Thailand Statistician
July 2019; 17(2): 198-211
http://statassoc.or.th
Contributed paper

New Population Total Estimators under Stratified Sampling Design in the Presence of Nonresponse

Chugiat Ponkaew and Nuanpan Lawson*
Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand.
*Corresponding author; e-mail: nuanpan.n@sci.kmutnb.ac.th

Received: 26 March 2018
Revised: 15 October 2018
Accepted: 10 December 2018

Abstract

In this paper we propose new estimators for population total in the presence of nonresponse with both a known and unknown response probability. The proposed estimators are investigated under stratified sampling with two assumptions within each stratum: uniform nonresponse and an overall negligible sampling fraction. We show in theory that one of our proposed estimators with a known response probability is unbiased, while the other estimators are asymptotically unbiased, and the variance estimation of the proposed estimators are all asymptotically unbiased. Finally, we compared the efficiency of the proposed estimators through a numerical comparison. The results showed that the proposed estimators performed well with a very small relative bias, especially for the proposed estimator with an unknown response probability which had a smaller relative root mean square error when compared to the others.

Keywords: Taylor linearization approach, nonresponse, reverse framework, variance estimator.

1. Introduction

In stratified sampling, a population is divided into non-overlapping subpopulation strata, each layer is referred to as a stratum. In each stratum a sample is randomly selected from the subpopulation according to a given sampling design which may be different between stratum or the same. The sample will be further studied to estimate the population total and population mean. Under unequal probability sampling without replacement, an unbiased estimator for estimating population total was proposed by Horvitz and Thompson (1952). This estimator is equal to the sum of the study variable's value in the selected sample multiplied by the inverse of the first-order inclusion probability (sampling weights). Moreover, they also derived the variance and an associated estimator for the proposed estimator. Later, Särndal, Swensson and Wretman (1992) proposed an estimator for population total under stratified sampling based on Horvitz and Thompson's (1952) estimator. They have adjusted Horvitz and Thompson (1952) to estimate the population total within each stratum. Consequently the population total estimator can be obtained by finding the sum of the estimator of population total within each stratum. Särndal and Lundström (2005) also discussed the population total estimator and associated variance estimator under stratified sampling.

Usually nonresponse can occur in sample surveys and subsequently this affects the population total estimator which may lead to bias and so decreased efficiency since the sample size is reduced as a result. Nonresponse was first introduced by Hansen and Hurvitz (1946) who encountered it while conducting mail surveys before subsequently classifying it into two items of nonresponse and unit nonresponse. In the presence of nonresponse, the population total estimators with the assumption of complete response will not work because some units of the study variable cannot be observed. Therefore Neyman (1938) proposed an ideal two-phase framework to estimate the population total in the presence of nonresponse comprised of two phases: the first phase is to form a sample by selecting a finite population total according to a sampling design and the second phase is to form a set of responses under unknown response probabilities. Under stratified sampling design Särndal and Lundström (2005) proposed a population total estimator along with a variance estimator under a two-phase framework with the assumption that the response probability is known but that the true probability of response is usually unknown in practice and so must be estimated using a suitable auxiliary variable (see also Bethelehem 2012; Dihidar 2014). In order to solve the problem, Fay (1991) proposed an alternative concept framework called 'reverse framework' to estimate population total. The procedure begins with a finite population total being randomly classified into subpopulations according to a nonresponse mechanism that includes both respondent and non-respondent population subtotals. Then a random sample will be selected from the two subpopulations. Shao and Steel (1999) developed a method under reverse framework to estimate the variance of the population total estimator. Lawson (2017) proposed a nonlinear estimator which is asymptotically unbiased to estimate both population mean and population total using probability proportional to size with replacement (PPSWR) sampling. Lawson's (2017) estimator is approximately unbiased and does not require a known or estimated response probability when a sampling fraction is negligible. Its variance can be obtained by using the Taylor linearization approach.

In this paper population total estimators in stratified sampling with unequal probabilities within each stratum based on Särndal and Lundström's (2005) estimator and Lawson's (2017) estimator are proposed. The overall sampling fraction is assumed to be negligible and the nonresponse mechanism in each stratum is assumed to have uniform nonresponse. Section 2 reviews the procedure for variance estimation under the reverse framework. The proposed population total estimators are discussed in Section 3. In Section 4, the variance and the estimated variance of the proposed estimators are investigated. Finally conclusions are given in Section 5.

2. Variance Estimation under Reverse Framework

This paper is aimed at developing an estimation method for population total along with a variance estimator for use in the presence of nonresponse. In this section, a procedure to estimate the variance of a population total estimator is discussed. Assume that the sampling design is a stratified sampling design with unequal probability without replacement within each stratum. In this sample, a finite population $U=\{1,2,\ldots,i,\ldots,N\}$ is partitioned into non-overlapping L subpopulation totals called strata and denoted $U_1,U_2,\ldots,U_h,\ldots,U_L$ where $U_h=\{i: i$ belongs to stratum $h\}$ and $U_1 \cup U_2 \cup \ldots \cup U_h \cup \ldots \cup U_L = U$. It is assumed that a sample s_h size n_h is selected without replacement from U_h of size N_h within each stratum and that unit i in each stratum is selected with probability π_{hi} and selected independently across the strata. We also assume that $H=\{1,2,\ldots,h,\ldots,L\}$. Let y be the study variable and y_{hi} be the value of y for any unit i in h stratum. The aim is to estimate the population total of y defined by $Y=\sum_{h\in H}\sum_{i\in U_h} y_{hi}$. Under stratified sampling the Särndal, Swensson and Wretman (1992) estimator to estimate Y is given by $\hat{Y}_\pi=\sum_{h\in H}\sum_{i\in s_h}\frac{y_{hi}}{\pi_{hi}}=\sum_{h\in H}\sum_{i\in U_h}\frac{y_{hi}I_{hi}}{\pi_{hi}}$ where $I_{hi}=1$ if $i\in s_h$ otherwise $I_{hi}=0$. Moreover, $E_S(I_{hi})=P(I_{hi}=1)=\pi_{hi}$ and $V_S(I_{hi})=\pi_{hi}(1-\pi_{hi})$.

Under nonresponse, let r_{hi} denote the response indicator variable of y_{hi} and defined by $r_{hi}=1$ if y_{hi} is observed otherwise $r_{hi}=0$. Let $\boldsymbol{R}=(\boldsymbol{R}_1\ \boldsymbol{R}_2\ \ldots\ \boldsymbol{R}_h\ldots\boldsymbol{R}_L)'$ denote the vector of response indicator where $\boldsymbol{R}_h=(r_{h1}\ r_{h2}\ \ldots\ r_{hNh})'$ is a vector of response indicator in stratum h. We assume that within each stratum the nonresponse machanism is of uniform nonresponse, i.e., $p_{hi}=P(r_{hi}=1)=p_h$, $E_R(r_{hi})=P(r_{hi}=1)=p_h$, $V_R(r_{hi})=p_h(1-p_h)$ and $E_R(r_{hi}r_{hj})=E_R(r_{hi})E_R(r_{hj})=p_h^2$.

