Confidence interval estimators for risk ratio under inverse sampling. Stat Med 27:3301–3324

Tian M, Tang ML, Ng HKT, Chan PS (2009). A comparative study of confidence intervals for negative binomial proportion. J Stat Compute Sim 33:241–249

Tianji Cai and Hongyu Wang (2018). Comparing Weighting Methods for Nonresponse in the HKPSSD Survey and the BCSPS', Chinese Sociological Review DOI: 10.1080/21620555.2017.1311772

Vandenplas C, Joye D, Staehli ME, Pollien A (2015). Identifying pertinent variables for non-response follow-up surveys: lessons learned from four cases in Switzerland. SURVEY RESEARCH METHODS.;9(3):141-158.

Vera Toepoel & Matthias Schonlau (2017) Dealing with nonresponse: Strategies to increase participation and methods for post-survey adjustments, Mathematical Population Studies, 24:2, 79-83, DOI: 10.1080/08898480.2017.1299988, https://doi.org/10.1080/08898480.2017.1299988 [accessed Mar 07 2019].

Wagner J (2012); A Comparison of Alternative Indicators for the Risk of Nonresponse Bias, Public Opinion Quarterly, Volume 76, Issue 3, 555–575, https://doi.org/10.1093/poq/nfs032

Ye Cong (2012). Adjustments for nonresponse, sample quality indicators, and nonresponse error in a total survey error context, Ph. D Dissertation, University of Maryland,

https://drum.lib.umd.edu/bitstream/handle/1903/13847/ye_umd_0117e_13720.pdf.html;jsessionid=F0B5E8370126643F590C5D9FBA0F112E?sequence=1

APPENDICES

Appendix (A): Unbiased Estimator for Population Mean using TPIS

We use the terminology of Sarjinder Singh (2003), for the post stratifications used in phase 2. In this phase, n_2' is subdivided into L homogeneous subgroups called strata such that the i^{th} stratum consist of n_{2i}' unites where $h = 1,2, \dots \dots \dots \dots, l$ and $\sum_{h=1}^{l} n_{2i}' = n_2'$.

From the i^{ti} stratum consisting of n_{2i}' units, a sample of size n_{2i} is drawn using any sampling technique such that $\sum_{i=1}^{l} n_{2i} = n_2$, the required sample size.

Let y_{1i}be the value of the study variable in the n_1' and the y_{2j} is the sample value of the study variable in non-respondents unites n_2'.

Then

$$\bar{y}_1 = \frac{1}{n_1'} \sum_{j=1}^{n_1'} y_j \tag{1}$$

$$\bar{y}_d = \frac{n_1'}{n'} \bar{y}_1' + \frac{n_2'}{n'} \bar{y}_{st}' \tag{2}$$

Let y_{2ii} be the i^{ti} population value of the study variable in the i^{ti} stratum , $i = 1, 2, \dots, n_{2i}'$ such that the i^{ti} stratum population mean is given by :

$$\bar{y'}_{2i} = \frac{1}{n_{2i}'} \sum_{i=1}^{n_{2i}'} y_{2ii} \qquad \text{for } i = 1,2, \dots, l$$

Using the concept of weighted average the true population mean at the whole strata can be written as:

All the through us i rather h

$$\bar{y}_2 = \frac{n_{21}'\bar{y}_{21} + n_{22}'\bar{y}_{22} + \cdots, + n_{2l}'\bar{y}_{2l}}{n_{21}' + n_{22}' + \cdots, n_{2l}'} \tag{3}$$

$$\bar{y}_2 = \frac{n_{21}'\bar{y}_{21} + n_{22}'\bar{y}_{22} + \cdots, + n_{2l}'\bar{y}_{2l}}{n_2'}$$

$$= \left(\frac{n'_{21}}{n'_2}\right)\bar{y}_{21} + \left(\frac{n'_{22}}{n'_2}\right)\bar{y}_{22+\cdots,}\left(\frac{n'_{2l}}{n'_2}\right)\bar{y}_{2l}$$

$$\bar{y}_2 = W_{21}\bar{y}'_{21} + W_{22}\bar{y}_{22+\cdots,} + W_{2l}\bar{y}_{2l} = \sum_{i=1}^{l} W_{2i}\,\bar{y}_{2i}$$

