DISSERTATION

UNBIASED RATIO ESTIMATION FOR FINITE POPULATIONS

Submitted by

Jehad Al-Jararha

Department of Statistics

In partial fulfillment of the requirements

for the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2008

COLORADO STATE UNIVERSITY

December 3, 2007

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY JEHAD AL-JARARHA ENTITLED UNBIASED RATIO ESTIMATION FOR FINITE POPULATIONS BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Jean Opsomer (Committee Member)

Hari K. Iyer (Committee Member)

Sanjay Ramchander (Committee Member)

F. Jay Breidt (Adviser and Department Head)

ABSTRACT OF DISSERTATION

UNBIASED RATIO ESTIMATION FOR FINITE POPULATIONS

In many sample surveys from finite populations, the value of an auxiliary variable x is available (at least in aggregate form) for the entire finite population, and is correlated with the study variable of interest y. This auxiliary variable can be used to improve the precision of the estimator of the y-total.

One method of improving precision is through finite population ratio estimation, which has been extensively discussed in the literature, especially under simple random sampling without replacement (SI). Hartley and Ross (1954) obtained an exactly unbiased estimator for the finite population ratio under SI, and hence an unbiased ratio estimator of the y-total. Other authors have obtained an almost unbiased estimator for the finite population ratio, or have considered alternative sampling designs to obtain an unbiased or an almost unbiased estimator for this parameter.

In this work, the Hartley and Ross (1954) estimator is generalized to unequal-probability sampling designs, under the condition of measurability (strictly positive second-order inclusion probabilities). This results in generalized Hartley and Ross (GHR) estimation. Two distinct versions are considered.

The first builds on the Horvitz and Thompson (1952) estimator. This GHR estimator is unbiased and an exact variance and an unbiased estimator for the exact variance are obtained. The computations for the exact variance and the unbiased variance estimator of the GHR require higher-order inclusion probabilities (up to fourth order), which are not easily obtained in general. To overcome this problem, two methods of approximation are given.

The GHR estimator is shown to be mean square consistent under mild conditions. These conditions are met, for example, by simple random sampling without replacement, simple random cluster sampling, and stratified sampling designs.

Central limit theorems (CLTs) are established for GHR under the SI design and under the Poisson sampling (PO) design. The asymptotic variance and a consistent estimator for the asymptotic variance are given under both designs.

The GHR is evaluated under a super-population model, and it is shown that the Godambe and Joshi (1965) lower bound is attainable for GHR under SI and PO sampling designs. The GHR is compared to other estimators analytically and via simulation.

The second version of GHR is derived using a Hansen and Hurwitz (1943) type estimator for with-replacement sampling. This estimator is unbiased. This estimator is discussed under two different asymptotic scenarios, when the population size N is fixed and number of independent draws m tends to infinity and when both m and N tend to infinity. Under each of the two cases, a CLT is established and the asymptotic variance and a consistent estimator for the variance are given. The Godambe and Joshi (1965) lower bound is shown to be attainable for the second case.

An important problem in applications is estimation of the population total t_y under a stratified sampling design when stratum x-totals are known, particularly in the case of small stratum sizes. If biased estimators are used to estimate within-stratum population y-totals, the bias may accumulate across strata. The unbiased GHR estimators can be used effectively in dealing with such situations by introducing a separate GHR estimator, analogous to the classic separate ratio estimator of survey statistics. A CLT is proven for the separate GHR estimator under a stratified sampling design when the stratum sizes are fixed and the number of strata

tends to infinity. Simulation results show that GHR under different sampling designs gives excellent results compared to other almost unbiased estimators proposed in the literature, even when the number of strata is not large.

Jehad Al-Jararha
Department of Statistics
Colorado State University
Fort Collins, Colorado 80523
Spring 2008

ACKNOWLEDGEMENTS

At the beginning, I would like to thank my advisor, Dr. Jay Breidt, for his suggestions, help, patience and friendship through the last few years I have been working with him. I also give thanks to my committee members, Dr. Jean Opsomer, Dr. Hari Iyer and Dr. Sanjay Ramchander. My thanks is also extended to my favorite teachers at Colorado State University, Dr. Peter Brockwell, Dr. David Bowden, Dr. Ronald Butler, and Dr. Richard Davis.

I am grateful for my father, mother, wife, daughters, and son for their help, support and patience. Also, I thank my friends from all over the world and every person who has supported me through my work.

DEDICATION

To Asma, Raghad, Rahaf, Sarah, Mohammad, Salsabeel.

CONTENTS

LIST OF TABLES

Chapter 1

INTRODUCTION

1.1 Sampling Designs

In surveys of finite populations, auxiliary information is often available for every element in the population. Ratio estimators use variables that are correlated with the variable of interest. Population registers in some countries contain age and taxable income for all residents. Studies of labor force characteristics or household expenditure patterns might benefit from these auxiliary data. Geographic information systems may contain measurements derived from satellite imagery for all locations. These spatially explicit data can be used in augmenting measurements obtained in agricultural surveys or natural resource inventories.

Consider a finite population U_N consisting of N units $\{1, 2, \ldots, k, \ldots, N\}$. A sample, denoted by s, is a subset from the population. Let us define the following concepts.

Definition 1.1.1 *Sampling design*, $p_N(\cdot)$, is a function mapping the set of all subsets of U_N to $[0, 1]$, where $p_N(s)$ is the probability of selecting the sample s.

Example 1.1.1 A *census* is a sampling design with

$$p_N(s) = \begin{cases} 1, & s = U_N \\ 0, & \text{otherwise.} \end{cases} \tag{1.1}$$

Definition 1.1.2 *First-order inclusion probability*, π_{Ni}, is the probability that a sample s will include the i^{th} element under the sampling design $p_N(\cdot)$:

$$\pi_{Ni} = Pr(i \in s) = \sum_{s \ni i} p_N(s),$$

where the sum is taken over all subsets s of U_N.

Definition 1.1.3 *A probability sampling design* is a sampling design such that $\pi_{Ni} > 0, \quad \forall i \in U_N$.

Definition 1.1.4 *Second-order inclusion probability*, π_{Nij}, is the probability that a sample s will include both i^{th}, j^{th} elements under the sampling design $p_N(\cdot)$,

$$\pi_{Nij} = Pr(ij \in s) = \sum_{s \ni ij} p_N(s).$$

Remark 1.1.1 In similar manner, we can define third-order, fourth-order, and higher-order inclusion probabilities.

Remark 1.1.2 It follows directly that the inclusion probability is invariant to permutations of the indices in the subscript, and that the order is reduced if any indices are repeated. For example, $\pi_{Nijikklj} = \pi_{Nijkl} = \pi_{Njkli}$.

Remark 1.1.3 In developing an asymptotic theory, we consider sequences of finite populations and associated sampling designs as $N \to \infty$. Therefore, the first-order inclusion probabilities π_{Ni}, second-order inclusion probabilities π_{Nij}, and higher-order inclusion probabilities are actually sequences depending on N. For the sake of simplicity in notations, we will drop the subscript N.

Definition 1.1.5 *A measurable sampling design* is a sampling design such that $\pi_{ij} > 0, \quad \forall ij \in U$.

Definition 1.1.6 In general, a *m^{th}-order measurable sampling design* is a sampling design such that all m^{th}-order inclusion probabilities are strictly positive.

In the following examples, we will discuss different sampling designs.

Example 1.1.2 *Simple random sampling without replacement design (SI)* (e.g. Särndal et al. (1992)) assigns equal probability $\binom{N}{n}^{-1}$ to every subset of U that contains exactly n distinct elements. SI can be implemented by drawing the first element from a uniform distribution on all N elements. Discard the first element and draw the second element from a uniform distribution on the remaining $N-1$ elements. Discard the second element and continue this process to select the n^{th} element from a uniform distribution on the last $N-n+1$ elements.

Example 1.1.3 *With-replacement sampling design (WR).* Suppose that $p_k = Pr$ (selecting element k on a single draw) for $k = 1, \ldots, N$, and $\sum_{k \in U_N} p_k = 1$. The WR design is implemented by using the p_k's to draw a first element. The selected element is replaced and the process is repeated until the m^{th} element is drawn. The sample size in this case is $n \leq m$.

Definition 1.1.7 *The sample membership indicators* are

$$I_{\{i \in s\}} = \begin{cases} 1 & \text{if } i \in s \\ 0 & \text{if } i \notin s. \end{cases}$$

Example 1.1.4 *Poisson sampling design, (PO).* Let π_i be the first-order inclusion probability for the i^{th} element $i = 1, \ldots, N$. Under PO, $I_{\{i \in s\}}$ are independent Bernoulli random variables with

$$Pr\left(I_{\{i \in s\}} = 1\right) = \pi_i.$$

To draw a random sample using the Poisson sampling design let $u_1, \ldots, u_N$ be independent *Uniform*(0, 1) random numbers; if $u_i < \pi_i$ then the element i is selected (Särndal et al. (1992)).

Remark 1.1.4 In the Poisson sampling design,

- The sample size $n_s = \sum_U I_{\{i \in s\}}$ is random with mean $E_{PO}(n_s) = \sum_U E_{PO}\left(I_{\{i \in s\}}\right) = \sum_U \pi_i$ and variance $var_{PO}(n_s) = \sum_U \pi_i (1 - \pi_i)$.

- Since the indicators are independent, $\pi_{ij} = \pi_i \pi_j$ for $i \neq j$. In a similar manner, we can define higher-order inclusion probabilities.

Example 1.1.5 *Probability proportional-to-size sampling with-replacement design (pps).* For given positive numbers $x_1, \ldots, x_N$, define

$$p_i = \frac{x_i}{\sum_U x_i}; \qquad i = 1, \ldots, N.$$

Draw a random sample with the same arguments as in the with-replacement sampling design of Example 1.1.3.

Example 1.1.6 *Probability proportional-to-size sampling without-replacement design (πps).* For given positive numbers $x_1, \ldots, x_N$, the first-order inclusion probabilities π_i are strictly proportional to x_i (Särndal et al. (1992)). Brewer and Hanif (1983) described a procedure to draw a πps sample of size $n = 2$. Define $t_x = \sum_U x_k$, $c_k = x_k (t_x - x_k) / t_x (t_x - 2x_k)$, $p_k = c_k / \sum_U c_k$ and assume $x_k < t_x/2$. Use the set of probabilities $p_k \quad (k = 1, \ldots, N)$ to draw a first element. Without replacing the first drawn element (say k_1), give the element l the probability

$$p_{l|k_1} = x_l / (t_x - x_i)$$

of being selected in the second draw. According to this scheme,

$$\pi_i = 2x_i / t_x \quad \text{for } k = 1, \ldots, N$$

and for $i \neq j$

$$\pi_{ij} = \frac{2x_i x_j}{t_x \sum_U c_k} \frac{t_x - x_i - x_j}{(t_x - x_i)(t_x - x_j)}.$$

Remark 1.1.5 For sample size $n > 2$, SAS `proc surveyselect` uses a method due to Hanurav (1967) and Vijayan (1968) to draw a πps sample. The procedure also produces the first-order and second-order inclusion probabilities.

Example 1.1.7 *Stratified sampling design (ST).* The finite population $U_N = \{1, \ldots, N\}$ is divided into H disjoint sub-populations, $U = \cup_{h=1}^{H} U_h$. For $h = 1, \ldots, H$ draw a probability sample s_h from U_h according to a design $p_h(\cdot)$, where the selection in one stratum is independent of the selections in all other strata, (Särndal et al. (1992)).

Example 1.1.8 *Simple random cluster sampling (SIC).* The finite population $U_N = \{1, \ldots, N\}$ is divided into N_I clusters, denoted by $U_1, \ldots, U_{N_I}$. The set of all clusters is a new finite population, denoted by $U_I = \{1, \ldots, N_I\}$. A sample s_I of fixed size n_I is selected from U_I via SI sampling design and all elements in each selected cluster are observed, (Särndal et al. (1992)).

1.2 Unbiased Estimation of Finite Population Total

One of the key interests in finite population sampling is to estimate the population total, $t_y = \sum_{i \in U} y_i$. For each $i \in s$, a value y_i is observed for element i.

Definition 1.2.1 The *Horvitz-Thompson* (HT) *estimator* (Horvitz and Thompson 1952) for the population total t_y, is defined by

$$\hat{t}_{y\pi} = \sum_{i \in s} \frac{y_i}{\pi_i} = \sum_{i \in U} \frac{y_i}{\pi_i} I_{\{i \in s\}}. \tag{1.2}$$

Remark 1.2.1 The Horvitz-Thompson estimator is an unbiased estimator for the population total t_y under any probability sampling design since

$$E_p\left[\hat{t}_{y\pi}\right] = \sum_{i \in U} \frac{y_i}{\pi_i} E_p\left[I_{\{i \in s\}}\right] = \sum_{i \in U} \frac{y_i}{\pi_i} \pi_i = t_y,$$

where $E_p[\cdot]$ is the average over all possible samples under the design. Also, the variance of HT with respect to the sampling design is

$$var_p\left(\hat{t}_{y\pi}\right) = \sum_{ij \in U} \frac{y_i}{\pi_i} \frac{y_j}{\pi_j} \Delta_{ij} \tag{1.3}$$

where $\Delta_{ij} = cov_p\left(I_{\{i\in s\}},\, I_{\{j\in s\}}\right) = \pi_{ij} - \pi_i\pi_j$. For a measurable sampling design, an unbiased estimator of $var\left(\hat{t}_y\pi\right)$ is

$$\hat{v}ar_p\left(\hat{t}_{y\pi}\right) \;=\; \sum_{ij\in s}\frac{y_i}{\pi_i}\frac{y_j}{\pi_j}\frac{\Delta_{ij}}{\pi_{ij}}. \tag{1.4}$$

Under the with-replacement sampling design of example 1.1.3,

$$\pi_k = 1 - (1-p_k)^m,$$

which can be used in constructing the HT estimator. An alternative unbiased estimator can also be derived. Let κ_i denote the element selected in the i^{th} draw, $i = 1, \ldots, m$. Define the indicator $I_{\{\kappa_i=k\}}$ to be one if the k^{th} element is selected in the i^{th} draw, and zero otherwise.

Definition 1.2.2 The *Hansen and Hurwitz (1943) (HH) estimator* for the population total t_y is defined by

$$\hat{t}_{HH} = \frac{1}{m}\sum_{i=1}^{m}\sum_{k\in U}\frac{y_k}{p_k}I_{\{\kappa_i=k\}}.$$

Remark 1.2.2 The Hansen-Hurwitz estimator is unbiased for the population total t_y, since

$$E_p\left[\hat{t}_{HH}\right] = \frac{1}{m}\sum_{i=1}^{m}\sum_{k\in U}\frac{y_k}{p_k}E_p\left[I_{\{\kappa_i=k\}}\right] = \frac{1}{m}\sum_{i=1}^{m}\sum_{k\in U}\frac{y_k}{p_k}p_k = t_y.$$

Furthermore, note that because the sampling is done with replacement, the random variables $\sum_{k\in U}\frac{y_k}{p_k}I_{\{\kappa_i=k\}}$ are independent and identically distributed (iid). See, for example, Section 2.9 of Särndal et al. (1992). It follows easily that

$$var_p\left(\hat{t}_{HH}\right) = \frac{1}{m}\sum_{k\in U}\left(\frac{y_k}{p_k} - t_y\right)^2 p_k,$$

and an unbiased estimator for this variance is

$$\hat{v}ar_p\left(\hat{t}_{HH}\right) = \frac{1}{m\,(m-1)}\sum_{i=1}^{m}\left(\frac{y_{\kappa_i}}{p_{\kappa_i}} - \hat{t}_{HH}\right)^2. \tag{1.5}$$

Since $\hat{t}_{HH}$ is the sample mean of iid random variables with finite variance, it follows from a standard CLT (Casella and Berger (2002) p.236) that

$$\frac{\hat{t}_{HH} - t_y}{\sqrt{\hat{var}_p\left(\hat{t}_{HH}\right)}} \xrightarrow{\mathcal{L}} (0,\, 1) \quad \text{as} \quad m \to \infty.$$

Remark 1.2.3 (with-replacement approximation).

If p_k is small then

$$\pi_k = 1 - (1 - p_k)^m \doteq m p_k,$$

so that the HT estimator under WR design becomes

$$\sum_{k \in U} \frac{y_k}{\pi_k} I_{\{k \in s\}} \doteq \sum_{k \in U} \frac{y_k}{m p_k} \sum_{i=1}^{m} I_{\{\kappa_i = k\}}, \tag{1.6}$$

since $Pr\left(\sum_{i=1}^{m} I_{\{\kappa_i = k\}} > 1\right)$ is very small for p_k small. But (1.6) is the HH estimator, so that HT and HH are expected to behave similarly under WR designs with small p_k.

It is common in practice to extend this approximation to without-replacement designs. Define $p_k = \pi_k / m$ and "pretend" that the sample was drawn with replacement with these probabilities and with $m = n$. equation (1.5) then provides a convenient (approximate) variance estimator, implemented in SAS and other survey software.

Remark 1.2.4 It is easy to establish asymptotic results when the Hansen-Hurwitz estimator is used, since under sampling with-replacement design the indicators $I_{\{\kappa_i = k\}}$ are independent random variables, due to the fact that we have independent draws. However, the asymptotic are not easy when a without-replacement sampling design and the Horvitz-Thompson estimator is used. The difficulties come from the fact that the indicator functions $I_{\{i \in s\}}$ are dependent random variables for

most designs. An exception is Poisson sampling, under which $I_{\{i\in s\}}$ are independent Bernoulli random variables each with with success probability π_i. A special case of Poisson sampling is the *Bernoulli sampling design*, with $\pi_i \equiv \pi \in (0, 1)$. Central limit theory for Poisson sampling has been established in Hájek (1960), and extended to central limit theory for SI in the same work. Additional results and references will be discussed later in this dissertation.

Up to this point, the only randomness that has been discussed is that introduced through the sampling design; in particular, the y_k values have been regarded as fixed, real numbers, not as random variables. To study further the properties of estimators, it is useful to introduce a probabilistic model for the y_k's. This model is referred to as a *superpopulation model,* and commonly denoted by ξ. Suppose that $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_N$ are known auxiliary vector values. Assume the relationship between y_k and $\boldsymbol{X}_k$ is given by

$$\xi: \quad y_k = \boldsymbol{X}_k'\boldsymbol{\beta} + \epsilon_k \tag{1.7}$$

where $E_\xi(\epsilon_k) = 0$, $E_\xi(\epsilon_k^2) = \sigma_k^2$ and for $k \neq l$ $E_\xi(\epsilon_k\epsilon_l) = 0$ where the expectation $E_\xi(\cdot)$ is the average over all realizations from the superpopulation model. If $\hat{t}_y$ is an estimator for t_y, the estimation error $\hat{t}_y - t_y$ can be examined jointly under the model ξ and the sampling design $p(\cdot)$. The anticipated variance (Särndal et al. (1992)) of $\hat{t}_y - t_y$ is

$$E_\xi E_p\left[\left(\hat{t}_y - t_y\right)^2\right] - \left[E_\xi E_p\left(\hat{t}_y - t_y\right)\right]^2.$$

If $E_\xi E_p\left(\hat{t}_y - t_y\right) = 0$, the anticipated variance is

$$E_\xi E_p\left[\left(\hat{t}_y - t_y\right)^2\right].$$

Result 1.2.1 *Godambe and Joshi lower bound (GJLB). Under the model (1.7), if*

$$E_\xi E_p\left(\hat{t_y} - t_y\right) = 0$$

then

$$E_\xi E_p \left(\hat{t_y} - t_y\right)^2 \geq \sum_{k \in U} \left(\frac{1}{\pi_k} - 1\right) \sigma_k^2$$

where $\hat{t_y}$ *is any estimator of the population parameter* t_y *(Godambe and Joshi (1965)).*

1.3 Estimation of a Population Ratio

In many survey applications, it is of interest to estimate the population ratio

$$\theta = \frac{t_y}{t_x} = \frac{\sum_{i \in U} y_i}{\sum_{i \in U} x_i}.$$

Example 1.3.1 Suppose the population consists of agricultural fields of different sizes. Let

$$\begin{aligned} y_i &= \text{bushels of grain harvested in field } i \\ x_i &= \text{acreage of field } i \end{aligned}$$

Therefore, we are interested in yield, which is the population ratio $\theta = t_y t_x^{-1} =$ bushels per acre.