Usually the variance estimation for a population total estimator is derived under a two-phase framework which requires a known probability of response, therefore an alternative framework namely, a reverse framework first pointed out by Fay (1991) was the focus for a solution. Later, Shao and Steel (1999) developed a method to estimate the variance of a population total estimator under the reverse framework. Assume that $\hat{Y}$ is the estimator of Y under a reverse framework with an overall negligible sampling fraction the variance of $\hat{Y}$ is approximately according to (see also, Shao and Steel 1999; Haziza 2010) $V(\hat{Y})\approx E_R V_S(\hat{Y}|\boldsymbol{R})$. The estimator of $V(\hat{Y})$ denote by $\hat{V}(\hat{Y})$ and defined by, $\hat{V}(\hat{Y})=\hat{V}_S(\hat{Y}|\boldsymbol{R})$ where $\hat{V}_S(\hat{Y}|\boldsymbol{R})$ is the estimator of $V_S(\hat{Y}|\boldsymbol{R})$, in other words $E_S\left[\hat{V}_S(\hat{Y}|\boldsymbol{R})\right]=V_S(\hat{Y}|\boldsymbol{R})$. Moreover, the overall expectation of $\hat{V}_S(\hat{Y}|\boldsymbol{R})$ is given by $E\left[\hat{V}_S(\hat{Y}|\boldsymbol{R})\right]=E_R\left[E_S\left[\hat{V}_S(\hat{Y}|\boldsymbol{R})\right]\right]=E_R\left[V_S(\hat{Y}|\boldsymbol{R})\right]$.

Lemma 1 *Under stratified sampling with unequal probability without replacement and uniform nonresponse within each stratum: Let $\hat{Y}$ be the estimator of Y under reverse framework with an overall negligible sampling fraction,*

(1) *The variance of $\hat{Y}$ is approximately by $V(\hat{Y})\approx E_R V_S(\hat{Y}|\boldsymbol{R})$,*

(2) *The estimator of $V(\hat{Y})$ is defined by $\hat{V}(\hat{Y})=\hat{V}_S(\hat{Y}|\boldsymbol{R})$.*

3. The Proposed Estimators

The estimation of population total in the presence of nonresponse may lead to a biased estimator since the sample size is reduced. In this section we focus on population total estimators in the presence of nonresponse under stratified sampling, uniform mechanism and with a negligible overall sampling fraction where the sampling design in each stratum is of unequal probability without replacement. This research proposes population total estimators based on Särndal and Lundström's (2005) estimator and Lawson's (2017) estimator, called the linear estimator and ratio estimator, respectively. The ratio estimator does not require the response probability whereas the linear estimator requires this value to be known or estimated. Therefore, we consider the linear estimator in two cases, one p_h is known for all $h\in H$ and the other p_h is unknown for all $h\in H$. We note that if p_h is unknown for some examples of $h\in H$ the linear estimator can be obtained by combining the estimators from the two previous cases. In the following development we also assume that;

(A) Each stratum nonresponse mechanism is uniformly nonresponsive,
(B) The overall sampling fraction is negligible.

3.1. The linear estimator

Särndal, Swensson and Wretman (1992), under stratified sampling with the assumption of complete response, proposed an unbiased estimator to estimate population total based on Horvitz and Thompson's (1952) work. Defined by $\hat{Y}_{\pi}=\sum_{h\in H}\sum_{i\in s_h}\frac{y_{hi}}{\pi_{hi}}$, in the presence of nonresponse, the Särndal, Swensson and Wretman (1992) estimator does not work because a few units in the random sample cannot be observed. Therefore, Särndal and Lundström (2005) proposed an unbiased estimator and associated variance estimator for use under a two-phase framework. The Särndal and Lundström (2005) estimator is denoted by $\hat{Y}=\sum_{i\in s}\frac{r_i y_i}{\pi_i p_i}$. Next, Särndal and Lundström's (2005) estimator was extended to include the reverse framework under stratified sampling and called the linear estimator. Under a reverse framework, the population total estimator for stratum h based on Särndal and Lundström's (2005) estimator is denoted by,

$$\hat{Y}_{h1}=\sum_{i\in s_h}\frac{r_{hi}y_{hi}}{\pi_{hi}p_{hi}}. \tag{1}$$

Therefore, the population total estimator based on Särndal and Lundström's (2005) estimator is defined by, $\hat{Y}_{STS}^{(1)}=\sum_{h\in H}\hat{Y}_{h1}=\sum_{h\in H}\sum_{i\in s_h}\frac{r_{hi}y_{hi}}{\pi_{hi}p_{hi}}$, where $\hat{Y}_{h1}$ is defined in (1). However, under the uniform nonresponse mechanism $p_{hi}=p_h$ for all $i\in U_h$ and $h\in H$ therefore, $\hat{Y}_{STS}^{(1)}=\sum_{h\in H}\sum_{i\in s_h}\frac{r_{hi}y_{hi}}{\pi_{hi}p_h}$.

Definition 1 *If assumption* (A) *is correct under a reverse framework with unequal probability sampling without replacement in each stratum the linear estimator for estimating a population total is defined by,*

$$\hat{Y}_{STS}^{(1)}=\sum_{h\in H}\sum_{i\in s_h}\frac{r_{hi}y_{hi}}{\pi_{hi}p_h}=\sum_{h\in H}\hat{Y}_{h1},$$

where $\hat{Y}_{h1}=\sum_{i\in s_h}\frac{r_{hi}y_{hi}}{\pi_{hi}p_h}$.

In Definition 1 we proposed a linear estimator with known response probability. However, in reality this value is unknown. Shao and Steel (1999) stated that the asymptotically unbiased estimator of p_h is defined by,

$$\hat{p}_h=\sum_{i\in s_h}\frac{r_{hi}}{\pi_{hi}}\Big/\sum_{i\in s_h}\frac{1}{\pi_{hi}}. \tag{2}$$

Next, in Lemma 2 we proposed a linear estimator with an unknown response probability.