Where $W_{2i} = \frac{n'_{2i}}{n'_2}$

$$\bar{y}_{2i} = \frac{1}{n'_{2i}}\sum_{i=1}^{n'_{2i}} y_{2ij}$$

Consider a sample of size of n_{2i} is drawn using sampling technique such that $\sum_{i=1}^{l} n_{2i} = n_2$, the required sample size from post stratified sampling.

$$\bar{y}_d = \frac{n'_1}{n'}\bar{y}'_1 + \frac{n'_2}{n'}\bar{y}'_{st}$$

$$\bar{y}_{st} = \sum_{i=1}^{l}\frac{n'_{2i}}{n'_2}\,\bar{y}_{2i}\ , \qquad \bar{y}_{2i} = \frac{1}{n'_{2i}}\sum_{j=1}^{n'_{2i}} y_{2ij}$$

$$E(\bar{y}_d) = \mu$$

$$E(\bar{y}_d) = E E_{II} E_{III}\,(\bar{y}_d|S_1, S_{II}\,) \tag{4}$$

$$E_{III}\,(\bar{y}_d|S_1, S_{II}\,) = E(\tfrac{n'_1}{n'}\bar{y}_1 + \tfrac{n'_2}{n'}\textstyle\sum_{i=1}^{L}\tfrac{n'_{2i}}{n'_2}\bar{y}_{2i}|S_I, S_{II})$$

$$= \frac{n'_1}{n'}\bar{y}_1 + \frac{n'_2}{n'}E_{II}(\textstyle\sum_{i=1}^{L}\tfrac{n'_{2i}}{n'_2}\bar{y}_{2i}) \tag{5}$$

According stratified Sampling theory:

$$E_{II}(\textstyle\sum_{i=1}^{L}\tfrac{n'_{2i}}{n'_2}\bar{y}_{2i}\,) = E_{II}(\,\textstyle\sum_{i=1}^{L}\tfrac{n'_{2i}}{n'_2}\tfrac{1}{n'_{2i}}\sum_{i=1}^{n_I^{2i}} y_{2ii}\,)$$

$$= E_{II}(\tfrac{1}{n'_2}\textstyle\sum_{i=1}^{L}\ \sum_{i=1}^{n_I^{2i}} y_{2ii}) = \bar{Y}'_2 \tag{6}$$

By substituting (6) in (5)

$$E_{III}\,(\bar{y}_d|S_1, S_{II}) = \frac{n'_1}{n'}\bar{y}_1 + \frac{n'_2}{n'}\bar{Y}'_2 \tag{7}$$

By substituting (7) in (4) $\quad (\bar{y}_d) = E E_{II}\left(\frac{n'_1}{n'}\bar{y}_1 + \frac{n'_2}{n'}\bar{Y}'_2 \,\middle|\, S_I\right) = \mu$

Appendix (B): Condition Variance of Estimator $\hat{\mu}$ in Second Phase

Following to the details in Appendix (A), we found the Condition variance of estimator in Second Phase as follow:

$$Var(\bar{y}_d) = Var_I E_{II}(\bar{y}_d|S_1) + E_I Var_{II}((\bar{y}_d|S_1)$$

$$= Var_I E_{II} E_{III}\,(\bar{y}_d|S_I, S_{II}) + \quad E_1 Var_{II} E_{II}(\bar{y}_d|S_1, S_{II})$$

$$+ E_I E_{II} Var_{II}(\bar{y}_d|S_I, S_{II}) \qquad (8)$$

By theorem,

$$Var_{11}(\bar{y}_d|S_1, S_2) = \frac{n'_2}{n'}\left(\frac{1}{n_{2i}} - \frac{1}{n'_{2i}}\right) S^2/n_{2i}$$

$$E_{II} var_{II}(\bar{y}_d|S_I, S_{II}) = E_{II}\left(\frac{\frac{n'_2}{n'}\left(\frac{1}{n_{2i}} - \frac{1}{n'_{2i}}\right) S^2}{n_{2i}}\right) =$$