Example 1.3.2 The goal of studies of labor force is to estimate the employment rate

$$\theta = \frac{t_y}{t_x} = \frac{\text{number of employed persons}}{\text{number of persons in labor force}}.$$

The availability of auxiliary information can vary from population to population. Consider the following situations:

- Aux_0 : x_i are available only for $i \in s$.
- Aux_1 : x_i are available only for $i \in s$ and $\bar{x}_{U_N}$ is known.

- Aux_2 : x_i are available only for $i \in s$, and $\bar{x}_{U_h}$ is known for stratum $h = 1, \ldots, H$ where $U = \cup_{h=1}^{H} U_h$.

- Aux_3 : x_i are known for all $i \in U$.

Note that if Aux_j holds the so does Aux_k, $k < j$. Under Aux_1, define $\hat{\theta}_{naive} = \hat{t}_{y\pi} t_x^{-1}$ as an estimator for $\hat{\theta}$. This is clearly an unbiased estimator for θ, with

$$var_p\left(\hat{\theta}_{naive}\right) = \frac{1}{t_x^2} \sum_{ij \in U} \frac{y_i}{\pi_i} \frac{y_j}{\pi_j} \Delta_{ij}. \tag{1.8}$$

This estimator is known to be relatively inefficient, in general.

Under Aux_0, define the *simple ratio estimator*

$$\hat{\theta} = \frac{\hat{t}_{y\pi}}{\hat{t}_{x\pi}} = \frac{\sum_{i \in s} \frac{y_i}{\pi_i}}{\sum_{i \in s} \frac{x_i}{\pi_i}}. \tag{1.9}$$

The estimator $\hat{\theta}$ is considered one of the most important estimators for the population ratio. This estimator is biased since it is a nonlinear function of the unbiased estimators $\hat{t}_{y\pi}$, $\hat{t}_{x\pi}$. It is often impossible to find exact bias or exact variance for this estimator. However, this estimator is asymptotically unbiased. Under a general sampling design, the properties of this estimator will be discussed. When this estimator is used in a separate ratio estimator under stratified sampling, the bias can accumulate, even for moderate numbers of strata. The estimator can then give very poor results as we will see in Section 3.4.

To study the asymptotic properties of $\hat{\theta}$, linearize $\hat{\theta}$ by first order Taylor expansion,

$$\begin{aligned} \hat{\theta} &= \frac{\hat{t}_{y\pi}}{\hat{t}_{x\pi}} \\ &\doteq \frac{t_y}{t_x} + \frac{1}{t_x}\left(\hat{t}_{y\pi} - t_y\right) - \frac{t_y}{t_x^2}\left(\hat{t}_{x\pi} - t_x\right) \\ &= \text{constant} + \frac{1}{t_x} \sum_{i \in U} \left(y_i - \theta x_i\right) \frac{I_{\{i \in s\}}}{\pi_i}. \end{aligned} \tag{1.10}$$

Therefore, the variance of $\hat{\theta}$ is given by

$$var_p\left(\hat{\theta}\right) = \frac{1}{t_x^2}\sum_{i,j\in U}\frac{y_i-\theta x_i}{\pi_i}\frac{y_j-\theta x_j}{\pi_j}\Delta_{ij}, \tag{1.11}$$

where $\Delta_{ij} = \pi_{ij} - \pi_i\pi_j$. Note that (1.11) should be smaller than (1.8) if θx_i explains some of the variation in y_i. For a measurable sampling design, an approximately unbiased estimator for $var_p\left(\hat{\theta}\right)$ is

$$\hat{var}_p\left(\hat{\theta}\right) = \frac{1}{\hat{t}_{x\pi}^2}\sum_{i,j\in s}\frac{y_i-\hat{\theta}x_i}{\pi_i}\frac{y_j-\hat{\theta}x_j}{\pi_j}\frac{\Delta_{ij}}{\pi_{ij}}.$$

This estimator can be shown to be consistent for the true variance under fairly mild conditions on the design, the x_i's and y_i's.

1.3.1 Bias of $\hat{\theta}$

In order to find the bias for $\hat{\theta}$, expand $\hat{\theta}$ to second order by Taylor expansion,

$$\begin{aligned}
\hat{\theta} &\doteq \frac{t_y}{t_x} + \frac{1}{t_x}\left(\hat{t}_{y\pi}-t_y\right) - \frac{t_y}{t_x^2}\left(\hat{t}_{x\pi}-t_x\right) + \left(\frac{-1}{t_x^2}\right)\left(\hat{t}_{x\pi}-t_x\right)\left(\hat{t}_{y\pi}-t_y\right) \\
&\quad + \frac{1}{2}*0 + \frac{1}{2}\left(\frac{2t_y}{t_x^3}\right)\left(\hat{t}_{x\pi}-t_x\right)^2 \\
&= \theta + \frac{1}{t_x}\left(\hat{t}_{y\pi}-t_y\right) - \frac{t_y}{t_x^2}\left(\hat{t}_{x\pi}-t_x\right) - \left(\frac{1}{t_x^2}\right)\left(\hat{t}_{x\pi}-t_x\right)\left(\hat{t}_{y\pi}-t_y\right) \\
&\quad + \left(\frac{t_y}{t_x^3}\right)\left(\hat{t}_{x\pi}-t_x\right)^2.
\end{aligned}$$

Then the bias of $\hat{\theta}$ is approximated by

$$\begin{aligned}
\text{Bias} &= E\left(\hat{\theta}-\theta\right) \\
&\doteq -\frac{1}{t_x^2}cov_p\left(\hat{t}_{x\pi},\,\hat{t}_{y\pi}\right) + \frac{t_y}{t_x^3}var_p\left(\hat{t}_{x\pi}\right) \\
&= -\frac{t_y}{t_x}\frac{cov_p\left(\hat{t}_{x\pi},\,\hat{t}_{y\pi}\right)}{\sqrt{var_p\left(\hat{t}_{x\pi}\right)var_p\left(\hat{t}_{y\pi}\right)}}\frac{\sqrt{var_p\left(\hat{t}_{x\pi}\right)var_p\left(\hat{t}_{y\pi}\right)}}{t_xt_y} + \frac{t_y}{t_x}\frac{var_p\left(\hat{t}_{x\pi}\right)}{t_x^2} \\
&= \theta\left[\left(cv_p\left(\hat{t}_{x\pi}\right)\right)^2 - \rho_{\hat{t}_{x\pi},\hat{t}_{y\pi}}cv_p\left(\hat{t}_{x\pi}\right)cv_p\left(\hat{t}_{y\pi}\right)\right] \\
&= \theta\left[cv_p\left(\hat{t}_{x\pi}\right) - \rho_{\hat{t}_{x\pi},\hat{t}_{y\pi}}cv_p\left(\hat{t}_{y\pi}\right)\right]cv_p\left(\hat{t}_{x\pi}\right)
\end{aligned} \tag{1.12}$$

where $cv_p(\cdot)$ denotes the coefficient of variation under the design.

Remark 1.3.1 Under the model ξ : y_i are independent $(\beta x_i, \sigma^2 x_i)$, it is an easy task to show that the right hand side of (1.12) is approximately zero, but this is not the general case.

1.3.2 Godambe-Joshi Lower Bound for $\hat{\theta}$

Under the model design ξ : y_i are independent $(\beta x_i, \sigma_i^2)$, and for probability sampling design, $p(\cdot)$. If we had the entire finite population, then the least squares estimate for β is $\theta = t_y t_x^{-1}$. Let $\tilde{\theta}$ be any estimator of θ satisfying

$$E_\xi E_p\left(\tilde{\theta} - \theta\right) = 0.$$

Godambe-Joshi (1965) showed that

$$E_\xi E_p\left(\tilde{\theta} - \theta\right)^2 \geq \frac{1}{t_x^2} \sum_{i \in U} \left(\frac{1 - \pi_i}{\pi_i}\right) \sigma_i^2 = GJLB.$$

Assume that $\pi_i \geq \pi_{ij} \geq \pi_{N*} > 0$. Then

$$GJLB \leq \frac{1}{\bar{x}_{U_N}^2} \frac{1}{N \pi_{N*}} \left[\frac{1}{N} \sum_{i \in U} (1 - \pi_i) \sigma_i^2\right],$$

which is order $O\left((N\pi_{N*})^{-1}\right)$ under mild conditions. In particular, for SI of size n, $GJLB = O\left(n^{-1}\right)$.

Under the model ξ : $y_i = \beta x_i + \epsilon_i$, where ϵ_i are independent $(0, \sigma_i^2)$, the Godambe and Joshi (1965) lower bound is asymptotically attained by $\hat{\theta}$. To see this, note that $\beta - \theta = -\bar{x}_{U_N}^{-1} \bar{\epsilon}_U$, where $\bar{\epsilon}_U = N^{-1} \sum_{i \in U} \epsilon_i$, and so from equation (1.11), recall that

$$\begin{aligned} var_p\left(\hat{\theta}\right) &= \frac{1}{t_x^2} \sum_{i,j \in U} \frac{y_i - \theta x_i}{\pi_i} \frac{y_j - \theta x_j}{\pi_j} \Delta_{ij} \\ &= \frac{1}{t_x^2} \sum_{i,j \in U} \frac{\Delta_{ij}}{\pi_i \pi_j} \left(y_i - \beta x_i + \beta x_i - \theta x_i\right) \left(y_j - \beta x_j + \beta x_j - \theta x_j\right) \\ &= \frac{1}{t_x^2} \sum_{i,j \in U} \frac{\Delta_{ij}}{\pi_i \pi_j} \left(\epsilon_i + (\beta - \theta) x_i\right) \left(\epsilon_j + (\beta - \theta) x_j\right) \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{t_x^2}\sum_{i,j\in U}\frac{\Delta_{ij}}{\pi_i\pi_j}\epsilon_i\epsilon_j - \frac{2}{t_x^2}\sum_{i,j\in U}\frac{\Delta_{ij}}{\pi_i\pi_j}\frac{x_j}{\bar{x}_{U_N}}\epsilon_i\bar{\epsilon}_U + \frac{1}{t_x^2}\sum_{i,j\in U}\frac{\Delta_{ij}}{\pi_i\pi_j}\frac{x_ix_j}{\bar{x}_{U_N}^2}\bar{\epsilon}_U^2 \\
E_\xi\left[var_p\left(\hat{\theta}\right)\right] &= \frac{1}{t_x^2}\sum_{i\in U}\frac{1-\pi_i}{\pi_i}\sigma_i^2 - \frac{2}{t_x^2}\sum_{i,j\in U}\frac{\Delta_{ij}}{\pi_i\pi_j}\frac{x_j}{N\bar{x}_{U_N}}\sigma_i^2 \\
&\quad + \frac{1}{t_x^2}\sum_{i,j\in U}\frac{\Delta_{ij}}{\pi_i\pi_j}\frac{x_ix_j}{N^2\bar{x}_{U_N}^2}\sum_{l\in U}\sigma_l^2 \\
&= GJLB + O\left(\frac{1}{N^2\pi_{N*}}\right) + O\left(\frac{1}{N^2\pi_{N*}}\right) \\
&= GJLB + O\left(\frac{1}{N^2\pi_{N*}}\right) \qquad (1.13)
\end{aligned}
$$

Since

$$GJLB = \frac{1}{t_x^2}\sum_{i\in U}\frac{1-\pi_i}{\pi_i}\sigma_i^2 \quad \text{of order} \quad O\left(\frac{1}{N\pi_{N*}}\right)$$

then

$$\frac{E_\xi\left[AV_p\left(\hat{\theta}\right)\right]}{GJLB} = 1 + O\left(\frac{1}{N}\right) \to 1 \quad \text{as} \quad N \to \infty.$$

Hence, Godambe and Joshi (1965) lower bound is asymptotically attainable.

On the other hand, from equation (1.8), recall that

$$
\begin{aligned}
var_p\left(\hat{\theta}_{naive}\right) &= \frac{1}{t_x^2}\sum_{ij\in U}\frac{y_i}{\pi_i}\frac{y_j}{\pi_j}\Delta_{ij} \\
&= \frac{1}{t_x^2}\left\{\sum_{i\in U}\frac{1-\pi_i}{\pi_i}y_i^2 + \sum_{i\neq j:ij\in U}\frac{\Delta_{ij}}{\pi_i\pi_j}y_iy_j\right\} \\
E_\xi\left[var_p\left(\hat{\theta}_{naive}\right)\right] &= \frac{1}{t_x^2}\left\{\sum_{i\in U}\frac{1-\pi_i}{\pi_i}\left[\sigma_i^2 + \beta^2x_i^2\right] + \sum_{i\neq j:ij\in U}\frac{\Delta_{ij}}{\pi_i\pi_j}\beta^2x_ix_j\right\} \\
&= \frac{1}{t_x^2}\left\{\sum_{i\in U}\frac{1-\pi_i}{\pi_i}\sigma_i^2 + \sum_{ij\in U}\frac{\Delta_{ij}}{\pi_i\pi_j}\beta^2x_ix_j\right\} \\
&= GJLB + \frac{\beta^2}{t_x^2}\sum_{ij\in U}\frac{\Delta_{ij}}{\pi_i\pi_j}x_ix_j \\
&= GJLB + O\left(\frac{1}{N\pi_{N*}}\right) \qquad (1.14)
\end{aligned}
$$

under mild conditions, so the reminder term is of the same order as $GJLB$, and the lower bound is not attainable asymptotically.

1.4 Ratio Estimation of a Finite Population Total

Once we have an estimate of the population ratio, we can use this estimate to estimate the population total t_y under Aux_1. In other words, estimate t_y by $\breve{t}_y = t_x\breve{\theta}$, where $\breve{\theta}$ is an estimate for θ. If $\breve{\theta}$ is unbiased for θ, then $\breve{t}_y$ is unbiased for t_y.

Ratio estimation is often used in the case of stratified sampling in which the population of N units is divided into H non-overlapping strata U_h, with $N_h = |U_h|$ units in the h^{th} stratum. Under this scenario, the population total is

$$t_y = \sum_{h=1}^{H} t_{yh} = \sum_{h=1}^{H} t_{xh}\theta_h, \tag{1.15}$$

where $t_{yh} = \sum_{k\in U_h} y_k$, $t_{xh} = \sum_{k\in U_h} x_k$, $\theta_h = t_{yh}t_{xh}^{-1}$, and $U_h \subset U$ is the h^{th} stratum, consisting of N_h units.

Consider a measurable stratified sampling design, sampling n_h units from the h^{th} stratum. To estimate the population total using ratio estimation, there are two approaches. The first one is is to combine all strata, estimate the population ratio and multiply by t_x and the second approach is to estimate the h^{th} stratum ratio, multiply by t_{xh}, and sum across all strata.

Definition 1.4.1 *The combined ratio estimator* for the population total t_y is defined by

$$\hat{t}_{yComb,\hat{\theta}} = t_x\hat{\theta} \tag{1.16}$$

where $\tilde{\theta}$ is any estimator of θ. This estimator requires auxiliary information Aux_1.

Definition 1.4.2 *The separate ratio estimator* for the population total t_y is defined by

$$\hat{t}_{ySep,\hat{\theta}} = \sum_{h=1}^{H} t_{xh}\hat{\theta}_h \tag{1.17}$$

where $\tilde{\theta}_h$ is any ratio estimator for the h^{th} stratum ratio. This estimator requires auxiliary information Aux_2.

For example, using $\hat{\theta}$ as an estimator for the population ratio θ in equation (1.16), we have the *simple combined ratio estimator*

$$\hat{t}_{ycom,\hat{\theta}} = t_x \hat{\theta} \tag{1.18}$$

under Aux_2. From equation (1.12), it follows that

$$\begin{aligned} Bias_p\left(\hat{t}_{ycom,\hat{\theta}}\right) &= E_p\left(\hat{t}_{ycom,\hat{\theta}}\right) - t_x\theta \\ &= t_x\left[E_p\left(\hat{\theta}\right) - \theta\right] \\ &= t_y\left[cv_p\left(\hat{t}_{x\pi}\right) - \rho_{\hat{t}_{x\pi},\hat{t}_{y\pi}} cv_p\left(\hat{t}_{y\pi}\right)\right] cv_p\left(\hat{t}_{x\pi}\right) \end{aligned}$$

and from equation (1.11),

$$var_p\left(\hat{t}_{ycom,\hat{\theta}}\right) \doteq \sum_{ij\in U} \frac{y_i - \theta x_i}{\pi_i} \frac{y_j - \theta x_j}{\pi_j} \Delta_{ij}. \tag{1.19}$$

Further, under the super-population model

$$\xi : y_i = \beta x_i + \epsilon_i, \text{ where } \epsilon_i \text{ are independent } \left(0, \sigma_i^2\right),$$

we have from earlier discussion that the Godambe-Joshi lower bound is asymptotically attainable for $\hat{t}_{ycom,\hat{\theta}}$.

As another example, using $\hat{\theta}$ as an estimator for the population ratio θ in equation (1.17), we have the *simple separate ratio estimator*

$$\hat{t}_{ysep,\hat{\theta}} = \sum_{h=1}^{H} t_{xh}\hat{\theta}_h$$

under Aux_2.

Applying equation (1.12) in each stratum, we have

$$\begin{aligned} Bias_p\left(\hat{t}_{ysep,\hat{\theta}}\right) &= E_p\left(\hat{t}_{ysep,\hat{\theta}}\right) - \sum_{h=1}^{H} t_{xh}\theta_h \\ &= \sum_{h=1}^{H} t_{xh}\left[E_p\left(\hat{\theta}_h\right) - \theta_h\right] \\ &= \sum_{h=1}^{H} t_{xh} Bias\left(\hat{\theta}_h\right) \end{aligned} \tag{1.20}$$

where

$$Bias\left(\hat{\theta}_h\right) = t_{yh}\left[cv_p\left(\hat{t}_{x\pi,h}\right) - \rho_{\hat{t}_{x\pi,h},\hat{t}_{y\pi,h}} cv_p\left(\hat{t}_{y\pi,h}\right)\right] cv_p\left(\hat{t}_{x\pi,h}\right)$$

and, using the independence across strata and equation (1.11) within strata,

$$var_p\left(\hat{t}_{ySep,\hat{\theta}}\right) \doteq \sum_{h=1}^{H} \sum_{ij \in U_h} \frac{y_i - \theta_h x_i}{\pi_i} \frac{y_j - \theta_h x_j}{\pi_j} \Delta_{ij}. \tag{1.21}$$

Remark 1.4.1 The use of stratification comes from our belief that there are differences between strata and homogeneity within strata. To use the combined ratio estimator to estimate the population total t_y will ignore much of the efficiency afforded by stratification. The variance in (1.21) is small in the case of a common ratio for all strata, so the simple common ratio estimator will work well only if the ratios do not vary much from stratum to stratum. When we have big differences in ratios from stratum to stratum, then the simple separate ratio estimator will be a better estimator of the population total t_y as can be seen from its approximate variance in equation (1.21).