Lemma 2 *If assumption* (A) *is correct under a reverse framework with unequal probability sampling without replacement within each stratum where* p_h *is unknown for all examples of* $h\in H$ *then the alternative linear estimator for estimating a population total is obtained by:*

$$\hat{Y}_{STS}^{(2)}=\sum_{h\in H}\left[\sum_{i\in s_h}\frac{r_{hi}y_{hi}}{\pi_{hi}}\sum_{i\in s_h}\frac{1}{\pi_{hi}}\Big/\sum_{i\in s_h}\frac{r_{hi}}{\pi_{hi}}\right]=\sum_{h\in H}\hat{Y}_{h2},$$

where $\hat{Y}_{h2}=\sum_{i\in s_h}\frac{r_{hi}y_{hi}}{\pi_{hi}}\sum_{i\in s_h}\frac{1}{\pi_{hi}}\Big/\sum_{i\in s_h}\frac{r_{hi}}{\pi_{hi}}.$

Proof: $\hat{Y}_{STS}^{(2)}$ can be obtained by estimate p_h in Definition 1 by $\hat{p}_h=\sum_{i\in s_h}\frac{r_{hi}}{\pi_{hi}}\Big/\sum_{i\in s_h}\frac{1}{\pi_{hi}}.$

In Definition 1 and Lemma 2 we proposed the linear estimators for estimating a population total under two situations, one where p_h is known for all $h\in H$ and another where p_h is unknown for all examples of $h\in H$. Next we show how the linear estimator works when p_h is known for all examples of $h\in H$ as an unbiased estimator and as an asymptotically unbiased estimator of Y.

Theorem 1 *If assumption* (A) *is correct under a reverse framework with unequal probability sampling without replacement within each stratum.*

(1) $\hat{Y}_{STS}^{(1)}$ *is an unbiased estimator of* Y.

(2) $\hat{Y}_{STS}^{(2)}$ *is an asymptotically unbiased estimator of* Y.

Proof:

(1) Under reverse framework if p_h is known for all examples of $h\in H$ then the overall expectation of $\hat{Y}_{STS}^{(1)}$ is obtained by,

$$E\left(\hat{Y}_{STS}^{(1)}\right)=E_R E_S\left[\sum_{h\in H}\sum_{i\in s_h}\frac{r_{hi}y_{hi}}{\pi_{hi}p_h}\middle|\boldsymbol{R}\right]=E_R E_S\left[\sum_{h\in H}\sum_{i\in U_h}\frac{r_{hi}y_{hi}I_{hi}}{\pi_{hi}p_h}\middle|\boldsymbol{R}\right]$$

$$=\sum_{h\in H}\sum_{i\in U_h}\frac{E_R(r_{hi})y_{hi}E_S(I_{hi})}{\pi_{hi}p_h}=\sum_{h\in H}\sum_{i\in U_h}\frac{p_h y_{hi}\pi_{hi}}{\pi_{hi}p_h}=\sum_{h\in H}\sum_{i\in U_h}y_{hi}=Y.$$

(2) Under reverse framework if p_h is unknown for all examples of $h\in H$ then the overall expectation of $\hat{Y}_{STS}^{(2)}$ is obtained by,

$$E\left(\hat{Y}_{STS}^{(2)}\right)=E_R E_S\left[\sum_{h\in H}\frac{\sum_{i\in s_h}\frac{r_{hi}y_{hi}}{\pi_{hi}}\sum_{i\in s_h}\frac{1}{\pi_{hi}}}{\sum_{i\in s_h}\frac{r_i}{\pi_{hi}}}\middle|\boldsymbol{R}\right]=E_R E_S\left[\sum_{h\in H}\frac{\sum_{i\in U_h}\frac{r_{hi}y_{hi}I_{hi}}{\pi_{hi}}\sum_{i\in U_h}\frac{I_{hi}}{\pi_{hi}}}{\sum_{i\in U_h}\frac{r_i I_{hi}}{\pi_{hi}}}\middle|\boldsymbol{R}\right]$$

$$\approx\sum_{h\in H}\frac{\sum_{i\in U_h}\frac{E_R(r_{hi})y_{hi}E_S(I_{hi})}{\pi_{hi}}\sum_{i\in U_h}\frac{E_S(I_{hi})}{\pi_{hi}}}{\sum_{i\in U_h}\frac{E_R(r_i)E_S(I_{hi})}{\pi_{hi}}}=\sum_{h\in H}\sum_{i\in U_h}y_{hi}=Y.$$

3.2. The ratio estimator

From Section 3.1, one can see that the linear estimator requires a known or estimated response probability, therefore in this section an alternative estimator based on Lawson's (2017) estimator, to solve the problem, is proposed. Under unequal probability sampling with a probability proportional to size with replacement Lawson (2017) proposed an estimator for population mean defined by $\hat{\bar{Y}}_r = \sum_{i\in s}\frac{r_i y_i}{\pi_i}\Big/\sum_{i\in s}\frac{r_i}{\pi_i}$. Consequently, the population total estimator is obtained by $\hat{Y}_r = N\hat{\bar{Y}}_r = N\sum_{i\in s}\frac{r_i y_i}{\pi_i}\Big/\sum_{i\in s}\frac{r_i}{\pi_i}$. The population total estimator proposed by Lawson (2017) does not require either known or estimated response probability under the uniform nonresponse mechanism with an overall negligible sampling fraction. Under stratified sampling with unequal probability sampling without replacement within each stratum the idea of Lawson's (2017) proposal was adapted. The population mean estimator of stratum h is defined by,

$$\hat{\bar{Y}}_{h3} = \sum_{i\in s_h}\frac{r_{hi} y_{hi}}{\pi_{hi}}\Big/\sum_{i\in s_h}\frac{r_{hi}}{\pi_{hi}}. \quad (3)$$

Consequently, the population total estimator of stratum h is obtained by, $\hat{Y}_{h3} = N_h\hat{\bar{Y}}_{h3}$, where $\hat{\bar{Y}}_{h3}$ is defined in (3). Therefore, the population total based on Lawson's (2017) estimator is defined by $\hat{Y}_{STS}^{(3)} = \sum_{h\in H}\hat{Y}_{h3} = \sum_{h\in H} N_h\hat{\bar{Y}}_{h3}$.

Definition 2 *If assumption (A) is correct under a reverse framework with unequal probability sampling without replacement within each stratum the ratio estimator is defined by:*

$$\hat{Y}_{STS}^{(3)} = \sum_{h\in H} N_h\hat{\bar{Y}}_{h3} = \sum_{h\in H}\hat{Y}_{h3},$$

where $\hat{\bar{Y}}_{h3}$ is defined in (3) and $\hat{Y}_{h3} = N_h\hat{\bar{Y}}_{h3}$.

Theorem 2 *If assumption (A) is correct under a reverse framework with unequal probability sampling without replacement within each stratum the ratio estimator $\hat{Y}_{STS}^{(3)}$ is an asymptotically unbiased estimator of Y.*

Proof: The proof of Theorem 2 is similar to Theorem 1.