$$= \frac{n'_2}{n'} \frac{\left(E\left(\frac{1}{n_{2i}}\right) - \frac{1}{n'_{2i}}\right) S^2}{n2i} \qquad (9)$$

$$E\,\left(\frac{1}{n_{2i}}\right) = ?$$

$$f(x) = \frac{1}{x} \quad ; a = E(X) \qquad ; x = n_{2i} \qquad a = n_2 \frac{n'_{2i}}{n'_2}$$

$$f(x) = f(a) + f'(a)(x - a) + f"\,(a)\,(x - a)^2\ +\ \ldots$$

$$f'(x) = \ -\frac{1}{x^2} \quad ;\ f"(x) = \frac{2}{x^3}$$

$$a = n_2 \frac{n'_{2i}}{n'_2}$$

$$f(n_{2i}) \approx \frac{n'}{n_2 n'_{2i}} - \frac{n'^2}{n_2^2 n{2'}_{2i}}\,(n_{2i} - n_2 \frac{n'_{2i}}{n'_2}) + \frac{2n{3'}_2}{n_2^3\, n{3'}_{2i}}\,\left(n_{2i} - \frac{n_2 n'_{2i}}{n'_2}\right)^2 +$$

$$E\left(\frac{1}{n_{2i}}\right) \approx \frac{n'}{n_2 n'_{2i}}\ + \frac{2n{3'}_2}{n_2^3\, n{3'}_{2i}}\ Var(n_{2i})$$

$$E\left(\frac{1}{n_{2i}}\right) \approx \frac{n'}{n_2 n'_{2i}}\ + \frac{2n{3'}_2}{n_2^3\, n{3'}_{2i}} \frac{n_2 n'_{2i}}{n'_2}\left(\frac{n'_2 - n\,'_{2i}}{n'_2}\ \right)(\frac{n'_2 - n_2}{n'_2 - 1})$$

$$E\left(\frac{1}{n_{2i}}\right) \approx \frac{n'}{n_2 n'_{2i}} + \frac{2n'}{n_2^2\, n^{2}{}'_{2i}} \quad (n'_2 - n'_{2i})(\frac{n'_2 - n_2}{n'_2 - 1})$$

$$E\left(\frac{1}{n_{2i}}\right) \approx \frac{n'}{n_2 n'_{2i}} + \frac{2n'_2}{n_2^2\, n^{2}{}'_{2i}} \left(\frac{n'_2 - n'_{2i}}{n'_2}\right)\left(\frac{n'_2 - n_2}{n'_2 - 1}\right) \qquad (10)$$

Substituting (10) in (9)

$$E_{II}Var_{II}(\bar{y}_d|S_I, S_{II}) == \frac{n'_2}{n'} \frac{\left(E\left(\frac{1}{n_{2i}}\right) - \frac{1}{n'_{2i}}\right)S^2}{n_{2i}}$$

$$E_{II}Var_{II}(\bar{y}_d|S_I, S_{II}) == \frac{n'_2}{n'}\frac{1}{n_{2i}} \left(\frac{n'}{n_2 n'_{2i}} + \frac{2n'_2}{n_2^2\, n^{2}{}'_{2i}}\left(\frac{n'_2 - n'_{2i}}{n'_2}\right)\left(\frac{n'_2 - n_2}{n'_2 - 1}\right) - \frac{1}{n'_{2i}}\right)S^2 \qquad (11)$$

$$E_{III}\,(\bar{y}_d|S_1, S_{II}) = \frac{n'_1}{n'}\bar{y}_1 + \frac{n'_2}{n'}\bar{Y}'_2$$

$$E_{II}E_{III}\,(\bar{y}_d|S_1, S_{II}) = E_{II}\left(\frac{n'_1}{n'}\bar{y}_1 + \frac{n'_2}{n'}\bar{Y}'_2\right) = \mu$$

$$Var_I E_{II} E_{III}\,(\bar{y}_d|S_I, S_{II}) = Var_1(\mu) = 0$$

$$E_{11}(\bar{y}_d|S_1, S_2) = \bar{Y}_2 = \frac{1}{n'_2}\sum_{i=1}^{n'_2} y_i \qquad (12)$$