On the other hand, if $\hat{t}_{ySep,\hat{\theta}}$ is used to estimate the population total t_y, then as shown by equation (1.20), the biases of the within-stratum ratio estimates may accumulate across strata, leading to poor performance. To overcome this problem, it is useful to have an exactly unbiased estimator for the within stratum ratios. Therefore, in this work, we will propose an exactly unbiased estimator for θ_h, which can be used for unbiased, efficient estimates of the within-stratum totals t_{yh}.

1.5 Ratio Estimators with Reduced Bias

Ratio estimation has been studied in the literature for more than fifty years. Most of the discussions are under the simple random sampling design, and only approximate variances of the corresponding estimators are given. Under a general

sampling design, the population ratio estimators are typically biased, though almost unbiased, and asymptotically unbiased. To eliminate or reduce the bias of ratio estimation, authors have either modified the estimator, or have modified the sampling design.

In particular, Lahiri (1951) proposed a sampling method that is an example of a rejective method. Start this method by choosing a number $M \geq \max(x_1, \ldots, x_N)$. With equal probability draw one of the N population elements. Let η be the selected element. Draw u from a Uniform $(0, 1)$. If $uM \leq x_\eta$, then the selected element is included in the sample; otherwise, start over. This method gives $\pi_k = x_k / \sum_{k \in U} x_k$, for $k = 1, \ldots, N$, a probability proportional to size sampling design with $n = 1$, even though $\sum_U x_k$ need not be known. Note that under this design, $\sum_{k \in s} (y_k / x_k)$ has expectation $\sum_{k \in U} (y_k / x_k) \left(x_k / \sum_{k \in U} x_k \right) = \sum_{k \in U} y_k / \sum_{k \in U} x_k$, the population ratio. That is, this particular combination of design and estimator gives an exactly unbiased estimator of the population ratio under Aux_0.

Mickey (1959) derived an estimator under simple random sampling without replacement of size n. Compute $\tilde{\theta}_{i-}$ by removing each unit i in turn from the sample, so that $\tilde{\theta}_{i-} = \frac{\sum y_k}{\sum x_k}$ is computed over the remaining $n-1$ members. Then the Mickey estimator is given by:

$$\hat{\theta}_M = \bar{\tilde{\theta}}_- + \frac{n(N-n+1)}{N \bar{x}_{U_N}} \left(\bar{y}_s - \bar{\tilde{\theta}}_- \bar{x}_s \right) \tag{1.22}$$

where $\bar{\tilde{\theta}}_-$ is the mean of n ratios $\tilde{\theta}_{i-}$, and $\bar{y}_s = n^{-1} \sum_s y_k$, $\bar{x}_s = n^{-1} \sum_s x_k$, and $\bar{x}_{U_N} = N^{-1} \sum_U x_k$. Mickey's estimator is an unbiased estimator for the population ratio θ under Aux_1.

Nieto de Pascual (1961) proposed an almost unbiased estimator, in which the bias is of order n^{-2}. This estimator is also under simple random sampling and Aux_1, and is given by:

$$\hat{\theta}_P = \frac{\bar{y}_s}{\bar{x}_s} + \frac{1}{(n-1)\bar{x}_{U_N}} (\bar{y}_s - \bar{r}_s \bar{x}_s) \tag{1.23}$$

where $r_k = y_k/x_k$ and $\bar{r}_s = n^{-1}\sum_s r_k$.

Murthy and Nanjamma (1959) proposed the following estimator,

$$\hat{\theta}_{MN} = \bar{r}_s + \frac{n}{(n-1)\,\bar{x}_s}\left(\bar{y}_s - \bar{r}_s\bar{x}_s\right) \tag{1.24}$$

under simple random sampling without replacement. This is an almost unbiased estimator for the population ratio under Aux_1.

Our work is directly motivated by that of Hartley and Ross (1954), who derived an exactly unbiased estimator for the population ratio under simple random sampling and Aux_1. Their estimator is given by

$$\hat{\theta}_{HR} = \bar{r}_s + \frac{n\,(N-1)}{N\,(n-1)\,\bar{x}_{U_N}}\left(\bar{y}_s - \bar{r}_s\bar{x}_s\right). \tag{1.25}$$

Hartley and Ross (1954) gave the asymptotic variance of this estimator as

$$var_p\left(\hat{\theta}_{HR}\right) = \frac{1}{n}\left(\frac{\bar{y}_U}{\bar{x}_{U_N}}\right)^2\left[\frac{\sigma_y^2}{\bar{y}_U^2} + \frac{\sigma_x}{\bar{x}_{U_N}^2} - 2\frac{cov_p\,(x,\,y)}{\bar{y}_U\bar{x}_{U_N}}\right]. \tag{1.26}$$

In this work, we will propose an exactly unbiased estimator under a general probability sampling design, which gives the Hartley and Ross (1954) estimator as a special case under SI. Furthermore, we will give an exact expression for the variance and an exactly unbiased estimator for the variance. Various extensions will be considered, including the stratified sampling case, in which the bias of ordinary ratio-type estimators is a serious issue.

1.6 Contributions of This Dissertation

We will give a brief description for the coming chapters. In Chapter 2, we will introduce the Generalized Hartley-Ross estimator, $\hat{\theta}_{GHR}$. We will show that this estimator is exactly unbiased under a general sampling design and that it gives the Hartley-Ross (1954) estimator under the special case of simple random sampling design without replacement (SI). As examples, we will investigate the properties of

$\hat{\theta}_{GHR}$ under SI, random sampling with replacement (WR), Poisson sampling (PO), and stratified sampling designs. We derive the exact variance of $\hat{\theta}_{GHR}$ and an exactly unbiased variance estimator under a fourth-order measurable sampling design. We will give the first through fourth-order inclusion probabilities for different sampling designs, which are needed to compute the variance and the unbiased variance estimator for $\hat{\theta}_{GHR}$. To find the first through fourth-order inclusion probabilities for a general sampling design is not an easy task in general; therefore, we will give two methods to approximate the computations of the unbiased estimator for the variance of $\hat{\theta}_{GHR}$. The two methods of approximations will be tested via simulations under SI, proportional to size sampling design (πps), and PO sampling designs. Furthermore, $\hat{\theta}_{GHR}$ will be written under stratified simple random without replacement sampling design (STSI) to estimate the population total t_y. Also, $\hat{\theta}_{GHR}$ is written under interpenetrating sub-samples, and under interpenetrating STSI by sampling one element from each strata via SI.

Under a stratified sampling design and Aux_2, $\hat{\theta}_{GHR}$ can be used in a separate ratio estimator to estimate the population ratio θ or to estimate the population total t_y.

An alternative unbiased ratio estimator $\tilde{\theta}_{GHR}$ under a with-replacement design using a Hansen and Hurwitz (1943) type estimator will be introduced. This estimator will be used to estimate the population ratio θ and the population total t_y. An exact variance and an unbiased estimator of the variance of this estimator will be given.

At the end of Chapter 2, we will introduce another estimator as a result of a linear combination between $\hat{\theta}_{GHR}$ and $\hat{\theta}$. The GHR estimator is unbiased but may have large variance, while the simple estimator is biased but has small variance. The goal of the combination is to obtain an estimator with variance less than the variance of $\hat{\theta}_{GHR}$. However, the new estimator is no longer unbiased (except in the trivial

special case when only $\hat{\theta}_{GHR}$ is included). From this combination and under the SI sampling design, we can produce Murthy and Nanjamma (1959), Nieto de Pascual (1961), Hartley and Ross (1954), and simple estimators.

In Chapter 3, asymptotic results involving the two unbiased estimators $\hat{\theta}_{GHR}$ and $\tilde{\theta}_{GHR}$ will be discussed. Central limit theorems (CLTs) for $\hat{\theta}_{GHR}$ will be discussed under SI and PO sampling designs. Further, a CLT for $\hat{t}_{ySep,GHR}$ under a general stratified sampling design will be established. In addition, CLT based on $\tilde{\theta}_{GHR}$ will be established; both for the case that the population size N is fixed and the number of independent draws tends to infinity, and for the case when both the population size and the number of independent draws tends to infinity. A consistent estimator of the asymptotic variance under the second case will be given. Godambe and Joshi (1965) lower bound will be discussed for $\hat{\theta}_{GHR}$, and $\tilde{\theta}_{GHR}$. At the end of the Chapter 3, $\hat{\theta}_{GHR}$ and $\hat{\theta}$ are compared through simulation and results are given for unstratified and stratified sampling designs. Chapter 4 includes some concluding discussion and an appendix assembles some technical details.

Chapter 2

GENERALIZED UNBIASED ESTIMATION OF RATIOS AND RATIO ESTIMATION

The need to find an exactly unbiased estimator for the population ratio is a serious issue especially in stratified sampling. In this chapter we will introduce an exactly unbiased estimator for the population ratio and derive its characteristics.

2.1 General Measurable Designs

The population ratio θ, is a non-linear function of two totals, the total of the study variable t_y and the total of the auxiliary variable t_x. In other words,

$$\theta = f(t_y, t_x) = \frac{t_y}{t_x} = \frac{\sum_{i \in U} y_i}{\sum_{i \in U} x_i}.$$

We will assume Aux_1 : the auxiliary values $x_i > 0$ are available for all sampled elements $i \in s$, and $\bar{x}_{U_N}$ is also available from some source external to the sample.

A probability sample s is drawn from a finite population U according to a measurable sampling design $p(\cdot)$; for this general probability sampling design, our goal is to obtain an exactly unbiased estimator for θ.

2.1.1 Generalized Hartley and Ross Estimator

We will generalize Hartley-Ross estimator under a measurable sampling design.

Theorem 2.1.1 *Under a measurable sampling design and Aux_1, the estimator*

$$\hat{\theta}_{GHR} = \frac{1}{N} \sum_{i \in s} r_i \frac{1}{\pi_i} + \frac{1}{N \bar{x}_{U_N}} \left[\sum_{i \in s} \frac{y_i}{\pi_i} - \frac{1}{N} \sum_{i \in s} \sum_{j \in s} r_i x_j \frac{1}{\pi_{ij}} \right], \tag{2.1}$$

where $r_i = y_i x_i^{-1}$, *is an unbiased estimator for* θ, *and* $\hat{t}_{y,GHR} = t_x \hat{\theta}_{GHR}$ *is an unbiased estimator for* t_y.

Proof: Rewrite $\hat{\theta}_{GHR}$ as

$$\hat{\theta}_{GHR} = \frac{1}{N}\hat{t}_{r\pi} + \frac{1}{t_x}\hat{t}_{y\pi} - \frac{1}{N^2 \bar{x}_{U_N}} \sum_{i \in U} \sum_{j \in U} r_i x_j \frac{I_{\{i,j \in s\}}}{\pi_{ij}} \tag{2.2}$$

where $\hat{t}_{r\pi} = \sum_{i \in U} r_i \frac{I_{\{i \in s\}}}{\pi_i}$. Therefore,

$$\begin{aligned} \mathrm{E}_p\left[\hat{\theta}_{GHR}\right] &= \frac{1}{N} t_r + \frac{t_y}{t_x} - \frac{1}{N t_x} \sum_{i \in U} \sum_{j \in U} r_i x_j \\ &= \frac{1}{N} t_r + \frac{t_y}{t_x} - \frac{1}{t_x}\left[\frac{1}{N} t_r\right] t_x \\ &= \theta. \end{aligned} \tag{2.3}$$

Hence, $\hat{\theta}_{GHR}$ is an unbiased estimator for θ, and $\hat{t}_{y,GHR}$ is an unbiased estimator for t_y.

■

In the following examples we will write $\hat{\theta}_{GHR}$ under different sampling designs.

Example 2.1.1 *Simple random sampling without replacement design (SI).*

Under SI design,

$$\pi_i = \frac{n}{N} \tag{2.4}$$

and

$$\pi_{ij} = \frac{n(n-1)}{N(N-1)} \quad \forall i \neq j. \tag{2.5}$$

Define $S_{D_t} = \{\text{all distinct } t\text{-tuples from } s\}$. From (2.1), we have

$$\begin{aligned}
\hat{\theta}_{GHR} &= \bar{r}_s + \frac{1}{N\bar{x}_{U_N}}\left[N\bar{y}_s - \frac{N-1}{n(n-1)}\sum_{ij\in s_{D_2}} r_i x_j - \bar{y}_s\right] \\
&= \bar{r}_s + \frac{1}{N\bar{x}_{U_N}}\left[N\bar{y}_s - \frac{N-1}{n(n-1)}\left\{\sum_{i,j\in s} r_i x_j - \sum_{i\in s} y_i\right\} - \bar{y}_s\right] \\
&= \bar{r}_s + \frac{1}{N\bar{x}_{U_N}}\left[\frac{n(N-1)}{n-1}\bar{y}_s - \frac{n(N-1)}{n-1}\bar{r}_s\bar{x}_s\right] \\
&= \bar{r}_s + \frac{n(N-1)}{N(n-1)\bar{x}_{U_N}}\left(\bar{y}_s - \bar{r}_s\bar{x}_s\right), \qquad (2.6)
\end{aligned}$$

which is the Hartley and Ross (1954) estimator.

Example 2.1.2 *With replacement sampling design (WR).*

In with-replacement sampling (WR), the selection is carried out by drawing a first element in such a way that

$$Pr\left(\text{selecting element } k\right) = p_k; \quad k = 1, \ldots, N$$

where $\sum_U p_k = 1$. The selected element is replaced and the second element is independently selected, continuing this process to the m^{th} element.

For m independent draws, the first-order inclusion probability is

$$\pi_k = 1 - (1 - p_k)^m. \qquad (2.7)$$

The second-order inclusion probability π_{kl}, for $k, l \in S_{D_2}$ is given by

$$\begin{aligned}
\pi_{kl} &= Pr\left(k, l \in s_{D_2}\right) \\
&= 1 - Pr\left(k \notin s \text{ or } l \notin s\right) \\
&= 1 - \left[Pr\left(k \notin s\right) + Pr\left(l \notin s\right) - Pr\left(k \text{ and } l \notin s\right)\right] \\
&= 1 - \left[(1-p_k)^m + (1-p_l)^m - (1-(p_k+p_l))^m\right] \qquad (2.8)
\end{aligned}$$

Expressions (2.7) and (2.8) can be plugged in to equation (2.1) to yield $\hat{\theta}_{GHR}$ under WR.

Example 2.1.3 *Poisson sampling design (PO).*

Let $\pi_k = Pr\left(I_{\{k\in s\}} = 1\right)$ be the predetermined positive inclusion probability for the k^{th} element in the population, where $k = 1, \ldots, N$. The sampling membership indicators $I_{\{k\in s\}}$ are independent; therefore, $\pi_{kl} = \pi_k \pi_l$ for any $k \neq l$. From equation (2.1), rewrite $\hat{\theta}_{GHR}$ under PO as

$$\hat{\theta}_{GHR} \overset{PO}{=} \frac{1}{N}\sum_{i\in s}\frac{r_i}{\pi_i} + \frac{1}{N^2\bar{x}_{U_N}}\sum_{i\in s}\left[N - 1 + \frac{1}{\pi_i}\right]\frac{y_i}{\pi_i} - \frac{1}{N^2\bar{x}_{U_N}}\left[\sum_{i\in s} r_i\frac{1}{\pi_i} \times \sum_{j\in s} x_j\frac{1}{\pi_j}\right]$$

Example 2.1.4 *Stratified sampling design.*

Under a general stratified sampling design, $\hat{\theta}_{GHR}$ from equation (2.1) is

$$\begin{aligned}\hat{\theta}_{GHR} &= \frac{1}{N}\sum_{h=1}^{H}\sum_{i\in s_h}\frac{r_i}{\pi_i} + \frac{1}{N\bar{x}_{U_N}}\sum_{h=1}^{H}\sum_{i\in s_h}\frac{y_i}{\pi_i} - \frac{1}{N^2\bar{x}_{U_N}}\sum_{h=1}^{H}\sum_{i,j\in s_h}\frac{r_i}{\pi_{ij}}x_j \\ &\quad - \frac{1}{N^2\bar{x}_{U_N}}\sum_{h,\acute{h}\in D_2}\sum_{i\in s_h}\sum_{j\in s_{\acute{h}}}\frac{r_i}{\pi_{ij}}x_j.\end{aligned}$$

In particular, under stratified simple random sampling without replacement (STSI), we have

$$\pi_i = \frac{n_h}{N_h} \tag{2.9}$$

and

$$\pi_{ij} = \begin{cases} \frac{n_h}{N_h} & \text{for } i = j \text{ and } i, j \in U_h \\ \frac{n_h}{N_h}\frac{n_h - 1}{N_h - 1} & \text{for } i \neq j \text{ and } ij \in U_h \\ \frac{n_h}{N_h}\frac{n_{\acute{h}}}{N_{\acute{h}}} & \text{for } i \in U_h,\ j \in U_{\acute{h}}, \text{ and } h, \acute{h} \in D_2. \end{cases} \tag{2.10}$$

Hence,

$$\begin{aligned}\hat{\theta}_{GHR} \overset{STSI}{=} &\sum_{h=1}^{H} w_h\left\{\bar{r}_h + \frac{1}{N\bar{x}_{U_N}}\left(\frac{N_h - n_h}{n_h - 1}\right)(\bar{y}_h - \bar{r}_h\bar{x}_h)\right\} + \frac{1}{\bar{x}_{U_N}}\sum_{h=1}^{H} w_h\bar{y}_h \\ &- \frac{1}{\bar{x}_{U_N}}\left(\sum_h w_h\bar{r}_h\right)\left(\sum_{\acute{h}} w_{\acute{h}}\bar{x}_{\acute{h}}\right).\end{aligned} \tag{2.11}$$

where $\bar{r}_h = n_h^{-1} \sum_{i \in S_h} y_i x_i^{-1}$ and $w_h = N_h N^{-1}$.

Example 2.1.5 $\hat{\theta}_{GHR}$ *under interpenetrating sub-samples.*

The idea of interpenetrating sub-samples is to draw a sub-sample of size m via SI, then return the m observations into the population and independently repeat this process G times; in this case, the sample size is $n \leq mG$. The first order-inclusion probability is then

$$\begin{aligned}
\pi_i &= Pr\left(i \in s\right) \\
&= 1 - Pr\left(i \notin s\right) \\
&= 1 - Pr\left(i \in \bigcap_g^G s_g^c\right) \\
&= 1 - \left[Pr\left(i \in s_1^c\right)\right]^G \\
&= 1 - \left[1 - \frac{m}{N}\right]^G
\end{aligned} \tag{2.12}$$

The second order-inclusion probability $(i, j \in D_2)$ is

$$\begin{aligned}
\pi_{ij} &= Pr\left(i, j \in s\right) \\
&= 1 - Pr\left(i \notin s \text{ or } j \notin s\right) \\
&= 1 - \left[Pr\left(i \notin s\right) + Pr\left(j \notin s\right) - Pr\left(i \notin s \text{ and } j \notin s\right)\right] \\
&= 1 - \left[2Pr\left(i \notin s\right) - Pr\left(i \notin s \text{ and } j \notin s\right)\right] \\
&= 1 - \left[2Pr\left(i \in \bigcap_g^G s_g^c\right) - Pr\left(i, j \in \bigcap_g^G s_g^c\right)\right] \\
&= 1 - \left[2\left(Pr\left(i \in s_1^c\right)\right)^G - \left(Pr\left(i, j \in s_1^c\right)\right)^G\right] \\
&= 1 - \left\{2\left[1 - \frac{m}{N}\right]^G - \left[1 - \left(2\frac{m}{N} - \frac{m\left(m-1\right)}{N\left(N-1\right)}\right)\right]^G\right\}
\end{aligned} \tag{2.13}$$

Expressions (2.12) and (2.13) can be plugged in to equation (2.1) to yield $\hat{\theta}_{GHR}$ under this design.