4. Proposed Variance Estimators

In Section 3 two population total estimators were proposed. One is a linear estimator and the other is a ratio estimator. In this section the variance and the estimated variance of the proposed estimators are investigated as follows:

4.1. The variance of the linear estimator

Recall from Definition 1 and Lemma 2 the following cases:

Case1: If p_h is known for all examples of $h\in H$ the linear estimator for estimating a population total is defined by; $\hat{Y}_{STS}^{(1)} = \sum_{h\in H}\sum_{i\in s_h}\frac{r_{hi} y_{hi}}{\pi_{hi} p_h}$, by using Lemma 1 the variance of $\hat{Y}_{STS}^{(1)}$ is approximately

$$V(\hat{Y}_{STS}^{(1)}) \approx E_R V_S\left[\sum_{h\in H}\sum_{i\in s_h}\frac{r_{hi} y_{hi}}{\pi_{hi} p_h}\middle|\boldsymbol{R}\right] = \sum_{h\in H} E_R V_S\left[\sum_{i\in s_h}\frac{r_{hi} y_{hi}}{\pi_{hi} p_h}\middle|\boldsymbol{R}\right] = \sum_{h\in H} E_R V_S\left[\sum_{i\in s_h}\frac{z_{h1i}}{\pi_{hi}}\middle|\boldsymbol{R}\right].$$

Therefore

$$V(\hat{Y}_{STS}^{(1)}) \approx \sum_{h\in H} E_R V_S\left[\sum_{i\in s_h}\frac{z_{h1i}}{\pi_{hi}}\middle|\boldsymbol{R}\right], \tag{4}$$

where $z_{h1i}=\frac{r_{hi}y_{hi}}{p_h}$.

Case2: If p_h is unknown for all examples of $h\in H$ the linear estimator is given by; $\hat{Y}_{STS}^{(2)}=\sum_{h\in H}\hat{Y}_{h2}$, by using Lemma 1 the variance of $\hat{Y}_{STS}^{(2)}$ is approximately

$$V(\hat{Y}_{STS}^{(2)}) \approx E_R V_S\left[\sum_{h\in H}\hat{Y}_{h2}\middle|\boldsymbol{R}\right] = \sum_{h\in H} E_R V_S\left[\hat{Y}_{h2}\middle|\boldsymbol{R}\right], \tag{5}$$

where $\hat{Y}_{h2}=\sum_{i\in s_h}\frac{r_{hi}y_{hi}}{\pi_{hi}}\sum_{i\in s_h}\frac{1}{\pi_{hi}}\bigg/\sum_{i\in s_h}\frac{r_i}{\pi_{hi}}$.

From (5), three random sums appear in $\hat{Y}_{h2}$ so therefore it is a nonlinear estimator. Hence the variance of $\hat{Y}_{h2}$ given by $\boldsymbol{R}$ respects sampling design or $V_S\left[\hat{Y}_{h2}\middle|\boldsymbol{R}\right]$ which can be obtained by using the Taylor linearization approach. Under this method the linear estimator of $\hat{Y}_{h2}$ is given by;

$$\hat{Y}_{h2} \approx \text{Constant}+\sum_{i\in s_h}\frac{z_{h2i}}{\pi_{hi}}, \tag{6}$$

where $z_{h2i}=\frac{1}{\sum_{i\in U_h} r_{hi}}\left[N_h r_{hi} y_{hi}+\sum_{i\in U_h} r_{hi}y_{hi}-N_h r_{hi}\frac{\sum_{i\in U_h} r_{hi}y_{hi}}{\sum_{i\in U_h} r_{hi}}\right]$.

Substituting (6) into (5) we have

$$V(\hat{Y}_{STS}^{(2)}) \approx \sum_{h\in H} E_R V_S\left[\sum_{i\in s_h}\frac{z_{h2i}}{\pi_{hi}}\middle|\boldsymbol{R}\right]. \tag{7}$$

4.2. The variance of the ratio estimator

Recall Definition 2 the ratio estimator for estimating a population total is defined by $\hat{Y}_{STS}^{(3)}=\sum_{h\in H}\hat{Y}_{h3}=\sum_{h\in H}N_h\hat{\bar{Y}}_{h3}$. By using Lemma 1 the variance of $\hat{Y}_{STS}^{(3)}$ is approximated by

$$V(\hat{Y}_{STS}^{(3)}) \approx E_R V_S\left[\sum_{h\in H}N_h\hat{\bar{Y}}_{h3}\middle|\boldsymbol{R}\right]. \tag{8}$$

From (8) one can see that $\hat{\bar{Y}}_{h3}$ takes the form of a nonlinear estimator. Therefore the linear estimator of $\hat{\bar{Y}}_{h3}$ can be obtained by using the Taylor linearization approach. Under this method the linear estimator of $\hat{\bar{Y}}_{h3}$ is obtained by,

$$\hat{\bar{Y}}_{h3} \approx \sum_{h\in H}\left[\text{Constant}+\frac{1}{\sum_{i\in U_h} r_{hi}}\sum_{i\in s_h}\frac{r_{hi}}{\pi_{hi}}\left(y_{hi}-\bar{Y}_{hr}\right)\right], \tag{9}$$

where $\bar{Y}_{hr}=\sum_{i\in U_h}r_{hi}y_{hi}\bigg/\sum_{i\in U_h}r_{hi}$.

Substituting (9) into (8) one has

$$V(\hat{Y}_{STS}^{(3)}) \approx \sum_{h\in H} E_R V_S \left[\left. \sum_{i\in s_h} \frac{N_h}{\sum_{i\in U_h} r_{hi}} \frac{r_{hi}}{\pi_{hi}} \left(y_{hi} - \bar{Y}_{hr} \right) \right| \boldsymbol{R} \right] = \sum_{h\in H} E_R V_S \left[\left. \sum_{i\in s_h} \frac{z_{h3i}}{\pi_{hi}} \right| \boldsymbol{R} \right], \quad (10)$$

where $z_{h3i} = \frac{N_h}{\sum_{i\in U_h} r_{hi}} r_{hi} \left(y_{hi} - \bar{Y}_{hr} \right).$

From Equations (4), (7), and (10) the general form of variance of the proposed estimators may be written as:

$$V(\hat{Y}_{STS}^{(m)}) \approx \sum_{h\in H} E_R V_S \left[\left. \sum_{i\in s_h} \frac{z_{hmi}}{\pi_{hi}} \right| \boldsymbol{R} \right], \quad (11)$$

where $z_{h1i} = \frac{r_{hi} y_{hi}}{p_h}$, $z_{h2i} = \frac{1}{\sum_{i\in U_h} r_{hi}} \left[N_h r_{hi} y_{hi} + \sum_{i\in U_h} r_{hi} y_{hi} - N_h r_{hi} \frac{\sum_{i\in U_h} r_{hi} y_{hi}}{\sum_{i\in U_h} r_{hi}} \right]$ and $z_{3hi} = \frac{N_h}{\sum_{i\in U_h} r_{hi}} r_{hi} \left(y_{hi} - \bar{Y}_{hr} \right).$

Theorem 3 *If assumption (A) and (B) are correct under a reverse framework with unequal probability sampling without replacement in each stratum.*