So,

$$Var_{II}E_{II}(\bar{y}_d|S_1, S_{II}) = Var_{II}(\frac{1}{n'_2}\sum_{i=1}^{n'_2} y_i) = \frac{1}{n'_2}\sigma_2^2 \qquad (13)$$

By substituting (11), (12),(13) we found that:

$$Var(\bar{y}_d) = Var(\hat{\mu}_n)$$

$$= \left(\frac{n'_2 - n'}{n' * n'_2}\right) \sum_{i=1}^{L}\left(\frac{n'_{2i}}{n'_2}\right)\sigma_{n'i}^2 + \frac{1}{n'^2}\left(\frac{n'_2 - n'}{n'_2 - 1}\right)\sum_{i=1}^{L}\left(\frac{n'_2 - n'_2}{n'_2}\right)\sigma_{ni}^2$$

Appendix (C): Survey Validity Using Confirmatory Factor Analysis

The survey structure validity was identified by Confirmatory Factor Analysis (CFA) Using AMOS program. There are some assumption should be achieved before using CFA, Alsouidi A. (2015, p.15) presented these assumptions as follow:

Table 5 Assumption of confirmatory factor analysis

fit indices	required values
Goodness of Fit Index (GFI)	<=0.90
Root Mean Square Error Approximation (RMSEA)	<= 0.08
X^2_{df}	<= 3
Comparative Fit Index (CFI)	≥ 0.90
Non-normed Fit Index (NNFI)	≥ 0.90

Source: (in Al-Swidi, A., 2015, p.15)

` To confirm the structure validity, factor loadings can be used to ensure that all the items designed to measure a construct should load highly and significantly on the constructs they were designed to measure (Chau & Hu ,2001 ; Hair et al., 2010, Al-Swidi et al, 2015).

To ensure that the structural validity is achieved, factor loadings are considered to ensure that all survey items designed to measure a particular factor must be heavily and significantly loaded on the corresponding factor (Chau & Hu, 2001; Hair et al., 2010, Sweden et al., 2015).)

Figure 5.1 shows that all assumptions are achieved and all items were highly and significantly loaded on corresponding factor which mean that the structural validity of the survey is acceptable.