Remark 2.1.1 For small p, expand $(1-p)^m$ by first-order Taylor expansion to get

$$(1-p)^m = 1 - mp.$$

We use this fact to approximate the first-order and second-order inclusion probabilities. From equation (2.12), we have

$$\pi_i = 1 - \left[1 - \frac{m}{N}\right]^G \doteq 1 - \left(1 - G\frac{m}{N}\right) = G\frac{m}{N}$$

and from equation (2.13), we have

$$\pi_{ij} \doteq 1 - \left\{2\left(1 - G\frac{m}{N}\right) - \left(1 - G\left(2\frac{m}{N} - \frac{m\,(m-1)}{N\,(N-1)}\right)\right)\right\} = G\frac{m}{N}\frac{m-1}{N-1}.$$

Therefore $\hat{\theta}_{GHR}$ under this sampling design is

$$\begin{aligned}\hat{\theta}_{GHR} &\doteq \frac{1}{G}\sum_{g=1}^{G}\left\{\left(\frac{1}{m}\sum_{i\in s_g} r_i\right) + \frac{m\,(N-1)}{N\,(m-1)\,\bar{x}_{U_N}}\left[\frac{1}{m}\sum_{i\in s_g} y_i\right.\right. \\ &\quad \left.\left. - \left(\frac{1}{m}\sum_{i\in s_g} r_i\right)\left(\frac{1}{m}\sum_{j\in s_g} x_j\right)\right]\right\} \\ &= \frac{1}{G}\sum_{g=1}^{G}\left\{\bar{r}_g + \frac{m\,(N-1)}{N\,(m-1)\,\bar{x}_{U_N}}\left[\bar{y}_g - \bar{x}_g\bar{r}_g\right]\right\}, \end{aligned} \tag{2.14}$$

which is the average of $\hat{\theta}_{GHR}$ under SI over G repetitions.

2.1.2 Exact Variance of $\hat{\theta}_{GHR}$

One of the interesting properties of $\hat{\theta}_{GHR}$ is an exact variance expression under any measurable sampling design.

Theorem 2.1.2 *Assume $x_i > 0$ for all $i \in U$. For a measurable sampling design, the variance of $\hat{\theta}_{GHR}$ is given by*

$$\begin{aligned} var_p\left(\hat{\theta}_{GHR}\right) &= \frac{1}{N^2}\sum_{i,j\in U}\frac{y_i^*}{\pi_i}\frac{y_j^*}{\pi_j}\Delta_{ij} + \frac{1}{N^4\bar{x}_{U_N}^2}\sum_{i,j,k,l\in U}\frac{r_i x_j}{\pi_{ij}}\frac{r_k x_l}{\pi_{kl}}\Delta_{ijkl} \\ &\quad - \frac{2}{N^3\bar{x}_{U_N}}\sum_{i,k,l\in U}\frac{y_i^*}{\pi_i}\frac{r_k x_l}{\pi_{kl}}\Delta_{ikl} \end{aligned} \tag{2.15}$$

where

$$\begin{aligned} y_i^* &= \left(\frac{1}{x_i} + \frac{1}{\bar{x}_{U_N}}\right) y_i \\ \Delta_{ij} &= cov_p\left(I_{\{i\in s\}},\, I_{\{j\in s\}}\right) = \pi_{ij} - \pi_i\pi_j \\ \Delta_{ikl} &= cov_p\left(I_{\{i\in s\}},\, I_{\{kl\in s\}}\right) = \pi_{ikl} - \pi_i\pi_{kl} \\ \Delta_{ijkl} &= cov_p\left(I_{\{ij\in s\}},\, I_{\{kl\in s\}}\right) = \pi_{ijkl} - \pi_{ij}\pi_{kl}. \end{aligned}$$

Further, the variance of $\hat{t}_{y,GHR}$ *is given by*

$$var_p\left(\hat{t}_{y,GHR}\right) = t_x^2 var_p\left(\hat{\theta}_{GHR}\right).$$

Proof: From equation (2.1), rewrite $\hat{\theta}_{GHR}$ as

$$\hat{\theta}_{GHR} = \frac{1}{N}\sum_{i\in U}\frac{y_i^*}{\pi_i} I_{\{i\in s\}} - \frac{1}{N^2\bar{x}_{U_N}}\sum_{i\in U}\sum_{j\in U}\frac{r_i x_j}{\pi_{ij}} I_{\{i,j\in s\}}$$

The theorem follows directly by taking the variance of both sides of this equation.

■

In the following examples, we will write the exact variance of $\hat{\theta}_{GHR}$. We need the following notation:

$$U_{D_t} = \{\text{Set of all distinct } t\text{-tuples } (i_1, i_2, \ldots, i_t) \text{ from } U\}.$$

Example 2.1.6 *Variance of* $\hat{\theta}_{GHR}$ *under SI sampling design.*

For a population of size $N > 3$ and under SI, the third-order inclusion probability for $i,\, j,\, k \in U_{D_3}$ is

$$\pi_{ijk} = \frac{n}{N}\frac{n-1}{N-1}\frac{n-2}{N-2} \tag{2.16}$$

and the fourth-order inclusion probability for $i,\, j,\, k,\, l \in U_{D_4}$ is

$$\pi_{ijkl} = \frac{n}{N}\frac{n-1}{N-1}\frac{n-2}{N-2}\frac{n-3}{N-3}. \tag{2.17}$$

The inclusion probabilities (2.16) and (2.17), along with (2.4) and (2.5), can then be plugged in to Theorem 2.1.2 to yield on exact variance. Note that Hartley and Ross (1954) provided only a variance approximation.

Example 2.1.7 *Variance of $\hat{\theta}_{GHR}$ under WR sampling design.*

For $i, j, k \in U_{D_3}$, and $i, j, k, l \in U_{D_4}$ it is possible to derive the third-order and fourth-order inclusion probabilities, using argument like those in Example 2.1.2 resulting in

$$\begin{aligned}\pi_{ijk} &= 1-[(1-p_i)^m+(1-p_j)^m+(1-p_k)^m-(1-(p_i+p_j))^m \\ &\quad -(1-(p_i+p_k))^m-(1-(p_j+p_k))^m+(1-(p_i+p_j+p_k))^m] \quad (2.18)\end{aligned}$$

and for $ijkl \in D_4$,

$$\begin{aligned}\pi_{ijkl} &= 1-[(1-p_i)^m+(1-p_j)^m+(1-p_k)^m+(1-p_l)^m-(1-(p_i+p_j))^m \\ &- (1-(p_i+p_k))^m-(1-(p_i+p_l))^m-(1-(p_j+p_k))^m-(1-(p_j+p_l))^m \\ &- (1-(p_k+p_l))^m+(1-(p_i+p_j+p_k))^m+(1-(p_i+p_j+p_l))^m \\ &+ (1-(p_i+p_k+p_l))^m+(1-(p_j+p_k+p_l))^m \\ &- (1-(p_i+p_j+p_k+p_l))^m].\end{aligned}$$

Alternatively, these higher-order inclusion probabilities can be approximated by zero using the second order Taylor expansion. In either cases, Theorem 2.1.2 can be used to compute the variance of $\hat{\theta}_{GHR}$ under this design, either exactly or approximately.

Example 2.1.8 *Variance of $\hat{\theta}_{GHR}$ under PO sampling design.*

The independence of sampling membership indicators $I_{\{k\in s\}}$ enables us to define easily the third-order and the fourth-order inclusion probabilities. For $i, j, k \in U_{D_3}$, the third-order inclusion probability for PO is

$$\pi_{ijk} = \pi_i\pi_j\pi_k$$

and for $i, j, k, l \in U_{D_4}$, the fourth-order inclusion probability is

$$\pi_{ijkl} = \pi_i\pi_j\pi_k\pi_l.$$

With these inclusion probabilities, Theorem 2.1.2 can be used to yield an exact variance of $\hat{\theta}_{GHR}$.

Example 2.1.9 *Variance of $\hat{\theta}_{GHR}$ under stratified sampling design.*

Under a general stratified sampling design,

$$var_p\left(\hat{\theta}_{GHR}\right) = \frac{1}{N^2}\sum_{h=1}^{H}\sum_{i,j\in U_h}\frac{y_i^*}{\pi_i}\frac{y_j^*}{\pi_j}\Delta_{ij} + \frac{1}{N^4\bar{x}_{U_N}^2}\sum_{h=1}^{H}\sum_{i,j,k,l\in U_h}\frac{r_i x_j}{\pi_{ij}}\frac{r_k x_l}{\pi_{kl}}\Delta_{ijkl} - \frac{2}{N^3\bar{x}_{U_N}}\sum_{h=1}^{H}\sum_{i,k,l\in U_h}\frac{y_i^*}{\pi_i}\frac{r_k x_l}{\pi_{kl}}\Delta_{ikl}, \tag{2.19}$$

by Theorem 2.1.2.

In particular, under STSI, for $i,\ j,\ k \in U_{D_3}$, and $i,\ j,\ k,\ l \in U_{D_4}$ it is possible to derive the third-order and fourth-order inclusion probabilities, using arguments like those in Example 2.1.4, equation (2.19) can then be used to compute the variance of $\hat{\theta}_{GHR}$.

2.1.3 Unbiased Variance Estimation

The existence of an exactly unbiased estimator of $var_p\left(\hat{\theta}_{GHR}\right)$ is another useful result for $\hat{\theta}_{GHR}$.

Theorem 2.1.3 *For a fourth-order measurable sampling design, an unbiased estimator for $var_p\left(\hat{\theta}_{GHR}\right)$ is given by*

$$\hat{var}_p\left(\hat{\theta}_{GHR}\right) = \frac{1}{N^2}\sum_{i,j\in s}\frac{y_i^*}{\pi_i}\frac{y_j^*}{\pi_j}\frac{\Delta_{ij}}{\pi_{ij}} + \frac{1}{N^4\bar{x}_{U_N}^2}\sum_{i,j,k,l\in s}\frac{r_i x_j}{\pi_{ij}}\frac{r_k x_l}{\pi_{kl}}\frac{\Delta_{ijkl}}{\pi_{ijkl}} - \frac{2}{N^3\bar{x}_{U_N}}\sum_{i,k,l\in s}\frac{y_i^*}{\pi_i}\frac{r_k x_l}{\pi_{kl}}\frac{\Delta_{ikl}}{\pi_{ikl}}. \tag{2.20}$$

Further, an unbiased estimator for $var_p\left(\hat{t}_{y,GHR}\right)$ is given by

$$\hat{var}_p\left(\hat{t}_{y,GHR}\right) = t_x^2\hat{var}_p\left(\hat{\theta}_{GHR}\right).$$

Proof: The proof of this theorem follows directly from Theorem 2.1.2, using fourth-order measurability to ensure that (2.20) is well-defined.

■

Though an exactly unbiased estimator for $\hat{var}_p\left(\hat{\theta}_{GHR}\right)$ exists if the conditions of Theorem 2.1.3 are satisfied, it is in general difficult to obtain the third-order and the fourth-order inclusion probabilities under general sampling designs. It is an easy task, however, to obtain these higher-order inclusion probabilities under SI (with $n \geq 4$ for fourth-order measurability) and PO sampling designs. For Brewer's method a special case of πps with $n = 2$, the third-order and fourth-order inclusion probabilities are zero when at least three of the indices are distinct. For $n > 4$, other πps methods exist as implemented, for example, in SAS `proc surveyselect`. This procedure will produce first and second -order inclusion probabilities, but not higher-order, a common limitation. It will be therefore be useful to consider approximate variance estimators that do not require higher-order inclusion probabilities. We begin however, with the case when the first through fourth-order inclusion probabilities are available. Let us return to our earlier examples.

Example 2.1.10 $\hat{var}_p\left(\hat{\theta}_{GHR}\right)$ *under SI design*

Let us rewrite the first through fourth-order inclusion probabilities under SI and for $N > 3$.

The first-order inclusion probability is $\pi_i = \frac{n}{N}$,

the second-order inclusion probability is $\pi_{ij} = \frac{n}{N}\frac{n-1}{N-1}$, for $i, j \in s_{D_2}$,

the third-order inclusion probability is $\pi_{ijk} = \frac{n}{N}\frac{n-1}{N-1}\frac{n-2}{N-2}$, for $i, j, k \in s_{D_3}$,

the fourth-order inclusion probability is $\pi_{ijkl} = \frac{n}{N}\frac{n-1}{N-1}\frac{n-2}{N-2}\frac{n-3}{N-3}$, for $i, j, k.l \in s_{D_4}$,

where s_{D_t} is the set of all distinct t-tuples $(i_1, i_2, \ldots, i_t)$ from s. Therefore, the computations of $\hat{var}_p\left(\hat{\theta}_{GHR}\right)$ in Theorem 2.1.3 can be done to yield an unbiased variance estimator.

Remark 2.1.2 As previously noted, $\hat{\theta}_{GHR}$ is exactly the Hartley and Ross (1954) estimator under SI. But what is new and not given by Hartley and Ross (1954)

are the exact variance of $\hat{\theta}_{GHR}$ and an exactly unbiased estimator for the variance. Hartley and Ross (1954) gave only an asymptotic result for the variance.

Example 2.1.11 $\hat{var}_p\left(\hat{\theta}_{GHR}\right)$ *under WR design.*

Under WR sampling design, the first through fourth-order inclusion probabilities can be computed (as given in Examples 2.1.2 and 2.1.7). Note that even a second-order Taylor approximation for such inclusion probabilities would not be sufficiently precise since the approximation of the third-order and fourth-order inclusion probabilities would be zero in this case.

Example 2.1.12 $\hat{var}_p\left(\hat{\theta}_{GHR}\right)$ *under PO design.*

Under PO sampling design, the first through fourth-order inclusion probabilities are available. Therefore, the computations of $\hat{var}_p\left(\hat{\theta}_{GHR}\right)$ can be done directly from Theorem 2.1.3.

Example 2.1.13 $\hat{var}_p\left(\hat{\theta}_{GHR}\right)$ *under stratified sampling design.*

Assuming a general fourth-order measurable stratified sampling design, an unbiased estimator of $var_p\left(\hat{\theta}_{GHR}\right)$ is given by

$$\begin{aligned}\hat{var}_p\left(\hat{\theta}_{GHR}\right) &= \frac{1}{N^2}\sum_{h=1}^{H}\sum_{i,j\in s_h}\frac{y_i^*}{\pi_i}\frac{y_j^*}{\pi_j}\frac{\Delta_{ij}}{\pi_{ij}} + \frac{1}{N^4\bar{x}_{U_N}^2}\sum_{h=1}^{H}\sum_{i,j,k,l\in s_h}\frac{r_i x_j}{\pi_{ij}}\frac{r_k x_l}{\pi_{kl}}\frac{\Delta_{ijkl}}{\pi_{ijkl}} \\ &\quad -\frac{2}{N^3\bar{x}_{U_N}}\sum_{h=1}^{H}\sum_{i,k,l\in s_h}\frac{y_i^*}{\pi_i}\frac{r_k x_l}{\pi_{kl}}\frac{\Delta_{ikl}}{\pi_{ikl}}. \qquad (2.21)\end{aligned}$$

In particular, under STSI with $n_h \geq 4$ in every stratum, it is easily to derive the third-order and fourth-order inclusion probabilities, using arguments like those in Example 2.1.4, equation (2.21) can then be used to compute $\hat{var}_p\left(\hat{\theta}_{GHR}\right)$ under STSI.

2.2 Separate Ratio Estimation for Stratified Sampling Designs

We now consider estimation of the population total or population ratio under Aux_2. The finite population is divided into H non-overlapping strata and $\hat{\theta}_{GHR}$ can be applied to each stratum since $\bar{x}_{U_h}$ is known for $h = 1, \ldots, H$. As noted in Chapter 1, this estimator is motivated by the fact that the classic separate ratio estimator may suffer from accumulation of bias across strata, so substitution of unbiased estimators within each stratum is warranted.

2.2.1 Separate Ratio Estimation Using $\hat{\theta}_{GHR}$

We apply $\hat{\theta}_{GHR}$ within each stratum to obtain a separate ratio estimator. Define

$$\hat{\theta}_{GHR,h} = \frac{1}{N_h}\sum_{i\in s_h} r_i \frac{1}{\pi_i} + \frac{1}{N_h \bar{x}_{U_h}}\left[\sum_{i\in s_h} y_i \frac{1}{\pi_i} - \frac{1}{N_h}\sum_{ij\in s_h} r_i x_j \frac{1}{\pi_{ij}}\right] \tag{2.22}$$

and

$$\hat{t}_{ySep,GHR} = \sum_{h=1}^{H} t_{xh}\hat{\theta}_{GHR,h}.$$

is estimating the population total.

Theorem 2.2.1 *Under a measurable stratified sampling design and* Aux_2,

$$\hat{\theta}_{GHR,Sep} = \sum_{h=1}^{H} \frac{t_{xh}}{t_x}\hat{\theta}_{GHR,h}$$

is an unbiased estimator of the population ratio θ, *and*

$$\hat{t}_{ySep,GHR} = \sum_{h=1}^{H} t_{xh}\hat{\theta}_{GHR,h}$$

is an unbiased estimator of the population total t_y.

Proof: The proof follows directly by applying $\hat{\theta}_{GHR}$ to each stratum under a measurable sampling design $p(\cdot)$, and using Theorem 2.1.1.

■

Remark 2.2.1 Many real surveys are stratified surveys, such as the National Resources Inventory (NRI). The sample design is based on a stratified two-stage area sample of all US lands. Strata are subtownship-level geographic subdivisions in the areas of the country covered by the public Land Survey and analogous geographic subdivisions elsewhere, amounting to tens of thousands of strata. Two primary sampling units are selected in most strata in the first stage of sampling. In the second stage of sampling, three points per Primary Sampling Unit are selected (Breidt (2002)).

The US Current Population Survey (CPS) is a multistage stratified sample. The first stage of the CPS sample design is the selection of counties (see http://www.census.gov/prod/2006pubs/tp-66.pdf). There are approximately 3,000 counties in the US.

From the above two examples, NRI and CPS are highly stratified, and such large numbers of strata enable us to use asymptotic results in which the number of strata goes to infinity.

In the following example, $\hat{\theta}_{GHR,h}$ will be derived under stratified simple random sampling.

Example 2.2.1 $\hat{\theta}_{GHR,Sep}$ *under stratified simple random sampling without-replacement (STSI).*

Assuming $n_h \geq 2$ in each stratum, estimate the population ratio θ by

$$\hat{\theta}_{GHR,Sep} \overset{STSI}{=} \sum_{h=1}^{H} \frac{t_{xh}}{t_x} \left\{ \bar{r}_{s_h} + \frac{n_h (N_h - 1)}{N_h (n_h - 1) \bar{x}_{U_h}} (\bar{y}_{s_h} - \bar{r}_{s_h} \bar{x}_{s_h}) \right\}$$

and the population total t_y by

$$\hat{t}_{ySep,GHR} \overset{STSI}{=} \sum_{h=1}^{H} t_{xh} \left\{ \bar{r}_{s_h} + \frac{n_h (N_h - 1)}{N_h (n_h - 1) \bar{x}_{U_h}} (\bar{y}_{s_h} - \bar{r}_{s_h} \bar{x}_{s_h}) \right\}$$

where $\bar{r}_{s_h} = n_h^{-1} \sum_{i \in S_h} r_i$. This result follows directly from applying the Hartley-Ross estimator of Example 2.1.1 to the h^{th} stratum.