(a) The variance of the proposed estimators can be derived from:

$$V\left(\hat{Y}_{STS}^{(m)}\right) \approx \sum_{h\in H} \left[\sum_{i\in U_h} \frac{1-\pi_{hi}}{\pi_{hi}} E_R\left(z_{hmi}^2\right) + \sum_{i\in U_h} \sum_{j\neq i\in U_h} D_{hij} E_R\left(z_{hmi}\right) E_R\left(z_{hmj}\right) \right],$$

where $D_{hij} = \frac{\pi_{hij} - \pi_{hi}\pi_{hj}}{\pi_{hi}\pi_{hj}}$ *and* $z_{hmi}; m=1,2,3$ *are defined in (11).*

(b) If p_h *is known for all examples of* $h\in H$ *then the variance of the linear estimator is defined by*

$$V\left(\hat{Y}_{STS}^{(1)}\right) \approx \sum_{h\in H} \left[\sum_{i\in U_h} \frac{1-\pi_{hi}}{\pi_{hi}} \frac{y_{hi}^2}{p_h} + \sum_{i\in U_h} \sum_{j\neq i\in U_h} D_{hij} y_{hi} y_{hj} \right].$$

(c) If p_h *is unknown for all examples of* $h\in H$ *then the variance of the linear estimator is given by*

$$V\left(\hat{Y}_{STS}^{(2)}\right) \approx \sum_{h\in H} \left[\sum_{i\in U_h} \frac{1-\pi_{hi}}{\pi_{hi}} \left(\frac{1}{p_h} (y_{hi} - \bar{Y}_h)^2 + (2y_{hi} - \bar{Y}_h)\bar{Y}_h \right) + \sum_{i\in U_h} \sum_{j\neq i\in U_h} D_{hij} y_{hi} y_{hj} \right].$$

(d) The variance of ratio estimator is defined by

$$V\left(\hat{Y}_{STS}^{(3)}\right) \approx \sum_{h\in H} \left[\sum_{i\in U_h} \frac{1-\pi_{hi}}{\pi_{hi} p_h} \left(y_{hi} - \bar{Y}_h \right)^2 + \sum_{i\in U_h} \sum_{j\neq i\in U_h} D_{hij} \left(y_{hi} - \bar{Y}_h \right)\left(y_{hj} - \bar{Y}_h \right) \right].$$

The proof of Theorem 3 is given in Appendix A.

Theorem 4 *If assumptions* (A) *and* (B) *are correct under a reverse framework with unequal probability sampling without replacement within each stratum.*

(a) The estimator of $V\left(\hat{Y}_{STS}^{(m)}\right)$ *can be obtained from*

$$\hat{V}\left(\hat{Y}_{STS}^{(m)}\right)\approx\sum_{h\in H}\left[\sum_{i\in s_h}\frac{1-\pi_{hi}}{\pi_{hi}^2}\hat{z}_{hmi}^2+\sum_{i\in s_h}\sum_{j\neq i\in s_h}\hat{D}_{hij}\hat{z}_{hmi}\hat{z}_{hmj}\right],$$

where $\hat{z}_{h1i}=\dfrac{r_{hi}y_{hi}}{p_h}$, $\hat{z}_{h2i}=\dfrac{1}{\sum_{i\in s_h}\frac{r_{hi}}{\pi_{hi}}}\left[N_h r_{hi}y_{hi}+\sum_{i\in s_h}\frac{r_{hi}y_{hi}}{\pi_{hi}}-N_h r_{hi}\dfrac{\sum_{i\in s_h}\frac{r_{hi}y_{hi}}{\pi_{hi}}}{\sum_{i\in s_h}\frac{r_{hi}}{\pi_{hi}}}\right]$,

$\hat{z}_{3hi}=\dfrac{N_h}{\sum_{i\in s_h}\frac{r_{hi}}{\pi_{hi}}}r_{hi}\left(y_{hi}-\hat{\bar{Y}}_{hr}\right)$, $\hat{\bar{Y}}_{hr}=\sum_{i\in s_h}\frac{r_{hi}y_{hi}}{\pi_{hi}}\Big/\sum_{i\in s_h}\frac{r_{hi}}{\pi_{hi}}$ *and* $\hat{D}_{hij}=\dfrac{\pi_{hij}-\pi_{hi}\pi_{hj}}{\pi_{hij}\pi_{hi}\pi_{hj}}$.

(b) If p_h *is known for all examples of* $h\in H$ *then the estimator of* $V\left(\hat{Y}_{STS}^{(1)}\right)$ *is defined by*

$$\hat{V}\left(\hat{Y}_{STS}^{(1)}\right)\approx\sum_{h\in H}\frac{1}{p_h}\left[\sum_{i\in s_h}\frac{1-\pi_{hi}}{\pi_{hi}^2}r_{hi}y_{hi}^2+\sum_{i\in s_h}\sum_{j\neq i\in s_h}\hat{D}_{hij}r_{hi}r_{hj}y_{hi}y_{hj}\right].$$

Moreover, $V\left(\hat{Y}_{STS}^{(1)}\right)$ *is asymptotically unbiased against* $V\left(\hat{Y}_{STS}^{(1)}\right)$.

(c) If p_h *is unknown for all examples of* $h\in H$ *then the estimator of* $\hat{V}\left(\hat{Y}_{STS}^{(2)}\right)$ *is given by*

$$\hat{V}\left(\hat{Y}_{STS}^{(2)}\right)\approx\sum_{h\in H}\left[\sum_{i\in s_h}\frac{1-\pi_{hi}}{\pi_{hi}^2}\hat{z}_{h2i}^2+\sum_{i\in s_h}\sum_{j\neq i\in s_h}\hat{D}_{hij}\hat{z}_{h2i}\hat{z}_{h2j}\right],$$

where $\hat{z}_{h2i}=\dfrac{1}{\sum_{i\in s_h}\frac{r_{hi}}{\pi_{hi}}}\left[N_h r_{hi}y_{hi}+\sum_{i\in s_h}\frac{r_{hi}y_{hi}}{\pi_{hi}}-N_h r_{hi}\dfrac{\sum_{i\in s_h}\frac{r_{hi}y_{hi}}{\pi_{hi}}}{\sum_{i\in s_h}\frac{r_{hi}}{\pi_{hi}}}\right]$.

Moreover, $\hat{V}\left(\hat{Y}_{STS}^{(2)}\right)$ *is asymptotically unbiased against* $\hat{V}\left(\hat{Y}_{STS}^{(2)}\right)$.