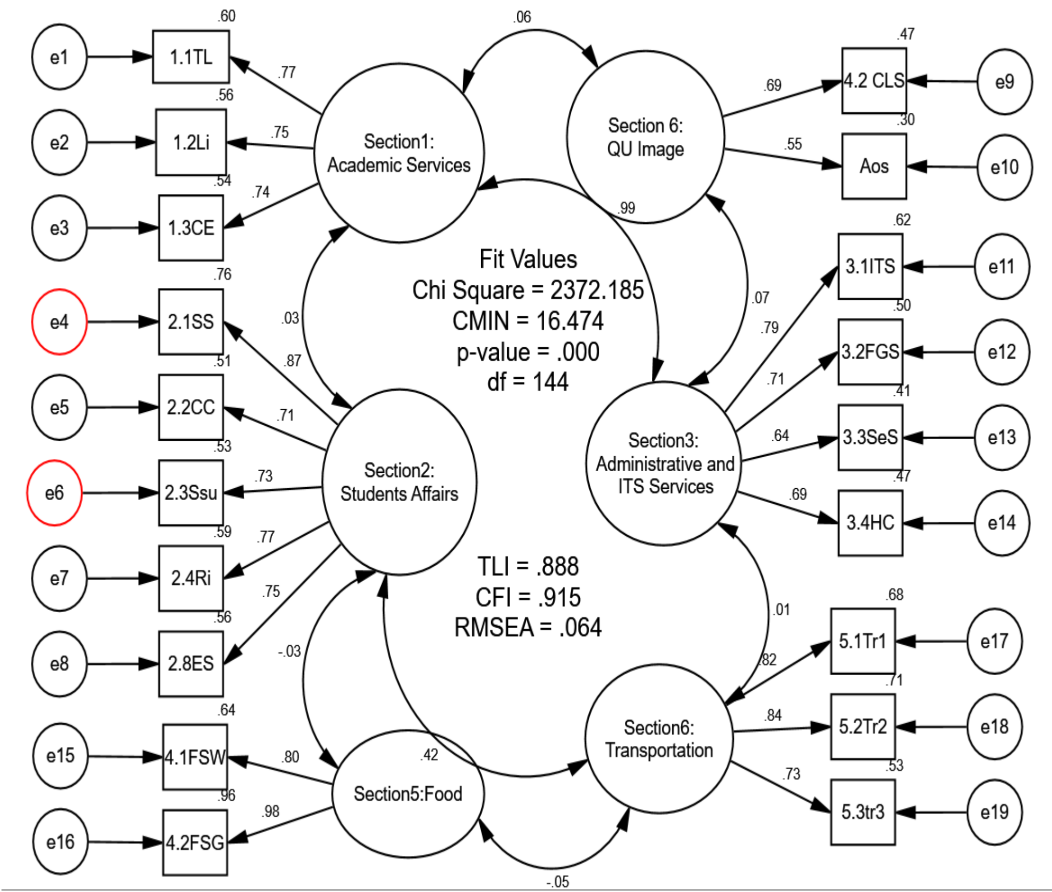

Figure 5 Confirmatory factor analysis for student satisfaction survey

Appendix (D): R-Code for Generating Non-response Subpopulation

```
## Generated non-response population

prog = read.table("D:\\prog.txt",header=F)
sm = sum(prog[,1])
vec = c()
for(i in 1:51){
  n = prog[,1][i]
  prob1 = prog[i,][-1]

  rmultinom(n, size = 1, prob = prob1)
  snr=t(rmultinom(n, size = 1, prob = prob1))
  snr1=data.frame(snr)

  snr1[,1] = snr1[,1]*1
  snr1[,2] = snr1[,2]*2
  snr1[,3] = snr1[,3]*3
  snr1[,4] = snr1[,4]*4

  vec = c(vec,apply(snr1, 1,sum))
}
vec
nam = rep(1:51, prog[,1])
rs = cbind(nam, vec)
rslt = write.table(rs, "D:\\rslt.xls", col.names=F,row.names=F,sep="\t")
```

Appendix (E): R-Code of Simulation Study

```
library(haven)
pop <- read_sav("pop.sav")
#View(pop)
dim(pop)
dim(pop)
POP2=data.frame(1:13701,pop[,7],pop[,5])
#View(POP2)
dim(POP2)
#PPstra(POP2,492,7)
PPstra=function(pop,n,nn){
  res=c()
  sa=sample(pop[,1],n)
  rem=pop[,1][-sa]
  sam=pop[sa,]
  nsam=pop[rem,]
  N=dim(pop)[1]
  Ni=as.vector(table(pop[,2]))
  ni=as.vector(table(sam[,2]))
  L=length(Ni)
  a=matrix(0,L,1)
  row.names(a)=1:L
  b=table(sam[,2])
  a[names(b),]=b
  a=nn-a
  a[,1][a[,1]<=0]=0
  a=as.vector(a)
  M=0
    me=0
    varmuh=0
    for(i in 1:L){
      if(a[i]>0){
      sta=nsam[,1][nsam[,2]==i]
      ssa=sample(sta,a[i])
        sam=rbind(sam,pop[ssa,])
      }
      y=sam[,3][sam[,2]==i]
      m=length(y)
      M=M+m
      me=me+(Ni[i]/N)*mean(y)
      varmuh=varmuh+(N-n)/(n*N)*(Ni[i]/N)*var(y)+(1/n^2)*(N-n)/(N-1)*(N-
Ni[i])/N*var(y)
    }
    res=c(me,varmuh,M)
  res
}

restt=c()
```

```
for(i in 1:1000)
restt=rbind(restt,PPstra(POP2,700,7))
apply (restt,2,mean)
must=mean(restt[,1])
nut=mean(restt[,3])
RV=(MSEt-MSESt)/MSEt
MSEt=sum((restt[,1]-mean(restt[,1]))^2)/(dim(restt)[1]-1)
MSESt=(mean(restt[,2])+(mean(POP2[,3])-must)^2)
#var(POP2[,3])
MSEt
MSESt
#hist(restt[,1])
#hist(restt[,2])
##ad.test(res15[,1])
```