2.2.2 Exact Variance for Separate Ratio Estimation Using GHR

Assuming a measurable stratified sampling design, the variance of $\hat{\theta}_{GHR,Sep}$ and $\hat{t}_{ySep,GHR}$ is given in the following theorem.

Theorem 2.2.2 *Under a measurable stratified sampling design,*

$$var_p\left(\hat{\theta}_{GHR,Sep}\right) = \sum_{h=1}^{H}\left(\frac{t_{xh}}{t_x}\right)^2 var_p\left(\hat{\theta}_{GHR,h}\right)$$

and

$$var_p\left(\hat{t}_{ySep,GHR}\right) = \sum_{h=1}^{H} t_{xh}^2 var_p\left(\hat{\theta}_{GHR,h}\right),$$

where

$$\begin{aligned} var_p\left(\hat{\theta}_{GHR,h}\right) &= \frac{1}{N_h^2}\sum_{i,j\in U_h}\frac{y_i^*}{\pi_i}\frac{y_j^*}{\pi_j}\Delta_{ij} + \frac{1}{N_h^4\bar{x}_{U_h}^2}\sum_{i,j,k,l\in U_h}\frac{r_i x_j}{\pi_{ij}}\frac{r_k x_l}{\pi_{kl}}\Delta_{ijkl} \\ &\quad -\frac{2}{N_h^3\bar{x}_{U_N}}\sum_{i,k,l\in U_h}\frac{y_i^*}{\pi_i}\frac{r_k x_l}{\pi_{kl}}\Delta_{ikl}, \end{aligned}$$

$$\begin{aligned} y_i^* &= \left(\frac{1}{x_i}+\frac{1}{\bar{x}_{U_h}}\right)y_i \\ \Delta_{ij} &= cov_p\left(I_{\{i\in s_h\}}, I_{\{j\in s_h\}}\right) = \pi_{ij}-\pi_i\pi_j \\ \Delta_{ikl} &= cov_p\left(I_{\{i\in s_h\}}, I_{\{kl\in s_h\}}\right) = \pi_{ikl}-\pi_i\pi_{kl} \end{aligned}$$

and

$$\Delta_{ijkl} = cov_p\left(I_{\{ij\in s_h\}}, I_{\{kl\in s_h\}}\right) = \pi_{ijkl}-\pi_{ij}\pi_{kl}.$$

Proof: The proof follows from Theorem 2.1.2 using the definition of $\hat{\theta}_{GHR,Sep}$ and $\hat{t}_{ySep,GHR}$ and the fact that the strata are independent.

■

Example 2.2.2 *$var_p\left(\hat{\theta}_{GHR,Sep}\right)$ and $var_p\left(\hat{t}_{ySep,GHR}\right)$ under STSI.*

Note that

$$\begin{aligned}
\pi_i &= \frac{n_h}{N_h}, \quad \text{for } i \in U_h \\
\pi_{ij} &= \frac{n_h}{N_h}\frac{n_h-1}{N_h-1}, \quad \text{for } i \neq j \in U_h \\
\pi_{ijk} &= \frac{n_h}{N_h}\frac{n_h-1}{N_h-1}\frac{n_h-2}{N_h-2}, \quad \text{for } ijk \in U_{h_{D_3}} \\
\pi_{ijkl} &= \frac{n_h}{N_h}\frac{n_h-1}{N_h-1}\frac{n_h-2}{N_h-2}\frac{n_h-3}{N_h-3}, \quad \text{for } ijkl \in U_{h_{D_4}}.
\end{aligned}$$

Plugging these expressions into Theorem 2.2.2 yields $var_p\left(\hat{\theta}_{GHR,Sep}\right)$ and $var_p\left(\hat{t}_{ySep,GHR}\right)$.

2.2.3 Unbiased Variance Estimation for Separate Ratio Estimators Using $\hat{\theta}_{GHR}$

For fourth-order measurable stratified sampling designs, unbiased estimators for $var_p\left(\hat{\theta}_{GHR,Sep}\right)$ and $var_p\left(\hat{t}_{ySep,GHR}\right)$ exist.

Theorem 2.2.3 *For a fourth-order measurable stratified sampling design, an unbiased estimator of* $var_p\left(\hat{\theta}_{GHR,Sep}\right)$ *is*

$$\hat{var}_p\left(\hat{\theta}_{GHR,Sep}\right) = \sum_{h=1}^{H}\left(\frac{t_{xh}}{t_x}\right)^2 \hat{var}_p\left(\hat{\theta}_{GHR,h}\right)$$

and an unbiased estimator of $var_p\left(\hat{t}_{ySep,GHR}\right)$ *is*

$$\hat{var}_p\left(\hat{t}_{ySep,GHR}\right) = \sum_{h=1}^{H} t_{xh}^2 \hat{var}_p\left(\hat{\theta}_{GHR,h}\right)$$

where

$$\begin{aligned}
\hat{var}_p\left(\hat{\theta}_{GHR,h}\right) &= \frac{1}{N_h^2}\sum_{i,j\in s_h}\frac{y_i^*}{\pi_i}\frac{y_j^*}{\pi_j}\frac{\Delta_{ij}}{\pi_{ij}} + \frac{1}{N_h^4\bar{x}_{U_h}^2}\sum_{i,j,k,l\in s_h}\frac{r_i x_j}{\pi_{ij}}\frac{r_k x_l}{\pi_{kl}}\frac{\Delta_{ijkl}}{\pi_{ijkl}} \\
&\quad - \frac{2}{N_h^3\bar{x}_{U_N}}\sum_{i,k,l\in s_h}\frac{y_i^*}{\pi_i}\frac{r_k x_l}{\pi_{kl}}\frac{\Delta_{ikl}}{\pi_{ikl}}.
\end{aligned}$$

Proof: The proof follows from Theorem 2.2.2 and Theorem 2.1.3.

■

Example 2.2.3 *$\hat{var}_p\left(\hat{\theta}_{GHR,Sep}\right)$ and $\hat{var}_p\left(\hat{t}_{ySep,GHR}\right)$ under STSI.*

If $n_h \geq 4$ for $h = 1, \ldots, H$, then STSI is fourth-order measurable. Using the inclusion probabilities defined in Example 2.2.2, $\hat{var}_p\left(\hat{\theta}_{GHR,Sep}\right)$ and $\hat{var}_p\left(\hat{t}_{ySep,GHR}\right)$ can be computed from Theorem 2.2.3.

2.3 Alternative Unbiased Ratio Estimation for With-Replacement Designs

We now consider another version of unbiased ratio estimation for the special case of with-replacement designs, using the Hansen and Hurwitz (1943) estimation idea. This alternative estimator is of interest of its own right, but is also of interest for producing an approximate variance estimator for $\hat{\theta}_{GHR}$ that does not require higher-order inclusion probabilities. We consider this approximation in Section 2.4. Recall that for a with-replacement design, $p_k = Pr$ (selecting element k on a single draw) > 0 for $k = 1, \ldots, N$, and $\sum_{k \in U_N} p_k = 1$.

Define

$$\left.\begin{array}{ll} Z_{Ni}(y) = \frac{1}{N}\sum_{k\in U}\frac{y_k}{p_k}I_{\{\kappa_i=k\}} & Z_{Ni}(x) = \frac{1}{N}\sum_{k\in U}\frac{x_k}{p_k}I_{\{\kappa_i=k\}} \\ Z_{Ni}(r) = \frac{1}{N}\sum_{k\in U}\frac{r_k}{p_k}I_{\{\kappa_i=k\}} & Z_{Ni}(\check{y}) = \frac{1}{N}\sum_{k\in U}\frac{\check{y}_k}{p_k}I_{\{\kappa_i=k\}} \end{array}\right\} \tag{2.23}$$

where $\check{y}_k = y_k/(Np_k)$ and $r_k = y_k/x_k$. Note that $Z_{Ni}(a)$ is a random variable with a discrete distribution assigning probability p_k to the values $N^{-1}a_k/p_k$, $k \in U$. It follows from the WR sampling scheme that $Z_{Ni}(a)$ are iid for $i = 1, \ldots, m$, with mean

$$E_p\left[Z_{Ni}(a)\right] = \frac{1}{N}\sum_{k\in U}\frac{a_k p_k}{p_k} = \frac{t_{aN}}{N}, \tag{2.24}$$

and variance

$$var_p\left[Z_{Ni}(a)\right] = \frac{1}{N^2}\left\{\sum_{k\in U}\left(\frac{a_k}{p_k}\right)^2 p_k - \left(\sum_{k\in U}\frac{a_k p_k}{p_k}\right)^2\right\}. \tag{2.25}$$

2.3.1 Unbiased Ratio Estimation Using Hansen-Hurwitz Estimators

Under a with-replacement sampling design and using Hansen and Hurwitz (1943) estimation ideas, define

$$\begin{aligned}\tilde{\theta}_{GHR} &= \bar{Z}_{Nm}(r) + \frac{1}{\bar{x}_{U_N}}\bar{Z}_{Nm}(y) + \frac{1}{(m-1)\bar{x}_{U_N}}\bar{Z}_{Nm}(\check{y}) \\ &\quad - \frac{m}{(m-1)\bar{x}_{U_N}}\bar{Z}_{Nm}(r)\,\bar{Z}_{Nm}(x) \end{aligned} \tag{2.26}$$

where

$$\bar{Z}_{Nm}(\cdot) = \frac{1}{m}\sum_{i=1}^{m} Z_{Ni}(\cdot),$$

and $Z_{Ni}(\cdot)$ are given by (2.23).

Theorem 2.3.1 *Under a with-replacement sampling design, $\tilde{\theta}_{GHR}$ defined in equation (2.26) is unbiased for the population ratio θ. Further,*

$$\tilde{t}_{y,GHR} = t_x\tilde{\theta}_{GHR}$$

is unbiased for the population total t_y.

Proof: First note that

$$\begin{aligned} E_p\left[\bar{Z}_{Nm}(r)\,\bar{Z}_{Nm}(x)\right] &= \frac{1}{N^2m^2}\sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k\in U}\sum_{l\in U}\frac{r_k}{p_k}\frac{x_l}{p_l}E_p\left[I_{\{\kappa_i=k\}}I_{\{\kappa_j=l\}}\right] \\ &= \frac{1}{N^2m^2}\sum_{i=1}^{m}\sum_{k\in U}\frac{y_k}{p_k} \\ &\quad + \frac{1}{N^2m^2}\sum_{i\neq j}\sum_{k,\,l\in U}\frac{r_k}{p_k}\frac{x_l}{p_l}E_p\left[I_{\{\kappa_i=k\}}\right]E_p\left[I_{\{\kappa_j=l\}}\right] \\ &= \frac{1}{N^2m^2}\left\{Nmt_{\check{y}N} + m(m-1)\,t_{rN}t_{xN}\right\} \\ &= \frac{1}{Nm}t_{\check{y}N} + \frac{m-1}{N^2m}t_{rN}t_{xN}. \end{aligned} \tag{2.27}$$

Also, $E_p\left[\bar{Z}_{Nm}(a)\right] = \frac{1}{N}t_{aN}$ by (2.24). Hence, from (2.26) and (2.27), we have

$$\begin{aligned} E_p\left(\tilde{\theta}_{GHR}\right) &= \frac{1}{N}t_{rN} + \frac{1}{N\bar{x}_{U_N}}t_{yN} + \frac{1}{N(m-1)\bar{x}_{U_N}}t_{\check{y}N} \\ &\quad - \frac{m}{N(m-1)\bar{x}_{U_N}}\left[\frac{1}{Nm}t_{\check{y}N} + \frac{m-1}{N^2m}t_{rN}t_{xN}\right] \\ &= \frac{1}{N\bar{x}_{U_N}}t_{yN} = \theta, \end{aligned} \tag{2.28}$$

and $E_p\left(\tilde{t}_{y,GHR}\right) = t_y$.

■

Remark 2.3.1 The exact variance of $\tilde{t}_{y,GHR}$ is not readily available. In Chapter 3, we will give an asymptotic variance and a consistent estimator of the asymptotic variance.

2.4 Simple Ways to Approximate $\hat{v}ar_p\left(\hat{\theta}_{GHR}\right)$

First through fourth-order inclusion probabilities are readily available for SI, PO, and WR sampling designs, but are not generally available. Standard software like SAS `proc surveyselct` will compute first and second-order inclusion probabilities under various designs, but not higher-order. In this section, we will introduce two methods of variance estimation that not require higher-order inclusion probabilities.

The first method, which we call the uncorrelated variance estimation method, treat the sample membership indicators as approximately uncorrelated, so that $\pi_{ij} \doteq \pi_i\pi_j$ for $i \neq j$. This method is discussed in Section 2.4.1. It requires first and second-order inclusion probabilities.

The second method, discussed in Section 2.4.2,approximates the sampling design as a with-replacement sampling design. This method requires only first-order inclusion probabilities.

2.4.1 Uncorrelated Variance Estimation

Assume

$$\pi_{ij} \geq \pi_{*N} > 0$$

and

$$\limsup_{N\to\infty} n_N \max_{ij\in U_N: i\neq j} |\pi_{ij} - \pi_i\pi_j| = O(1).$$

This assumption enable us to approximate π_{ij} by $\pi_i\pi_j$. Such an assumption holds exactly under PO sampling, and approximately under SI, WR sampling designs. Under this assumption,

$$\begin{aligned}
|\text{diff}| &= \left| \frac{1}{N^2\bar{x}_{U_N}} \sum_{i,j\in s:\, i\neq j} r_i x_j \frac{1}{\pi_{ij}} - \frac{1}{N^2\bar{x}_{U_N}} \sum_{i,j\in s:\, i\neq j} r_i x_j \frac{1}{\pi_i\pi_j} \right| \\
&\leq \frac{1}{N^2\bar{x}_{U_N}} \left| \sum_{i,j\in s:\, i\neq j} r_i x_j \frac{\pi_i\pi_j - \pi_{ij}}{\pi_{ij}\pi_i\pi_j} \right| \\
&\leq \frac{1}{n_N N^2\bar{x}_{U_N}} n_N \max_{i\neq j} |\pi_i\pi_j - \pi_{ij}| \frac{1}{\pi^3_{*N}} \sum_{i\in s} |r_i| \sum_{j\in s} x_j \\
&= O\left(\frac{1}{n_N\pi^3_{*N}}\right), \qquad (2.29)
\end{aligned}$$

which goes to zero under mild additional assumptions. Thus, from equation (2.1), we have

$$\begin{aligned}
\hat{\theta}_{GHR} &\doteq \frac{1}{N}\sum_{i\in U} \frac{y_i^*}{\pi_i} I_{\{i\in s\}} - \frac{1}{N^2\bar{x}_{U_N}} \left[\sum_{ij\in U_{D_2}} r_i x_j \frac{I_{\{i,j\in s\}}}{\pi_i\pi_j} + \sum_{i\in U} y_i \frac{I_{\{i\in s\}}}{\pi_i} \right] \\
&= \frac{1}{N}\sum_{i\in U} y_i^* \frac{I_{\{i\in s\}}}{\pi_i} - \frac{1}{N^2\bar{x}_{U_N}} \left[\sum_{ij\in U} r_i x_j \frac{I_{\{i\in s\}} I_{\{j\in s\}}}{\pi_i\pi_j} - \sum_{i\in U} y_i \frac{I_{\{i\in s\}}}{\pi_i^2} + \sum_{i\in U} y_i \frac{I_{\{i\in s\}}}{\pi_i} \right] \\
&= \frac{1}{N}\sum_{i\in U} \left(1 + \frac{x_i}{\bar{x}_{U_N}}\right) r_i \frac{I_{\{i\in s\}}}{\pi_i} - \frac{1}{N^2\bar{x}_{U_N}} \left[\sum_{i\in U} r_i \frac{I_{\{i\in s\}}}{\pi_i} \sum_{i\in U} x_j \frac{I_{\{j\in s\}}}{\pi_j} \right. \\
&\quad \left. - \sum_{i\in U} \frac{y_i}{\pi_i} \frac{I_{\{i\in s\}}}{\pi_i} + \sum_{i\in U} y_i \frac{I_{\{i\in s\}}}{\pi_i} \right] \\
&= \frac{1}{N}\hat{t}_{r\pi} + \frac{N-1}{Nt_x}\hat{t}_{y\pi} + \frac{1}{Nt_x}\hat{t}_{\frac{y}{\pi}\cdot\pi} - \frac{1}{Nt_x}\hat{t}_{x\pi}\hat{t}_{r\pi} \qquad (2.30) \\
&\doteq \frac{1}{N}\hat{t}_{r\pi} + \frac{N-1}{Nt_x}\hat{t}_{y\pi} + \frac{1}{Nt_x}\hat{t}_{\frac{y}{\pi}\cdot\pi} - \frac{1}{Nt_x}\left[t_x t_r + t_r\left(\hat{t}_{x\pi} - t_x\right) + t_x\left(\hat{t}_{r\pi} - t_r\right)\right] \\
&= \frac{1}{N}t_r + \frac{1}{Nt_x}\sum_{i\in U}\left[(N-1)\,y_i + \frac{y_i}{\pi_i} - t_r x_i\right]\frac{I_{\{i\in s\}}}{\pi_i}. \qquad (2.31)
\end{aligned}$$

The variance of $\hat{\theta}_{GHR}$ can be approximated by taking the variance of the right hand side of equation (2.31), we have

$$var_{app}\left(\hat{\theta}_{GHR}\right) \doteq var\left\{\frac{1}{N}t_r + \frac{1}{Nt_x}\sum_{i\in U}\left[(N-1)\,y_i + \frac{y_i}{\pi_i} - t_r x_i\right]\frac{I_{\{i\in s\}}}{\pi_i}\right\}$$

$$= \frac{1}{N^2 t_x^2} \sum_{ij \in U} \frac{w_i}{\pi_i} \frac{w_j}{\pi_j} \Delta_{ij} \tag{2.32}$$

where

$$w_i = (N-1)\, y_i + \frac{y_i}{\pi_i} - t_r x_i.$$

An unbiased estimator for $var_{app}\left(\hat{\theta}_{GHR}\right)$ is

$$\hat{var}_{app}\left(\hat{\theta}_{GHR}\right) = \frac{1}{N^2 t_x^2} \sum_{ij \in s} \frac{\hat{w}_i}{\pi_i} \frac{\hat{w}_j}{\pi_j} \frac{\Delta_{ij}}{\pi_{ij}} \tag{2.33}$$

where

$$\hat{w}_i = (N-1)\, y_i + \frac{y_i}{\pi_i} - \hat{t}_{r\pi} x_i.$$

Note that (2.33) requires first and second-order inclusion probabilities.

Remark 2.4.1 In Chapter 3, we will show that $\hat{var}_{app}\left(\hat{\theta}_{GHR}\right)$ is consistent for $var_p\left(\hat{\theta}_{GHR}\right)$ under SI and PO sampling.

2.4.2 With-Replacement Approximation

This method requires the first-order inclusion probabilities to be known and "pretends" that sampling is done via a with-replacement design.

Define $p_k =: m^{-1}\pi_k$ where π_k are the first-order inclusion probabilities for the original design.