(d) The estimator of $V\left(\hat{Y}_{STS}^{(3)}\right)$ *is defined by*

$$\hat{V}\left(\hat{Y}_{STS}^{(3)}\right)\approx\sum_{h\in H}\left(\frac{N_h}{\sum_{i\in s_h}\frac{r_{hi}}{\pi_{hi}}}\right)^2\left[\sum_{i\in s_h}\frac{1-\pi_{hi}}{\pi_{hi}^2}r_{hi}\left(y_{hi}-\hat{\bar{Y}}_h\right)^2+\sum_{i\in s_h}\sum_{j\neq i\in s_h}\hat{D}_{hij}r_{hi}r_{hj}\left(y_{hi}-\hat{\bar{Y}}_h\right)\left(y_{hj}-\hat{\bar{Y}}_h\right)\right].$$

Moreover, $V\left(\hat{Y}_{STS}^{(3)}\right)$ *is asymptotically unbiased against* $V\left(\hat{Y}_{STS}^{(3)}\right)$.

The proof of Theorem 4 is given in Appendix B.

5. Efficiency Comparisons of the Proposed Estimators

In this section we compare the efficiency of the linear estimators in Definition 1 and Lemma 2 with the ratio estimator in Definition 2 by using their simulated relative biases (RB) and relative root-mean-square errors (RRMSE) (Dihidar, 2014; Särndal and Lundström, 2005), respectively defined as

$$RB(\hat{Y}_{STS}^{(m)})=\frac{\frac{1}{B}\sum_{b=1}^{B}\hat{Y}_{STS[b]}^{(m)}-Y}{Y}, \tag{12}$$

$$RRMSE(\hat{Y}_{STS}^{(m)}) = \frac{\sqrt{\frac{1}{B}\sum_{b=1}^{B}(\hat{Y}_{STS[b]}^{(m)} - Y)^2}}{Y}, \tag{13}$$

where $\hat{Y}_{STS[b]}^{(m)}$ is the value of proposed estimators in the b^{th} iteration, for $m = 1,2,3;$ Y is the population total; and B is the number of replications. We use the simple procedure of unequal probability sampling without replacement (Midzuno 1952) to select sample s_h of size n_h from population U_h of size N_h in each stratum. Under this scheme, the first unit of sample s_h of size n_h in each stratum is selected by a probability proportional to size from population U_h size N_h. Following this, the other units of the sample of size $n_h - 1$ are selected by using simple random sampling without replacement from population size $N_h - 1$. Next, the first and second order inclusion probabilities are given by

$$\pi_{hi} = f(\boldsymbol{k_h}) = \frac{k_{hi}}{K_h}\left(\frac{N_h - n_h}{N_h - 1}\right) + \frac{n_h - 1}{N_h - 1}, \tag{14}$$

$$\pi_{hij} = h(\boldsymbol{k_h}) = \left(\frac{k_{hi} + k_{hj}}{K_h}\right)\frac{N_h - n_h}{N_h - 1}\frac{n_h - 1}{N_h - 2} + \frac{n_h - 1}{N_h - 1}\frac{n_h - 2}{N_h - 2}, \tag{15}$$

where f and h are functions of auxiliary variable k, $K_h = \sum_{i \in U_h} k_{hi}$ and $\boldsymbol{k_h} = [k_{h1} \ \ k_{h2} \ \cdots \ k_{hN_h}]$.

We used real data as per Valliant et al. (2000) and Dihidar (2014) from the US Labor Force Population. We note that in program R, this data called “labor” is contained in the package “PrackTools” and was obtained from the US Current Population Survey carried out in September 1976, which contains information on demographic and economic variables of 478 persons in three strata with population sizes $N = (210, 212, 56)$. In the procedure to compute $RB(\hat{Y}_{STS}^{(m)})$ and $RRMSE(\hat{Y}_{STS}^{(m)})$, we used two variables from this data: the usual amount of weekly wages (WklyWage) and the usual number of hours worked per week (HoursPerWk). Let $y =$ WklyWage, $k =$ HoursPerWk, and $H = \{1,2,3\}$. In the numerical comparison, the following steps were considered:

Step 1. Select a sample of 5% from each stratum by using Midzuno’s (1952) scheme.

Step 2. Generate a response indicator from the uniform nonresponse mechanism with different levels of response probability.

Step 3. Compute $\hat{Y}_{STS}^{(m)}$, where $m = 1,2,3$.

Step 4. Repeat 10,000 times from Steps 1 to 3.

Step 5. Compute $RB(\hat{Y}_{STS}^{(m)})$ and $RRMSE(\hat{Y}_{STS}^{(m)})$, where $m = 1,2,3$.

The simulated RB and RRMSE results of the estimators are reported in Table 1, from which we can see that the overall simulated RB results of the proposed estimators were small and very close to zero, especially for the proposed estimator with an unknown response probability. In addition, the simulated RRMSE of this estimator was smaller than the others, and this value decreased when the response probability increased.

Table 1 The simulated RB and RRMSE of the proposed estimators based on 10,000 samples

$p=(p_1, p_2, p_3)$	$RB(\hat{Y}_{STS}^{(m)})$			$RRMSE(\hat{Y}_{STS}^{(m)})$		
	$RB(\hat{Y}_{STS}^{(1)})$	$RB(\hat{Y}_{STS}^{(2)})$	$RB(\hat{Y}_{STS}^{(3)})$	$RRMSE(\hat{Y}_{STS}^{(1)})$	$RRMSE(\hat{Y}_{STS}^{(2)})$	$RRMSE(\hat{Y}_{STS}^{(3)})$
(0.5, 0.5 ,0.5)	0.0262	-0.0051	-0.0038	0.2615	0.1863	0.1880
(0.7, 0.5 ,0.5)	0.0316	0.0021	0.0034	0.2418	0.1766	0.1786
(0.5, 0.7 ,0.5)	0.0314	0.0015	0.0028	0.2309	0.1691	0.1711
(0.5, 0.5 ,0.7)	0.0127	0.0018	0.0030	0.2602	0.1882	0.1900
(0.7, 0.7 ,0.5)	0.0258	-0.0005	0.0006	0.1999	0.1534	0.1555
(0.7, 0.5 ,0.7)	0.0056	-0.0017	-0.0005	0.2370	0.1771	0.1791
(0.5, 0.7 ,0.7)	0.0062	-0.0017	0.0004	0.2256	0.1688	0.1706
(0.7, 0.7 ,0.7)	0.0059	0.000-4	0.0009	0.1979	0.1552	0.1574
(0.8, 0.7 ,0.7)	0.0059	-0.0007	0.0005	0.1858	0.1491	0.1512
(0.7, 0.8 ,0.7)	0.0073	-0.0005	0.0007	0.1823	0.1480	0.1502
(0.7, 0.7 ,0.8)	-0.0009	-0.0013	-0.0001	0.1961	0.1534	0.1555
(0.8, 0.8 ,0.7)	0.0050	-0.0009	0.0003	0.1708	0.1417	0.1440
(0.8, 0.7 ,0.8)	0.0006	-0.0001	-0.0004	0.1844	0.1464	0.1486
(0.7, 0.8 ,0.8)	0.0017	-0.0003	0.0009	0.1821	0.1475	0.1497
(0.8, 0.8 ,0.8)	0.0004	-0.0005	0.0001	0.1711	0.1418	0.1441