We now construct a with-replacement version of the approximation (2.31). Define

$$Z_i = \frac{1}{N t_x} \sum_{k \in U} \left[\left(N - 1 + \frac{1}{\pi_k} \right) y_k - t_r x_k \right] \frac{I_{\{\kappa_i = k\}}}{p_k},$$

where $I_{\{\kappa_i = k\}}$ is one when the k^{th} element is selected in the i^{th} draw and zero otherwise. Now,

$$EZ_i \overset{WR}{=} \frac{1}{N t_x} \sum_{k \in U} \left[\left(N - 1 + \frac{1}{\pi_k} \right) y_k - t_r x_k \right]$$

$$= t_z$$

$$\begin{aligned}
var\left(Z_i\right) &\overset{WR}{=} E\left(Z_i - t_z\right)^2 \\
&= E\left(\frac{1}{Nt_x}\sum_{k\in U}\left[\left(N-1+\frac{1}{\pi_k}\right)y_k - t_r x_k\right]\frac{I_{\{\kappa_i=k\}}}{p_k} - \sum_{k\in U} t_z I_{\{\kappa_i=k\}}\right)^2 \\
&= E\left(\sum_{k\in U}\left\{\frac{1}{Nt_x}\left[\left(N-1+\frac{1}{\pi_k}\right)y_k - t_r x_k\right]\frac{1}{p_k} - t_z\right\} I_{\{\kappa_i=k\}}\right)^2 \\
&= \sum_{kl\in U}\left\{\frac{1}{Nt_x}\left[\left(N-1+\frac{1}{\pi_k}\right)y_k - t_r x_k\right]\frac{1}{p_k} - t_z\right\} \\
&\quad \left\{\frac{1}{Nt_x}\left[\left(N-1+\frac{1}{\pi_l}\right)y_l - t_r x_l\right]\frac{1}{p_l} - t_z\right\} E I_{\{\kappa_i=k\}} I_{\{\kappa_i=l\}} \\
&= \sum_{k\in U}\left\{\frac{1}{Nt_x}\left[\left(N-1+\frac{1}{\pi_k}\right)y_k - t_r x_k\right]\frac{1}{p_k} - t_z\right\}^2 p_k \\
&= V_1. \qquad (2.34)
\end{aligned}$$

Since Z_i are $iid\left(t_z, V_1\right)$, then

$$\hat{t}_{pwr} = \bar{Z} = \frac{1}{n}\sum_{i=1}^{n} Z_i$$

is an unbiased estimator, under WR, for t_z and $var_{pwr}\left(\hat{t}_{pwr}\right) = n^{-1}V_1$. Therefore,

$$\hat{v}ar_{pwr}\left(\hat{t}_{pwr}\right) = \frac{1}{n\left(n-1\right)}\sum_{i=1}^{n}\left(Z_i - \bar{Z}\right)^2 \qquad (2.35)$$

ia an unbiased estimator for $var_{pwr}\left(\hat{t}_{pwr}\right)$. Estimate Z_i by

$$\hat{Z}_i = \frac{1}{Nt_x}\sum_{k\in U}\left[\left(N-1+\frac{1}{\tilde{\pi}_k}\right)y_k - \hat{t}_{r\tilde{\pi}}x_k\right]\frac{I_{\{\Psi_i=k\}}}{p_k}$$

and compute $\hat{v}ar_{pwr}\left(\hat{t}_{pwr}\right)$ as a with-replacement approximation to the variance of (2.31), and hence as a WR approximation to the variance of $\hat{\theta}_{GHR}$.

2.4.3 Simulation Results

We will compare the two methods of variance approximation through simulations. Let x_i be iid $Gamma(\alpha = 3, \beta = 2)$ with mean 6 and variance

12, ϵ_i iid $N(0, 25x_i)$, and $y_i = 3x_i + \epsilon_i$. The entire population consists of $(x_1, y_1), \ldots, (x_{1000}, y_{1000})$.

Define the following terms

$$var_{emp}\left(\hat{\theta}_{GHR}\right) \text{ is the empirical (simulation) variance of } \hat{\theta}_{GHR}$$

$$\hat{v}ar_{app}\left(\hat{\theta}_{GHR}\right) \text{ is defined in equation(2.33)}$$

$$\hat{v}ar_{pwr}\left(\hat{\theta}_{GHR}\right) \text{ is defined in equation (2.35)}$$

and

$$\%RVB = 100\frac{E\left(\hat{v}ar(\cdot)\right) - var_{emp}(\cdot)}{var_{emp}(\cdot)}.$$

Table 2.1 shows the simulation performance of the two methods as defined in equations (2.33) and (2.35) based on 1500 replications of SI and different sample sizes. The values of $\%RVB$ when the second method is used are higher consistently than the values of $\%RVB$ when the first method is used. For both methods, the values of the approximate variance estimators are fairly conservative, but track the actual variance well.

n	10	15	20	25	30	35	40	75
$var_{emp}\left(\hat{\theta}_{GHR}\right)$	0.456	0.290	0.218	0.178	0.137	0.122	0.108	0.047
$\hat{v}ar_{app}\left(\hat{\theta}_{GHR}\right)$	0.567	0.341	0.246	0.190	0.156	0.130	0.113	0.056
$\%RVB$	24.342	17.586	12.844	6.742	13.869	6.557	4.630	19.149
$\hat{v}ar_{pwr}\left(\hat{\theta}_{GHR}\right)$	0.573	0.347	0.251	0.195	0.161	0.134	0.118	0.061
$\%RVB$	25.658	19.655	15.138	9.551	17.518	9.836	9.259	29.787

Table 2.1: Performance of two variance estimation approximations under SI Based on 1500 simulated simple random samples from a fixed finite population.

We next consider the performance of the two approximations under πps sampling. Let $z_i = 10 + x_i + \eta_i$, where ϵ_i and η_i are independent and η_i are iid $N(0, 1)$,

be the size variable to be used in the probability proportional to size sampling design. The first and second-order inclusion probabilities are obtained from the `out=JTPROBS` option in SAS `proc surveyselect`. Table 2.2 shows the simulation performance of the same two methods of approximations of $\hat{v}ar_p\left(\hat{\theta}_{GHR}\right)$, but in this case under πps sampling. In this case, the two methods perform similarly throughout, tracking the true variance fairly well throughou, through non-conservatively at higher sample sizes.

n	10	15	20	25	30	35	40	75
$Var_{Emp}\left(\hat{\theta}_{GHR}\right)$	0.428	0.271	0.211	0.176	0.145	0.141	0.117	0.066
$\hat{V}ar_{app}\left(\hat{\theta}_{GHR}\right)$	0.520	0.325	0.233	0.184	0.153	0.132	0.114	0.059
$\%RVB$	21.495	19.926	10.427	4.546	5.517	-6.383	-2.564	-10.606
$\hat{V}ar_{pwr}\left(\hat{\theta}_{GHR}\right)$	0.523	0.326	0.235	0.186	0.151	0.129	0.112	0.058
$\%RVB$	22.196	20.295	11.374	5.682	4.138	-8.51	-4.274	-12.121

Table 2.2: Performance of two variance estimation approximations under πps sampling, based on 1500 simulated πps samples from a fixed finite population.

Finally, we consider the performance of the two variance estimation approximations under PO. The first-order inclusion probabilities are computed from

$$\pi_i = b\frac{\exp\left(1.5x_i\right)}{1+\exp\left(1.5x_i\right)} \quad \text{for} \quad b > .0$$

Because sample size is random for PO, choose b such that $Nb\bar{E} = 10,\ 15,\ \ldots, 40,\ 75$, where $\bar{E} = 1500^{-1}\sum_{i=1}^{1500}\left\{\exp\left(1.5x_i\right)\left(1+\exp\left(1.5x_i\right)\right)^{-1}\right\}$. Furthermore, under PO sampling design, we have $\pi_{ij} = \pi_i\pi_j$. Therefore, one can expect especially the first method of approximation will give excellent results. Table 2.3 shows the values of $\%RVB$. In all cases, both methods of approximate variance estimation work extremely well.

$E(n_s)$	10	15	20	25	30	35	40	75
$Var_{Emp}\left(\hat{\theta}_{GHR}\right)$	0.267	0.173	0.125	0.095	0.078	0.066	0.056	0.026
$\hat{V}ar_{app}\left(\hat{\theta}_{GHR}\right)$	0.270	0.166	0.119	0.092	0.075	0.063	0.054	0.026
$\%RVB$	1.124	-4.046	-4.8	-3.158	-3.846	-4.546	-3.571	0
$\hat{V}ar_{pwr}\left(\hat{\theta}_{GHR}\right)$	0.259	0.164	0.120	0.094	0.078	0.066	0.058	0.030
$\%RVB$	-2.996	-5.202	-4	-1.053	0	0	3.571	15.385

Table 2.3: Performance of two variance estimation approximations under PO sampling, based on 1500 simulated PO samples from a fixed finite population.

The first method of approximation of $\hat{v}ar_p\left(\hat{\theta}_{GHR}\right)$ will be used under stratified sampling in Section 3.4.

2.5 Combining $\hat{\theta}_{GHR}$ and $\hat{\theta}$

The estimator $\hat{\theta}_{GHR}$ is exactly unbiased but may have large variance, while the estimator $\hat{\theta}$ is asymptotically unbiased and has variance less than the variance of $\hat{\theta}_{GHR}$. It is natural to consider convex combinations of the two estimators, to produce asymptotically unbiased estimator with potentially smaller mean square error than either $\hat{\theta}_{GHR}$ or $\hat{\theta}$. Furthermore, under SI sampling design, it turns out that the estimators for Murthy and Nanjamma (1959), Nieto de Pascual (1954), and other estimators can be obtained from such combinations.

2.5.1 Optimal Combination

For $\lambda \in [0, 1]$ define

$$\breve{\theta} = \lambda\hat{\theta} + (1 - \lambda)\,\hat{\theta}_{GHR}.$$

Therefore,

$$E_p\breve{\theta} \;=\; \theta + \lambda bias\left(\hat{\theta}\right)$$

$$
\begin{aligned}
MSE_p\left(\breve{\theta}\right) &= E_p\left(\breve{\theta}-\theta\right)^2 \\
&= E_p\left[\left(\breve{\theta}-E_p\breve{\theta}\right)+\left(E_p\breve{\theta}-\theta\right)\right]^2 \\
&= var_p\left(\breve{\theta}\right)+\lambda^2 bias_p^2\left(\hat{\theta}\right) \\
&= \lambda^2 var_p\left(\hat{\theta}\right)+(1-\lambda)^2 var_p\left(\hat{\theta}_{GHR}\right)+2\lambda(1-\lambda)\,cov_p\left(\hat{\theta},\hat{\theta}_{GHR}\right) \\
&\quad +\lambda^2 bias^2\left(\hat{\theta}\right). \qquad (2.36)
\end{aligned}
$$

Differentiate (2.36) with respect to λ, equate to zero, and solve for λ, we have

$$
\lambda=\frac{var_p\left(\hat{\theta}_{GHR}\right)-cov_p\left(\hat{\theta},\hat{\theta}_{GHR}\right)}{var_p\left(\hat{\theta}_{GHR}\right)+var_p\left(\hat{\theta}\right)-2cov_p\left(\hat{\theta},\hat{\theta}_{GHR}\right)+bias^2\left(\hat{\theta}\right)}. \qquad (2.37)
$$

Since the equation (2.36) is quadratic in λ and the coefficient of λ^2 is positive $\left(var_p\left(\hat{\theta}-\hat{\theta}_{GHR}\right)=var_p\left(\hat{\theta}_{GHR}\right)+var_p\left(\hat{\theta}\right)-2cov_p\left(\hat{\theta},\hat{\theta}_{GHR}\right)\right)$, it follows that the given value of λ in (2.37) minimizes equation (2.36). This optimal value of λ is unknown in practice but might be estimated from the sample.

2.5.2 Relationship to Earlier Literature

Under SI design and for different choices of λ we can obtain different estimators given in the literature:

$$
\begin{aligned}
\breve{\theta} &\overset{SI}{=} \frac{\bar{y}_s}{\bar{x}_s}\lambda+(1-\lambda)\left[\bar{r}_s+\frac{n(N-1)}{N(n-1)\bar{x}_{U_N}}(\bar{y}_s-\bar{r}_s\bar{x}_s)\right] \\
&= \bar{r}_s+\frac{n(N-1)}{N(n-1)\bar{x}_{U_N}}(\bar{y}_s-\bar{r}_s\bar{x}_s)+\lambda\left[\frac{\bar{y}_s}{\bar{x}_s}-\left(\bar{r}_s+\frac{n(N-1)}{N(n-1)\bar{x}_{U_N}}(\bar{y}_s-\bar{r}_s\bar{x}_s)\right)\right] \\
&= \bar{r}_s+\frac{n(N-1)}{N(n-1)\bar{x}_{U_N}}(\bar{y}_s-\bar{r}_s\bar{x}_s)+\lambda\left[\frac{1}{\bar{x}_s}-\frac{n(N-1)}{N(n-1)\bar{x}_{U_N}}\right](\bar{y}_s-\bar{r}_s\bar{x}_s). \qquad (2.38)
\end{aligned}
$$

The estimator due to Murthy and Nanjamma (1959) can be obtained from (2.38) by taking

$$
\lambda=\frac{n(\bar{x}_{U_N}-\bar{x}_s)}{(n-1)\bar{x}_{U_N}-n\bar{x}_s}
$$

and using the approximate $(N-1)N^{-1}\cong 1$.

Furthermore, the estimator due to Nieto de Pascual (1961) can be obtained from (2.38) by taking

$$\lambda = 1 + \frac{N\bar{x}_s}{N(n-1)\bar{x}_{U_N} - n(N-1)\bar{x}_s}.$$

Of course, the Hartley and Ross (1954) estimator is obtained by letting $\lambda = 0$, and the simple estimator is obtained by letting $\lambda = 1$.

Chapter 3

ASYMPTOTIC RESULTS

In this chapter we will discuss the asymptotic properties for estimators derived in earlier sections, including results on mean square consistency, central limit theory, and the Godambe and Joshi (1965) lower bound.

3.1 Asymptotic Results for $\hat{\theta}_{GHR}$

To study asymptotic properties of estimators in finite populations, we can imagine that we have sequences of finite populations and associated probability samples (eg., Hájek (1960), Breidt and Opsomer (2000), Isaki and Fuller (1982), Krewski and Rao (1981), Bickel and Freedman (1984)). We will assume that the N^{th} finite population, $U_N = \{1, \ldots, N\}$, contains N elements. Therefore, the first-order, second-order, and higher order inclusion probabilities are sequences that depend on N. For simplicity of notation, we will drop the index N.

Definition 3.1.1 Given sequences of finite populations, the estimator $\tilde{\theta}$ is *design consistent* for the finite population parameter θ if for every $\epsilon > 0$

$$\lim_{N\to\infty} Pr\left(\left|\tilde{\theta} - \theta\right| \geq \epsilon\right) = 0$$

where the probabilities are computed with respect to the sequence of sampling designs.

Definition 3.1.2 Given sequences of finite populations, the estimator $\tilde{\theta}$ is *mean square consistent* (MSC) under the design for the finite population parameter θ if

$$\lim_{N\to\infty} E_p\left(\tilde{\theta} - \theta\right)^2 = 0,$$

where the expectations are computed with respect to the sequence of sampling designs.

Remark 3.1.1 By Chebychev's inequality,

$$Pr\left(\left|\tilde{\theta}-\theta\right| \geq \epsilon\right) \leq \frac{E_p\left(\tilde{\theta}-\theta\right)^2}{\epsilon^2}.$$

Therefore, if $\tilde{\theta}$ is MSC for the population parameter θ then it is a design consistent estimator for θ.

Consider the following conditions,

A_1- $\pi_i \geq \pi_{ij} \geq \pi_{N*} > 0$ for all $ij \in U_N$

A_2- $\sum_{i,j\in U} \Delta_{ij}^2 = O\left(N^{\eta^*}\right)$ for some $\eta^* < 2$

A_3- $\sum_{i,j,k,l\in U} \Delta_{ijkl}^2 = O\left(N^{\eta}\right)$ for some $\eta < 4$

A_4- $N^{\min\left\{1,\frac{2-\eta^*}{4},\frac{4-\eta}{4}\right\}}\pi_{N*} \to \infty$ as $N \to \infty$

A_5- $\limsup_{N\to\infty} \frac{1}{N}\sum_{i\in U} y_i^2 < \infty$

A_6- $\limsup_{N\to\infty} \frac{1}{N}\sum_{i\in U} x_i^2 < \infty$

A_7- $\limsup_{N\to\infty} \frac{1}{N}\sum_{i\in U} \frac{y_i^2}{x_i} < \infty$

A_8- $\limsup_{N\to\infty} \frac{1}{N}\sum_{i\in U} \left(\frac{y_i}{x_i}\right)^2 < \infty$

A_9- $\liminf_{N\to\infty} \frac{1}{N}\sum_{i\in U} x_i > 0.$

Since we have sequences of finite populations, the first-order and second-order inclusion probabilities are sequences based on sequences of sampling designs. To keep $\hat{\theta}_{GHR}$ defined through all the sequences of finite populations, condition A_1 is assumed and simply says that the first and second-order inclusion probabilities are bounded away from 0, ensuring that the designs are all measurable.

Since $|\Delta_{ij}| \leq 2$, the highest order for $\sum_{i,j\in U} \Delta_{ij}^2$ is $O\left(N^2\right)$. Therefore, Condition A_2 is assumed to exclude this case and ensure weaker dependence among sample membership indicators. For similar reasons we have condition A_3.

The expected sample size is $\sum_{i\in U}\pi_i \geq N\pi_{N*}$, and we need to guarantee that this expected sample size will tend to infinity as $N \to \infty$, even if the minimum sampling rate $\pi_{N*} \to 0$. If η or η^* are large, meaning strong dependence in the design, then we need the sample size to go to infinity even more rapidly. Therefore, we assume condition A_4. Assumptions $A_5 - A_9$ are moment conditions for the finite population. They will be satisfied, for example, if we assume $0 < l_x \leq x_i \leq u_x < \infty$ and $y_i \leq u_y < \infty$. Later, we will demonstrate that conditions $A_1 - A_4$ hold for simple random sampling without replacement, simple random cluster sampling (SIC) and general stratified sampling designs.

Theorem 3.1.1 *Under $A_1 - A_9$, $\hat{\theta}_{GHR}$ is a mean square consistent estimator for θ.*

Proof: Since the design is measurable under A_1, $\hat{\theta}_{GHR}$ is unbiased for θ by Theorem 2.1.1. It suffices to show that $var_p\left(\hat{\theta}_{GHR}\right) \to 0$ as $N \to \infty$. Rewrite $var_p\left(\hat{\theta}_{GHR}\right) = B_N + C_N - D_N$, where

$$B_N = \frac{1}{N^2}\sum_{i,j\in U}\frac{y_i^*}{\pi_i}\frac{y_j^*}{\pi_j}\Delta_{ij} \tag{3.1}$$

$$C_N = \frac{1}{N^4\bar{x}_{U_N}^2}\sum_{i,j,k,l\in U}\frac{r_i x_j}{\pi_{ij}}\frac{r_k x_l}{\pi_{kl}}\Delta_{ijkl} \tag{3.2}$$

$$D_N = \frac{2}{N^3\bar{x}_{U_N}}\sum_{i,k,l\in U}\frac{y_i^*}{\pi_i}\frac{r_k x_l}{\pi_{kl}}\Delta_{ikl} \tag{3.3}$$

and

$$y_i^* = \left(\frac{1}{x_i} + \frac{1}{\bar{x}_{U_N}}\right) y_i. \tag{3.4}$$

It is enough to show that $B_N \to 0$ and $C_N \to 0$ as $N \to \infty$, since $|D_N| \leq B_N^{1/2}C_N^{1/2}$ by the Cauchy-Schwarz inequality. Now

$$B_N = \frac{1}{N^2}\left[\sum_{i\in U}\left(\frac{y_i^*}{\pi_i}\right)^2 \pi_i\left(1-\pi_i\right) + \sum_{i,j\in U_{D_2}}\frac{y_i^*}{\pi_i}\frac{y_j^*}{\pi_j}\Delta_{ij}\right]$$

$$
\begin{aligned}
&\leq \frac{1}{N\pi_{N*}}\left(\frac{1}{N}\sum_{i\in U} y_i^{*2}\right) + \frac{1}{N\pi_{N*}^2}\left(\sum_{i,j\in U_{D_2}} \Delta_{ij}^2\right)^{\frac{1}{2}}\left(\sum_{i,j\in U_{D_2}} \frac{y_i^{*2}}{N}\frac{y_j^{*2}}{N}\right)^{\frac{1}{2}} \\
&\leq \frac{1}{N\pi_{N*}}\left(\frac{1}{N}\sum_{i\in U} y_i^{*2}\right) + \frac{1}{N\pi_{N*}^2}\left(\sum_{i,j\in U} \Delta_{ij}^2\right)^{\frac{1}{2}}\left(\sum_{i,j\in U} \frac{y_i^{*2}}{N}\frac{y_j^{*2}}{N}\right)^{\frac{1}{2}} \\
&\leq \frac{1}{N\pi_{N*}}\left(\frac{1}{N}\sum_{i\in U} y_i^{*2}\right) + \frac{1}{N\pi_{N*}^2}\left(\sum_{i,j\in U} \Delta_{ij}^2\right)^{\frac{1}{2}}\left(\frac{1}{N}\sum_{i\in U} y_i^{*2}\right) \\
&= \frac{1}{N\pi_{N*}}\left(\frac{1}{N}\sum_{i\in U} y_i^{*2}\right) + \frac{1}{N^{\frac{2-\eta^*}{2}}\pi_{N*}^2}\left(\sum_{i,j\in U} \frac{\Delta_{ij}^2}{N^{\eta^*}}\right)^{\frac{1}{2}}\left(\frac{1}{N}\sum_{i\in U} y_i^{*2}\right). \qquad (3.5)
\end{aligned}
$$

By assumptions $A_1 - A_9$, $B_N \to 0$ as $N \to \infty$.

Furthermore,

$$
\begin{aligned}
C_N &\leq \frac{1}{N^4\pi_{N*}^2\bar{x}_{U_N}^2}\sum_{i,j,k,l\in U}|r_i x_j r_k x_l \Delta_{ijkl}| \\
&\leq \frac{1}{N^{2-\frac{\eta}{2}}\pi_{N*}^2\bar{x}_U^2}\left(\frac{\sum_{i,j,k,l\in U}\Delta_{ijkl}^2}{N^\eta}\right)^{\frac{1}{2}}\left(\frac{1}{N}\sum_{i\in U} r_i^2\right)^{\frac{1}{2}}\left(\frac{1}{N}\sum_{j\in U} x_j^2\right)^{\frac{1}{2}} \\
&\quad \times \left(\frac{1}{N}\sum_{k\in U} r_k^2\right)^{\frac{1}{2}}\left(\frac{1}{N}\sum_{l\in U} x_l^2\right)^{\frac{1}{2}} \\
&= \frac{1}{N^{2-\frac{\eta}{2}}\pi_{N*}^2\bar{x}_U^2}\left(\frac{\sum_{i,j,k,l\in U}\Delta_{ijkl}^2}{N^\eta}\right)^{\frac{1}{2}}\left(\frac{1}{N}\sum_{i\in U}\left(\frac{y_i}{x_i}\right)^2\right)\left(\frac{1}{N}\sum_{j\in U} x_j^2\right). \qquad (3.6)
\end{aligned}
$$

By assumptions $A_1 - A_9$, $C_N \to 0$ as $N \to \infty$. This concludes the proof.

■

There are a lot of cases in the quadruple sum of condition A_3. To reduce the number of cases to be checked in determining the value of η, consider the following results.

Result 3.1.1 *Consider*

$$
\begin{aligned}
A_{2.1}: &\quad \textstyle\sum_{i\in U}\pi_i^2 = O\left(N^{\eta^*}\right) \\
A_{2.2}: &\quad \textstyle\sum_{ij\in U_{D_2}}(\pi_{ij} - \pi_i\pi_j)^2 = O\left(N^{\eta^*}\right).
\end{aligned} \qquad (3.7)
$$

If both $A_{2.1}$ and $A_{2.2}$ hold, then A_2 holds, where D_t is the set of all distinct t-tuples $(i_1, i_2, \ldots, i_t)$.

Proof: The result follows from the fact that $(1-\pi_i)^2 \sim O(1)$ and the definition of Δ_{ij}.

Result 3.1.2 *Consider*

$$\begin{aligned} A_{3.1}: &\quad \textstyle\sum_{ij\in U_{D_2}} \left(N\pi_{ij}^2\right) = O\left(N^{\eta}\right) \\ A_{3.2}: &\quad \textstyle\sum_{ijkl\in U_{D_4}} \left(\pi_{ijkl} - \pi_{ij}\pi_{kl}\right)^2 = O\left(N^{\eta}\right). \end{aligned} \tag{3.8}$$

If both $A_{3.1}$ and $A_{3.2}$ hold, then A_3 holds.

Proof: Consider all possible cases under condition $A3$.

1. If $i = j = k = l \in U$ then this case is covered under condition $A2$.

2. If $i = j;\ ikl \in U_{D_3}$, then

$$\begin{aligned} \sum_{ikl\in U_{D_3}} \Delta_{iikl}^2 &= \sum_{ikl\in U_{D_3}} \left(\pi_{ikl} - \pi_i\pi_{kl}\right)^2 \\ &\leq \sum_{ikl\in U_{D_3}} \left(\pi_{kl} - \pi_i\pi_{kl}\right)^2 \\ &\leq \sum_{ikl\in U_{D_3}} \pi_{kl}^2 \left(1-\pi_i\right)^2 \\ &\leq N \sum_{kl\in U_{D_2}} \pi_{kl}^2. \end{aligned} \tag{3.9}$$

   This is covered by $A_{3.1}$.

3. If $i = j,\ k = l;\ ik \in U_{D_2}$ then $\Delta_{iikk} = \pi_{ik} - \pi_i\pi_k$, and this case is covered under condition $A2$.

4. If $i = k;\ ijl \in U_{D_3}$ then

$$\begin{aligned} \Delta_{ijil} &= \pi_{ijl} - \pi_{ij}\pi_{il} \\ &\leq \pi_{ij}\left(1-\pi_{il}\right) \\ &\leq \pi_{ij}, \end{aligned}$$

and

$$\sum_{ijl \in U_{D_3}} \Delta_{ijil}^2 \leq \sum_{ijl \in U_{D_2}} \pi_{ij}^2 \leq N \sum_{ij \in U_{D_3}} \pi_{ij}^2. \tag{3.10}$$

This is covered by $A_{3.1}$.

5. If $i = k;\ j = l;\ ij \in U_{D_2}$ then $\Delta_{ijij} = \pi_{ij}(1 - \pi_{ij})$. This is covered by $A_{3.1}$.

6. If $i = l;\ j = k;\ ij \in U_{D_2}$ then $\Delta_{ijji} = \pi_{ij}(1 - \pi_{ij})$. This is covered by $A_{3.1}$.

7. If $i = l;\ ijk \in U_{D_3}$ then

$$\begin{aligned} \Delta_{ijki} &= \pi_{ijk} - \pi_{ij}\pi_{ik} \\ &\leq \pi_{ij}(1 - \pi_{ik}). \end{aligned} \tag{3.11}$$

This is covered by $A_{3.1}$.

8. If $j = k;\ ijl \in U_{D_3}$ then

$$\begin{aligned} \Delta_{ijjl} &= \pi_{ijl} - \pi_{ij}\pi_{jl} \\ &\leq \pi_{ij}(1 - \pi_{jl}). \end{aligned} \tag{3.12}$$

This is covered by $A_{3.1}$.

9. If $j = l;\ ijk \in U_{D_3}$ then

$$\begin{aligned} \Delta_{ijkj} &= \pi_{ijk} - \pi_{ij}\pi_{kj} \\ &\leq \pi_{ij}(1 - \pi_{jk}). \end{aligned} \tag{3.13}$$

This is covered by $A_{3.1}$.

10. If $k = l;\ ijk \in U_{D_3}$ then

$$\begin{aligned} \Delta_{ijkk} &= \pi_{ijk} - \pi_{ij}\pi_{k} \\ &\leq \pi_{ij}(1 - \pi_{k}). \end{aligned} \tag{3.14}$$

This is covered by $A_{3.1}$.

11. If $i = j = k$; $il \in U_{D_2}$ then $\Delta_{iiil} = \pi_{il}(1 - \pi_i)$. This is covered by $A_{3.1}$.

12. If $i = j = l$; $i, k \in U_{D_2}$: $\Delta_{iiki} = \pi_{ik}(1 - \pi_i)$. This is covered by $A_{3.1}$.

13. If $i = k = l$; $ij \in U_{D_2}$ then $\Delta_{ijii} = \pi_{ij}(1 - \pi_i)$. This is covered by $A_{3.1}$.

14. If $j = k = l$; $ij \in U_{D_2}$ then $\Delta_{ijjj} = \pi_{ij}(1 - \pi_j)$. This is covered by $A_{3.1}$.

15. If $ijkl \in U_{D_4}$ then $\Delta_{ijkl} = \pi_{ijkl} - \pi_{ij}\pi_{kl}$. This is condition $A_{3.2}$.

■

Consider the following examples.

Example 3.1.1 Assume $n \sim O\left(N^{\delta}\right)$, for $\frac{5}{6} < \delta \leq 1$. Then $\hat{\theta}_{GHR}$ is a mean square consistent estimator for θ under SI sampling design. If $\delta = 1$, then the finite population correction $\left(fpc = 1 - \frac{n}{N}\right)$ cannot be ignored and we can ignore it if $\delta < 1$.

Example 3.1.2 Consider simple random cluster sampling design (SIC). Under this design, M is the number of clusters, C is the cluster size, $N = MC$ is the population size, and we draw m clusters from the M clusters via SI design and observe all elements in each selected cluster. Assume $m \sim O\left(M^{\delta}\right)$, for $\frac{5}{6} < \delta \leq 1$, then $\hat{\theta}_{GHR}$ is mean square consistent for θ.

Example 3.1.3 For $0 < \delta \leq 1$, $\hat{\theta}_{GHR}$ is a mean square consistent estimator for θ under stratified sampling design, assuming that $H_N \sim O\left(N^{\delta}\right)$ is the number of strata and $N_h \sim O\left(N^{1-\delta}\right)$ is the h^{th} stratum size if $N^{\frac{\delta}{4}}\pi_{N*} \to \infty$. Consider $STSI$ sampling design. Then

$$\pi_{N*} = \min_h \left\{ \frac{n_h}{N_h} \frac{n_h - 1}{N_h - 1} \right\}.$$

If $N^{\frac{\delta}{4}}\pi_{N*} \to \infty$, where $N = \sum_{h=1}^{H_N} N_h$, then $STSI$ is a mean square consistent estimator.

Details for Examples 3.1.1, 3.1.2, and 3.1.3 will be given in technical details appendix.

In the following sections, we will discuss the asymptotic distribution of $\hat{\theta}_{GHR}$ under SI and PO sampling designs.

3.1.1 CLT for $\hat{\theta}_{GHR}$ Under Simple Random Sampling Without Replacement

Under SI sampling design, we will show that $\hat{\theta}_{GHR}$ is asymptotically normal. Further, we will give a consistent variance estimator and show that the Godambe and Joshi (1965) lower bound is asymptotically attainable. Consider the following assumptions:

$$
\begin{aligned}
&D_1: \quad 0 < l_x \leq x_i \leq u_x < \infty \quad \text{and} \quad |y_i| \leq u_y < \infty, \\
&D_2: \quad \lim_{N\to\infty} \bar{r}_{U_N} = \mu_r, \quad \lim_{N\to\infty} \bar{y}_{U_N} = \mu_y, \quad \text{and} \quad \lim_{N\to\infty} \bar{x}_{U_N} = \mu_x. \qquad (3.15) \\
&D_3: \quad \lim_{N\to\infty} S^2_{w,U_N} = \sigma^2_w > 0
\end{aligned}
$$

where

$$
S^2_{w,U_N} = \frac{1}{N-1} \sum_{i\in U_N} \left(w_i - \bar{w}_{U_N}\right)^2.
$$

It follows by using Fuller (1996) Corollary 5.1.1.1 p. 220 that

$$
\begin{aligned}
\bar{x}_{s_N} - \bar{x}_{U_N} &= O_p\left(n^{-1/2}\right), \\
\bar{y}_{s_N} - \bar{y}_{U_N} &= O_p\left(n^{-1/2}\right), \qquad (3.16) \\
\bar{r}_{s_N} - \bar{r}_{U_N} &= O_p\left(n^{-1/2}\right)
\end{aligned}
$$

where $\bar{x}_{s_N} = n^{-1}\sum_s x_k$. It follows from Stuart and Ord (1987) p. 422, Exercise 12.11, and and straightforward computations that

$$
\begin{aligned}
var_p\left[S^2_{w,s_N}\right] &= E_p\left[S^2_{w,s_N} - S^2_{w,U_N}\right]^2 \\
&= c_N \left\{ a_N \left[\frac{1}{N} \sum_{i\in U_N} \left(w_i - \bar{w}_{U_N}\right)^4 \right] \right. \\
&\quad \left. - b_N \left[\frac{1}{N} \sum_{i\in U_N} \left(w_i - \bar{w}_{U_N}\right)^2 \right] \right\}, \quad \text{for} \quad n \geq 4 \qquad (3.17)
\end{aligned}
$$

where

$$S^2_{w,s_N} = \frac{1}{n-1}\sum_{i\in s_N}\left(w_i - \bar{w}_{s_N}\right)^2,$$

$$a_N = \left(1 - \frac{1}{n} - \frac{1}{N} - \frac{1}{nN}\right)\left(1 - \frac{n}{N}\right) = O(1),$$

$$b_N = 1 - \frac{3}{n} + \frac{6}{nN} - \frac{3}{N^2} - \frac{3}{nN^2} = O(1),$$

and

$$\begin{aligned} c_N &= \frac{N^3(N-n)}{(n-1)(N-1)^2(N-2)(N-3)} \\ &= \frac{1}{n-1}\frac{N^4}{(N-1)^2(N-2)(N-3)}\left(1 - \frac{n}{N}\right) \\ &= O\left(\frac{1}{n}\right). \end{aligned} \tag{3.18}$$

It follows from Fuller (1996) Corollary 5.1.1.1 p. 220

$$S^2_{w,s_N} - S^2_{w,U_N} = O_p\left(\frac{1}{\sqrt{n}}\right).$$

Now we are ready for the following theorems,

Theorem 3.1.2 *Under $D_1 - D_3$ and simple random sampling without-replacement,*

$$\widehat{AV}_{SI}\left(\hat{\theta}_{GHR}\right) = \frac{1}{\bar{x}^2_{U_N}}\frac{1}{n}\left(1 - \frac{n}{N}\right)S^2_{y-\bar{r}_s x,s}$$

is a consistent estimator for

$$AV_{SI}\left(\hat{\theta}_{GHR}\right) = \frac{1}{\bar{x}^2_{U_N}}\frac{1}{n}\left(1 - \frac{n}{N}\right)S^2_{y-\bar{r}_U x,U}.$$

Proof: First note that

$$\begin{aligned}
S^2_{y-\bar{r}_{U_N}x,U_N} &= \frac{1}{N-1}\sum_{i\in U_N}\left[(y_i-\bar{y}_{U_N})-\bar{r}_{U_N}(x_i-\bar{x}_{U_N})\right]^2 \\
&= \frac{1}{N-1}\sum_{i\in U_N}\left\{(y_i-\bar{y}_{U_N})^2+\bar{r}^2_{U_N}(x_i-\bar{x}_{U_N})^2\right. \\
&\quad \left.-\bar{r}_{U_N}(y_i-\bar{y}_{U_N})(x_i-\bar{x}_{U_N})\right\} \\
&= S^2_{y,U_N}+\bar{r}^2_{U_N}S^2_{x,U_N}-2\bar{r}_{U_N}S_{xy,U_N} \\
S^2_{y-\bar{r}_{s_N}x,s_N} &= \frac{1}{n-1}\sum_{i\in s_N}\left[(y_i-\bar{y}_{s_N})-\bar{r}_{s_N}(x_i-\bar{x}_{U_N})\right]^2 \\
&= S^2_{y,s_N}+\bar{r}^2_{s_N}S^2_{x,s_N}-2\bar{r}_{s_N}S_{xy,s_N}+\frac{n}{n-1}\bar{r}_{s_N}(\bar{x}_{s_N}-\bar{x}_{U_N})^2
\end{aligned}$$

so that

$$\begin{aligned}
S^2_{y-\bar{r}_{s_N}x,s_N}-S^2_{y-\bar{r}_{U_N}x,U_N} &= \left(S^2_{y,s_N}-S^2_{y,U_N}\right)+(\bar{r}_{s_N}-\bar{r}_{U_N})(\bar{r}_{s_N}+\bar{r}_{U_N})S^2_{x,s_N} \\
&\quad +\bar{r}^2_{U_N}\left(S^2_{x,s_N}-S^2_{x,U_N}\right)+2\bar{r}_{U_N}\left(S_{xy,U_N}-S_{xy,s_N}\right) \\
&\quad -2(\bar{r}_{s_N}-\bar{r}_{U_N})S_{xy,s_N}+\frac{n}{n-1}\bar{r}_{s_N}(\bar{x}_{s_N}-\bar{x}_{U_N})^2 \\
&\quad \xrightarrow{p} 0 \quad \text{as } N,n\to\infty. \qquad (3.19)
\end{aligned}$$

Note that $\left(S_{xy,U_N}-S_{xy,s_N}\right)\xrightarrow{p}0$ as $N,n\to\infty$ by Cauchy-Schwarz inequality. Now

$$\frac{\frac{1}{t_x^2}\frac{N^2}{n}\left(1-\frac{n}{N}\right)S^2_{y-\bar{r}_s x,s}}{\frac{1}{t_x^2}\frac{N^2}{n}\left(1-\frac{n}{N}\right)S^2_{y-\bar{r}_U x,U}}-1=\frac{S^2_{y-\bar{r}_s x,s}-S^2_{y-\bar{r}_U x,U}}{S^2_{y-\bar{r}_U x,U}}\xrightarrow{p}0 \quad \text{as } N,n\to\infty.$$

■

Theorem 3.1.3 *Assume D_1-D_3 and assume that n, $N\to\infty$ and $N-n\to\infty$. Then under simple random sampling without replacement*

$$\frac{\hat{\theta}_{GHR}-\theta}{\sqrt{AV_{SI}\left(\hat{\theta}_{GHR}\right)}}\xrightarrow{\mathcal{L}}\mathcal{N}(0,\ 1) \quad as \quad N\to\infty,$$

where

$$AV_{SI}\left(\hat{\theta}_{GHR}\right)=\frac{1}{\bar{x}^2_{U_N}}\frac{1}{n}\left(1-\frac{n}{N}\right)S^2_{y-\bar{r}_{U_N}x,U_N}.$$

Proof: From Example 2.1.1, $\hat{\theta}_{GHR}$ is the Hartley and Ross (1954) estimator and is given by

$$\hat{\theta}_{GHR} = \bar{r}_{s_N} + \frac{n(N-1)}{N(n-1)\bar{x}_{U_N}}\left(\bar{y}_{s_N} - \bar{r}_{s_N}\bar{x}_{s_N}\right).$$