6. Conclusions

In this paper both linear and ratio estimators for estimating population total in the presence of nonresponse under a reverse framework were successfully delivered. It was assumed that the sampling design was stratified with unequal probability without replacement and uniform nonresponse across all strata. Moreover we assumed that the overall sampling fraction was negligible. Under these conditions the ratio estimator has neither known nor estimated response probabilities while conversely the linear estimator requires known or estimated response probabilities unlike ours. Therefore the response probabilities for these estimators are classified into the two cases of knowing all strata and knowing no strata. We note that if the response probabilities are unknown for some strata the linear estimator and associated variance estimator can be obtained by combining the estimators from the two previous cases. We showed in theory that the proposed estimator with a known response probability is unbiased whereas the other estimators are asymptotically unbiased and that the variance estimators of the proposed estimators are all asymptotically unbiased. The results showed that for all levels of response probability the proposed estimators all had low simulated RB. In addition, when the response probability is unknown the proposed estimator had a lower simulated RRMSE than the other estimators for all levels of response probability.

Under the assumption of uniform nonresponse all units in the population have the same response probabilities. The estimated values for our response probabilities can be easily estimated without using any auxiliary information. However, in general, a uniform nonresponse mechanism is not realistic in practice so therefore we aim to extend the proposed estimators to include a missing-at-random (MAR) mechanism which will perform more realistically than a uniform nonresponse mechanism. Under this MAR mechanism if response probabilities are unknown then they can be estimated by some function of an auxiliary variable value using probit or logistic regression.

In this paper the method for estimating the variance of our proposed estimators are derived from Horvitz and Thompson (1952) with the limitation of requiring joint inclusion probabilities which is difficult to implement. In practice the simple method proposed by Hansen and Hurwitz (1943) is good enough to estimate variance if the overall sampling fraction is negligible. Therefore in further work we will focus on variance estimation for the population total estimator in the presence of nonresponse based on Hansen and Hurwitz (1943) estimator.

Acknowledgements

We are grateful to Dr.Yves Berger (University of Southampton, UK) for suggesting this topic and to the referees for their helpful comments.

References

Berger YG. A simple variance estimator for unequal probability sampling without replacement. J Appl Stat. 2004; 31: 305-315.

Bethlehem JG. Using response probabilities for assessing representativity. Statistics Netherlands, Discussion Paper. 2012; 382-399.

Dihidar K. Estimating population mean with missing data in unequal probability sampling. SiT. 2014; 15(3): 369-388.

Fay RE. A design-based perspective on missing data variance. Proceedings of the 1991 Annual Research Conference. US Bureau of the Census. 1991; 429-440.

Hansen MH, Hurwitz WN. The problem of non-response in sample surveys. J Am Stat Assoc. 1964; 41: 517-529.

Haziza D. Resampling methods for variance estimation in the presence of missing survey data. Proceedings of the Annual Conference of the Italian Statistical Society. 2010.

Horvitz DF, Thompson DJ. A generalization of sampling without replacement from a finite universe. J Am Stat Assoc. 1952; 47(260): 663-685.

Lawson N. Variance estimation in the presence of nonresponse under probability proportional to size sampling. Proceedings of 6th Annual International Conference on Computational Mathematics, Computational Geometry & Statistics (CMCGS 2017), Singapore. 2017.

Neyman J .Contribution to the theory of sampling human populations. J Am Stat Assoc. 1938; 33: 101-116.

Särndal CE, Lundström S. Estimation in surveys with nonresponse. New York: John Wiley & Sons. 2005.

Särndal CE, Swensson B, Wretman J. Model assisted survey sampling. New York: Springer-Verlag. 1992.

Shao J, Steel P. Variance estimation for survey data with composite imputation and nonnegligible sampling fractions. J Am Stat Assoc. 1999; 94: 254-265.

Valliant R, Dorfman AH, Royall RM. Finite population sampling and inference: A prediction approach. New York: John Wiley & Sons. 2000.

Appendix A: **Proof of Theorem 3**

Assuming that (A) and (B) hold.

(a) From (11) one has

$$V(\hat{Y}_{STS}^{(m)}) \approx \sum_{h\in H} E_R V_S\left[\sum_{i\in s_h}\frac{z_{hmi}}{\pi_{hi}}\middle|\boldsymbol{R}\right], \tag{A1}$$

where z_{hmi}; $m=1,2,3$ are defined in (11). First of all, we consider $V_S\left[\sum_{i\in s_h}\frac{z_{hmi}}{\pi_{hi}}\middle|\boldsymbol{R}\right]$ in (A1).

Since $\sum_{i\in s_h}\frac{z_{hmi}}{\pi_{hi}}$ has the form of Horvitz and Thompson (1952) estimator therefore

$$V_S\left[\sum_{i\in s_h}\frac{z_{hmi}}{\pi_{hi}}\middle|\boldsymbol{R}\right]=\sum_{h\in H}\left[\sum_{i\in U_h}\frac{1-\pi_{hi}}{\pi_{hi}}z_{hmi}^2+\sum_{i\in U_h}\sum_{j\neq i\in U_h}D_{hij}z_{hmi}z_{hmj}\right]. \quad \text{(A2)}$$

Substituting (A2) into (A1) one has:

$$V(\hat{Y}_{STS}^{(m)})\approx\sum_{h\in H}E_R\left[\sum_{i\in U_h}\frac{1-\pi_{hi}}{\pi_{hi}}z_{hmi}^2+\sum_{i\in U_h}\sum_{j\neq i\in U_h}D_{hij}z_{hmi}z_{hmj}\right]$$

$$\approx\sum_{h\in H}\left[\sum_{i\in U_h}\frac{1-\pi_{hi}}{\pi_{hi}}E_R(z_{hmi}^2)+\sum_{i\in U_h}\sum_{j\neq i\in U_h}D_{hij}E_R(z_{hmi})E_R(z_{hmj})\right].$$

Therefore

$$V(\hat{Y}_{STS}^{(m)})\approx\sum_{h\in H}\left[\sum_{i\in U_h}\frac{1-\pi_{hi}}{\pi_{hi}}E_R(z_{hmi}^2)+\sum_{i\in U_h}\sum_{j\neq i\in U_h}D_{hij}E_R(z_{hmi})E_R(z_{hmj})\right]. \quad \text{(A3)}$$

(b) First of all note that $r_{hi}^2=r_{hi}$, recall from (11) we have; $z_{h1i}=\frac{r_{hi}y_{hi}}{p_h}$ substituting z_{h1i} into (A3) one has

$$V(\hat{Y}_{STS}^{(1)})\approx\sum_{h\in H}\left[\sum_{i\in U_h}\frac{1-\pi_{hi}}{\pi_{hi}}E_R\left(\frac{r_{hi}y_{hi}}{p_h}\right)^2+\sum_{i\in U_h}\sum_{j\neq i\in U_h}D_{hij}E_R\left(\frac{r_{hi}y_{hi}}{p_h}\right)E_R\left(\frac{r_{hj}y_{hj}}{p_h}\right)\right]$$

$$=\sum_{h\in H}\left[\sum_{i\in U_h}\frac{1-\pi_{hi}}{\pi_{hi}}\frac{p_hy_{hi}^2}{p_h^2}+\sum_{i\in U_h}\sum_{j\neq i\in U_h}D_{hij}\frac{p_hy_{hi}}{p_h}\frac{p_hy_{hj}}{p_h}\right]$$

$$=\sum_{h\in H}\left[\sum_{i\in U_h}\frac{1-\pi_{hi}}{\pi_{hi}}\frac{y_{hi}^2}{p_h}+\sum_{i\in U_h}\sum_{j\neq i\in U_h}D_{hij}y_{hi}y_{hj}\right].$$

The proof of (c) and (d) are similar to (b).

Appendix B: **Proof of Theorem 4**

Assuming that (A) and (B) hold.

(a) From Lemma 1, (2) and (11) $V(\hat{Y}_{STS}^{(m)})$ can be estimated by

$$\hat{V}(\hat{Y}_{STS}^{(m)})\approx\sum_{h\in H}\left[\sum_{i\in s_h}\frac{1-\pi_{hi}}{\pi_{hi}^2}\hat{z}_{hmi}^2+\sum_{i\in s_h}\sum_{j\neq i\in s_h}\hat{D}_{hij}\hat{z}_{hmi}\hat{z}_{hmj}\right], \quad \text{(B1)}$$

where $\hat{D}_{hij}=\frac{\pi_{hij}-\pi_{hi}\pi_{hj}}{\pi_{hij}\pi_{hi}\pi_{hj}}$, $\hat{z}_{h1i}=\frac{r_{hi}y_{hi}}{p_h}$, $\hat{z}_{h2i}=\frac{1}{\sum_{i\in s_h}\frac{r_{hi}}{\pi_{hi}}}\left[N_hr_{hi}y_{hi}+\sum_{i\in s_h}\frac{r_{hi}y_{hi}}{\pi_{hi}}-N_h\frac{\sum_{i\in s_h}\frac{r_{hi}y_{hi}}{\pi_{hi}}}{\sum_{i\in s_h}\frac{r_{hi}}{\pi_{hi}}}r_{hi}\right]$

and $\hat{z}_{3hi}=\frac{N_h}{\sum_{i\in U_h}r_{hi}}r_{hi}\left(y_{hi}-\hat{\bar{Y}}_{hr}\right)$.

(b) Recall from (B1) the estimator of $V(\hat{Y}_{STS}^{(1)})$ is defined by

$$\hat{V}(\hat{Y}_{STS}^{(1)}) \approx \sum_{h\in H}\left[\sum_{i\in s_h}\frac{1-\pi_{hi}}{\pi_{hi}^2}\hat{z}_{h1i}^2 + \sum_{i\in s_h}\sum_{j\neq i\in s_h}\hat{D}_{hij}\hat{z}_{h1i}\hat{z}_{h1j}\right]$$

$$= \sum_{h\in H}\left[\sum_{i\in s_h}\frac{1-\pi_{hi}}{\pi_{hi}^2}\left(\frac{r_{hi}y_{hi}}{p_h}\right)^2 + \sum_{i\in s_h}\sum_{j\neq i\in s_h}\hat{D}_{hij}\left(\frac{r_{hi}y_{hi}}{p_h}\right)\left(\frac{r_{hj}y_{hj}}{p_h}\right)\right]$$

$$= \sum_{h\in H}\frac{1}{p_h^2}\left[\sum_{i\in s_h}\frac{1-\pi_{hi}}{\pi_{hi}^2}r_{hi}y_{hi}^2 + \sum_{i\in s_h}\sum_{j\neq i\in s_h}\hat{D}_{hij}r_{hi}r_{hj}y_{hi}y_{hj}\right]. \tag{B2}$$

From (B2), the overall expectation of $\hat{V}(\hat{Y}_{STS}^{(1)})$ is obtained by:

$$E\left[\hat{V}(\hat{Y}_{STS}^{(1)})\right] = E_R E_S\left[\hat{V}(\hat{Y}_{STS}^{(1)})\middle|\boldsymbol{R}\right]$$

$$= E_R E_S\left[\sum_{h\in H}\frac{1}{p_h^2}\left[\sum_{i\in s_h}\frac{1-\pi_{hi}}{\pi_{hi}^2}r_{hi}y_{hi}^2 + \sum_{i\in s_h}\sum_{j\neq i\in s_h}\hat{D}_{hij}r_{hi}r_{hj}y_{hi}y_{hj}\right]\middle|\boldsymbol{R}\right]$$

$$= E_R E_S\left[\sum_{h\in H}\frac{1}{p_h^2}\left[\sum_{i\in U_h}\frac{1-\pi_{hi}}{\pi_{hi}^2}r_{hi}I_{hi}y_{hi}^2 + \sum_{i\in U_h}\sum_{j\neq i\in U_h}\hat{D}_{hij}r_{hi}r_{hj}y_{hi}I_{hi}y_{hj}I_{hj}\right]\middle|\boldsymbol{R}\right]$$

$$= \sum_{h\in H}\frac{1}{p_h^2}\left[\sum_{i\in U_h}\frac{1-\pi_{hi}}{\pi_{hi}^2}E_R(r_{hi})E_S(I_{hi})y_{hi}^2\right.$$

$$\left. + \sum_{i\in U_h}\sum_{j\neq i\in U_h}\frac{\pi_{hij}-\pi_{hi}\pi_{hj}}{\pi_{hij}\pi_{hi}\pi_{hj}}E_R(r_{hi})E_R(r_{hj})y_{hi}E_S(I_{hi}I_{hj})y_{hj}\right]$$

$$= \sum_{h\in H}\frac{1}{p_h^2}\left[\sum_{i\in U_h}\frac{1-\pi_{hi}}{\pi_{hi}^2}p_h\pi_{hi}y_{hi}^2 + \sum_{i\in U_h}\sum_{j\neq i\in U_h}\frac{\pi_{hij}-\pi_{hi}\pi_{hj}}{\pi_{hij}\pi_{hi}\pi_{hj}}p_h p_h \pi_{hij}y_{hi}y_{hj}\right]$$

$$= \sum_{h\in H}\left[\sum_{i\in U_h}\frac{1-\pi_{hi}}{\pi_{hi}}\frac{y_{hi}^2}{p_h} + \sum_{i\in U_h}\sum_{j\neq i\in U_h}D_{hij}y_{hi}y_{hj}\right] = V(\hat{Y}_{STS}^{(1)}).$$

The proof of (c) and (d) are similar to (b).