Define

$$\begin{aligned} d_N &= \frac{n(N-1)}{N(n-1)} \\ 1-d_N &= -\frac{1}{n-1}\left(1-\frac{n}{N}\right) \\ &= O\left(\frac{1}{n}\right). \end{aligned} \tag{3.20}$$

Now

$$\begin{aligned} \hat{\theta}_{GHR} - \theta &= \bar{r}_{s_N} + \frac{d_N}{\bar{x}_{U_N}}\left(\bar{y}_{s_N} - \bar{r}_{s_N}\bar{x}_{s_N}\right) - \frac{\bar{y}_{U_N}}{\bar{x}_{U_N}} \\ &= \bar{r}_{s_N} + \frac{d_N}{\bar{x}_{U_N}}\left(\left(\bar{y}_{s_N} - \bar{y}_{U_N}\right) + \bar{y}_{U_N} - \bar{r}_{s_N}\left(\left(\bar{x}_{s_N} - \bar{x}_{U_N}\right) + \bar{x}_{U_N}\right)\right) - \frac{\bar{y}_{U_N}}{\bar{x}_{U_N}} \\ &= (1-d_N)\bar{r}_s + (d_N-1)\frac{1}{\bar{x}_{U_N}}\left(\bar{y}_{s_N} - \bar{y}_{U_N}\right) + (d_N-1)\frac{\bar{y}_{U_N}}{\bar{x}_{U_N}} \\ &\quad - (d_N-1)\frac{\bar{r}_{s_N}}{\bar{x}_{U_N}}\left(\bar{x}_{s_N} - \bar{x}_{U_N}\right) + \frac{1}{\bar{x}_{U_N}}\left(\bar{y}_{s_N} - \bar{y}_{U_N}\right) \\ &\quad - \frac{\bar{r}_{s_N} - \bar{r}_{U_N} + \bar{r}_{U_N}}{\bar{x}_{U_N}}\left(\bar{x}_{s_N} - \bar{x}_{U_N}\right) \\ &= (1-d_N)\bar{r}_{s_N} + (d_N-1)\frac{1}{\bar{x}_{U_N}}\left(\bar{y}_{s_N} - \bar{y}_{U_N}\right) + (d_N-1)\frac{\bar{y}_{U_N}}{\bar{x}_{U_N}} \\ &\quad - (d_N-1)\frac{\bar{r}_{s_N}}{\bar{x}_{U_N}}\left(\bar{x}_{s_N} - \bar{x}_{U_N}\right) - \frac{1}{\bar{x}_{U_N}}\left(\bar{r}_{s_N} - \bar{r}_{U_N}\right)\left(\bar{x}_{s_N} - \bar{x}_{U_N}\right) \\ &\quad + \frac{1}{\bar{x}_{U_N}}\left(\bar{y}_{s_N} - \bar{y}_{U_N}\right) - \frac{\bar{r}_{U_N}}{\bar{x}_{U_N}}\left(\bar{x}_{s_N} - \bar{x}_{U_N}\right) \\ &= O\left(\frac{1}{n}\right)O_p(1) + O\left(\frac{1}{n}\right)o_p\left(n^{-\frac{1}{2}}\right) + O\left(\frac{1}{n}\right)O_p(1) \\ &\quad + O\left(\frac{1}{n}\right)o_p\left(n^{-\frac{1}{2}}\right) + o_p\left(n^{-1}\right) + \frac{1}{\bar{x}_{U_N}}\left(\bar{y}_{s_N} - \bar{y}_{U_N}\right) - \frac{\bar{r}_{U_N}}{\bar{x}_{U_N}}\left(\bar{x}_{s_N} - \bar{x}_{U_N}\right) \\ &= o_p\left(n^{-1}\right) + \frac{1}{t_x}\sum_{k\in s}\left[\frac{y_{Nk} - \bar{y}_{U_N} - \bar{r}_{U_N}\left(x_{Nk} - \bar{x}_{U_N}\right)}{n/N}\right] \\ &= o_p\left(n^{-1}\right) + \frac{1}{\bar{x}_{U_N}}\sum_{k\in s}\left[\frac{y_{Nk} - \bar{y}_{U_N} - \bar{r}_{U_N}\left(x_{Nk} - \bar{x}_{U_N}\right)}{n}\right] \\ &= o_p\left(n^{-1}\right) + \frac{1}{\bar{x}_{U_N}}\bar{w}_N. \end{aligned} \tag{3.21}$$

Since

$$\begin{aligned} |y_{Nk} - \bar{y}_{U_N} - \bar{r}_{U_N}(x_{Nk} - \bar{x}_{U_N})| &\leq |y_{Nk}| + |\bar{y}_{U_N}| + |\bar{r}_{U_N}|\,[|x_{Nk}| + |\bar{x}_{U_N}|] \\ &\leq 2u_y + 2\frac{u_y}{l_x}u_x < \infty \quad \forall k, \end{aligned}$$

it follows that the Lindeberg-type condition,

$$\lim_{N\to\infty} \frac{\sum_{k\in U_N} w_{Nk}^2 I_{\{|w_{Nk}|>\epsilon c_N\}}}{\sum_{k\in U_N} w_{Nk}^2} = 0 \quad \text{for all} \quad \epsilon > 0. \tag{3.22}$$

holds (Ash (2000)). Hence, by Hájek (1960) Theorem 3.1, $\bar{w}_N$ has an asymptotic normal distribution and so

$$\frac{\sqrt{n}\left(\hat{\theta}_{GHR} - \theta\right)}{\sqrt{\frac{1}{\bar{x}_{U_N}^2}\left(1 - \frac{n}{N}\right) S_{y-\bar{r}_U x,U}^2}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,\ 1),$$

where

$$\begin{aligned} AV_{SI}\left(\hat{\theta}_{GHR}\right) &= var_{SI}\left\{\frac{1}{t_x}\sum_{k\in s}\left[\frac{y_k - \bar{y}_{U_N} - \bar{r}_{U_N}(x_k - \bar{x}_{U_N})}{n/N}\right]\right\} \\ &= \frac{1}{\bar{x}_{U_N}^2}\frac{1}{n}\left(1 - \frac{n}{N}\right) S_{y-\bar{r}_U x,U}^2. \end{aligned} \tag{3.23}$$

■

Corollary 3.1.1 *Under conditions of Theorem 3.1.3,*

$$\frac{\hat{\theta}_{GHR} - \theta}{\sqrt{\widehat{AV}_{SI}\left(\hat{\theta}_{GHR}\right)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,\ 1) \quad N \to \infty, \quad as \quad N \to \infty \tag{3.24}$$

Proof: Since

$$\frac{\hat{\theta}_{GHR} - \theta}{\sqrt{AV_{SI}\left(\hat{\theta}_{GHR}\right)}} = \frac{\sqrt{\widehat{AV}_{SI}\left(\hat{\theta}_{GHR}\right)}}{\sqrt{AV_{SI}\left(\hat{\theta}_{GHR}\right)}} \frac{\hat{\theta}_{GHR} - \theta}{\sqrt{\widehat{AV}_{SI}\left(\hat{\theta}_{GHR}\right)}},$$

then the result follows from Theorems 3.1.2 and 3.1.3.

■

Result 3.1.3 *Under the model,*

$$\xi : y_i \text{ are independent } \left(\beta x_i, \sigma_i^2\right) \qquad \sigma_i^2 > 0, \ \forall i,$$

$E_\xi\left[AV_{SI}\left(\hat{\theta}_{GHR}\right)\right]$ *asymptotically attains the Godambe and Joshi (1965) lower bound.*

Proof: It follows from the model that

$$\left.\begin{array}{ll} E_\xi\left(\bar{y}_{U_N}\right) = \beta \bar{x}_{U_N} & E_\xi\left(\bar{r}_U\right) = \beta \\ var_\xi\left(\bar{y}_{U_N}\right) = \frac{1}{N^2}\sum_{i\in U}\sigma_i^2 & var_\xi\left(\bar{r}_U\right) = \frac{1}{N^2}\sum_{i\in U}\frac{\sigma_i^2}{x_i^2} \end{array}\right\} \tag{3.25}$$

Recall that, $w_i = y_i - \bar{y}_{U_N} - \bar{r}_U\left(x_i - \bar{x}_{U_N}\right)$, then $E_\xi\left(w_i\right) = 0$, and

$$\sum_{i\in U} var_\xi\left(w_i\right) = \frac{N-1}{N}\sum_{i\in U}\sigma_i^2 - \left[\frac{1}{N}\sum_{i\in U}\left(x_i - \bar{x}_{U_N}\right)^2\right]\left[\frac{1}{N}\sum_{j\in U}\frac{\sigma_i^2}{x_j^2}\right].$$

Therefore, the model expectation of the approximate design variance is given by

$$\begin{aligned} E_\xi\left[\frac{1}{t_x^2}\frac{N^2}{n}\left(1-\frac{n}{N}\right)S^2_{y-\bar{r}_U x,U}\right] &= E_\xi\left[\frac{1}{\bar{x}_{U_N}^2}\frac{1}{n}\left(1-\frac{n}{N}\right)\frac{1}{N-1}\sum_{i\in U} var_\xi\left(w_i\right)\right] \\ &= \frac{1}{\bar{x}_{U_N}^2}\frac{1}{n}\left(1-\frac{n}{N}\right)\left\{\frac{1}{N}\sum_{i\in U}\sigma_i^2\right. \\ &- \left.\frac{1}{N-1}\left[\frac{1}{N}\sum_{i\in U}\left(x_i-\bar{x}_{U_N}\right)^2\right]\left[\frac{1}{N}\sum_{j\in U}\frac{\sigma_i^2}{x_j^2}\right]\right\} \\ &= GJLB + O\left(\frac{1}{nN}\right). \end{aligned}$$

Since GJLB is of exact order $O\left(n^{-1}\right)$, it follows that

$$\frac{E_\xi\left[\frac{1}{t_x^2}\frac{N^2}{n}\left(1-\frac{n}{N}\right)S^2_{y-\bar{r}_U x,U}\right]}{GJLB} = 1 + O\left(N^{-1}\right) \to 1 \quad \text{as} \quad N \to \infty,$$

and so $\hat{\theta}_{GHR}$ asymptotically attains the GJLB.

■

From Section 2.4.1, recall the discussions of approximating $\hat{v}ar_p\left(\hat{\theta}_{GHR}\right)$. It was given that

$$\hat{v}ar_{app}\left(\hat{\theta}_{GHR}\right) = \frac{1}{N^2 t_x^2}\sum_{ij\in s}\frac{\hat{w}_i}{\pi_i}\frac{\hat{w}_j}{\pi_j}\frac{\Delta_{ij}}{\pi_{ij}}, \tag{3.26}$$

where

$$\hat{w}_i = (N-1)\,y_i + \frac{y_i}{\pi_i} - \hat{t}_{r\pi}x_i$$

will be used to approximate $\hat{v}ar_p\left(\hat{\theta}_{GHR}\right)$. Under SI, this method of approximation is asymptotically equivalent to $\widehat{AV}_{SI}\left(\hat{\theta}_{GHR}\right)$.

Result 3.1.4 *Under conditions of Theorem 3.1.3,*

$$\begin{aligned}\hat{v}ar_{app}\left(\hat{\theta}_{GHR}\right) &\overset{SI}{=} \frac{1}{N^2\bar{x}_{U_N}^2}\frac{1}{n}\left(1-\frac{n}{N}\right)\left[\frac{1}{n-1}\sum_{i\in s}\left(\hat{w}_i-\bar{\hat{w}}\right)^2\right]\\ &= \widehat{AV}_{SI}\left(\hat{\theta}_{GHR}\right)+O_p\left(\frac{1}{n^2}\right).\end{aligned}$$

Proof: Since

$$\hat{w}_i-\bar{\hat{w}} \overset{SI}{=} N\left[\left(y_i-\bar{y}_{s_N}\right)-\left(x_i-\bar{x}_{s_N}\right)\bar{r}_{s_N}+\frac{1}{n}\left(1-\frac{n}{N}\right)\left(y_i-\bar{y}_{s_N}\right)\right];$$

therefore,

$$\begin{aligned}\sum_{i\in s}\left(\hat{w}_i-\bar{\hat{w}}\right)^2 &= N^2\sum_{i\in s}\left[\left(y_i-\bar{y}_{s_N}\right)-\left(x_i-\bar{x}_{s_N}\right)\bar{r}_{s_N}\right]^2\\ &+ \frac{N^2}{n^2}\left(1-\frac{n}{N}\right)^2\sum_{i\in s}\left(y_i-\bar{y}_{s_N}\right)^2\\ &+ \frac{N^2}{n}\left(1-\frac{n}{N}\right)\left\{\sum_{i\in s}\left(y_i-\bar{y}_{s_N}\right)^2-\bar{r}_{s_N}\sum_{i\in s}\left(y_i-\bar{y}_{s_N}\right)\left(x_i-\bar{x}_{s_N}\right)\right\}.\end{aligned}$$

Hence,

$$\begin{aligned}\hat{v}ar_{app}\left(\hat{\theta}_{GHR}\right) &\overset{SI}{=} \frac{1}{N^2\bar{x}_{U_N}^2}\frac{1}{n}\left(1-\frac{n}{N}\right)\left[\frac{1}{n-1}\sum_{i\in s}\left(\hat{w}_i-\bar{\hat{w}}\right)^2\right]\\ &= \frac{1}{\bar{x}_{U_N}^2}\frac{1}{n}\left(1-\frac{n}{N}\right)\left[\frac{1}{n-1}\sum_{i\in s}\left[\left(y_i-\bar{y}_{s_N}\right)-\left(x_i-\bar{x}_{s_N}\right)\bar{r}_{s_N}\right]^2\right]\end{aligned}$$

$$
\begin{aligned}
&+ \frac{1}{\bar{x}_{U_N}^2}\frac{1}{n^3}\left(1-\frac{n}{N}\right)^3\left[\frac{1}{n-1}\sum_{i\in s}\left(y_i-\bar{y}_{s_N}\right)^2\right] \\
&+ \frac{1}{\bar{x}_{U_N}^2}\frac{2}{n^2}\left(1-\frac{n}{N}\right)^2\left[\frac{1}{n-1}\sum_{i\in s}\left(y_i-\bar{y}_{s_N}\right)^2\right] \\
&- \frac{\bar{r}_{s_N}}{\bar{x}_{U_N}^2}\frac{2}{n^2}\left(1-\frac{n}{N}\right)^2\left[\frac{1}{n-1}\sum_{i\in s}\left(y_i-\bar{y}_{s_N}\right)\left(x_i-\bar{x}_{s_N}\right)\right] \\
&= \widehat{AV}_{SI}\left(\hat{\theta}_{GHR}\right)+O_p\left(\frac{1}{n^3}\right)+O_p\left(\frac{1}{n^2}\right)++O_p\left(\frac{1}{n^2}\right) \\
&= \underbrace{\widehat{AV}_{SI}\left(\hat{\theta}_{GHR}\right)}_{O_p\left(\frac{1}{n}\right)}+O_p\left(\frac{1}{n^2}\right).
\end{aligned}
$$

■

In the following sections we will discuss the asymptotic results of $\hat{\theta}_{GHR}$ under PO; $\hat{\theta}_{GHR}$ is asymptotically normal. Also, a consistent variance estimator is given, and the Godambe and Joshi (1965) lower bound is asymptotically attainable for the asymptotic design variance.

3.1.2 CLT for $\hat{\theta}_{GHR}$ Under Poisson Sampling

From Example 2.1.3, recall the definition of $\hat{\theta}_{GHR}$ under PO sampling design,

$$
\hat{\theta}_{GHR}=\frac{1}{N}\hat{t}_{r\pi}+\frac{N-1}{Nt_x}\hat{t}_{y\pi}+\frac{1}{N\bar{x}_{U_N}}\hat{t}_{\check{y}\pi}-\frac{1}{Nt_x}\hat{t}_{r\pi}\hat{t}_{x\pi} \tag{3.27}
$$

where $\check{y}_k=y_k/\left(N\pi_k\right)$. Consider the following assumptions

$E_1: 0<l_x\leq x_i\leq u_x<\infty$ and $|y_i|\leq u_y<\infty$.

$E_2: 0<\pi_{N*}\leq\pi_i\leq\pi_N^*<1$.

$E_3: N\pi_{*N}^3\left(1-\pi_N^*\right)^2\to\infty \quad \text{as} \quad N\to\infty$.

$E_4: \liminf_{N\to\infty}\frac{1}{N}\sum_{k\in U_N}\left(y_k-\bar{r}_{U_N}x_k\right)^2>0$.

The term $N^{-1}\hat{t}_{\check{y}\pi}$ in (3.27) asymptotically is ignorable as we will see in the following lemma.

Lemma 3.1.1 *Under $E_1 - E_4$ and under Poisson sampling,*

$$\frac{1}{N}\hat{t}_{\check{y}\pi} = O\left(\frac{1}{N\pi_{N*}}\right).$$

Proof: Since

$$\begin{aligned}
E_p\left[\frac{1}{N}\hat{t}_{\check{y}\pi}\right] &= \frac{1}{N}\sum_{k\in U_N}\frac{y_k}{N\pi_k} \\
&\leq \frac{1}{N\pi_{N*}}\left|\frac{1}{N}\sum_{k\in U_N} y_k\right| \\
&= O_p\left(\frac{1}{N\pi_{N*}}\right) \\
var_p\left[\frac{1}{N}\sum_{k\in U_N}\frac{y_k}{N\pi_k}\frac{I_{\{k\in s\}}}{\pi_k}\right] &= \frac{1}{N^2}\sum_{k\in U_N}\left(\frac{y_k}{N\pi_k}\right)^2\frac{1-\pi_k}{\pi_k} \\
&\leq \frac{1}{N^3\pi_{N*}^3}\left[\frac{1}{N}\sum_{k\in U_N} y_k^2\right] \\
&= O\left(\frac{1}{N^3\pi_{N*}^3}\right). \qquad (3.28)
\end{aligned}$$

Therefore,

$$E_p\left[\frac{1}{N}\hat{t}_{\check{y}\pi}\right]^2 = var_p\left[\frac{1}{N}\sum_{k\in U_N}\frac{y_k}{N\pi_k}\frac{I_{\{k\in s\}}}{\pi_k}\right] + \left[E_p\left[\frac{1}{N}\hat{t}_{\check{y}\pi}\right]\right]^2 = O_p\left(\frac{1}{(N\pi_{N*})^2}\right).$$

The Lemma follows by using Fuller (1996) Corollary 5.1.1.1 p. 220.

■

Lemma 3.1.2 *Under $E_1 - E_4$ and under Poisson sampling design,*

$$\frac{\widehat{AV}_{PO}\left(\hat{\theta}_{GHR}\right)}{AV_{PO}\left(\hat{\theta}_{GHR}\right)} \xrightarrow{p} 1 \quad as \quad N \to \infty,$$

where

$$\widehat{AV}_{PO}\left(\hat{\theta}_{GHR}\right) = \frac{1}{N^2\bar{x}_{U_N}^2}\sum_{i\in U_N}\left(y_i - \frac{1}{N}\hat{t}_{r\pi}x_i\right)^2\frac{1-\pi_{iN}}{\pi_{iN}}\frac{I_{\{i\in s\}}}{\pi_{iN}} \qquad (3.29)$$

and

$$AV_{PO}\left(\hat{\theta}_{GHR}\right) = \frac{1}{N^2\bar{x}_{U_N}^2}\sum_{i\in U_N}\left(y_i - \bar{r}_{U_N}x_i\right)^2\frac{1-\pi_{iN}}{\pi_{iN}}. \tag{3.30}$$

Proof: Since

$$\begin{aligned}
N^2\bar{x}_{U_N}^2\widehat{AV}_{PO}\left(\hat{\theta}_{GHR}\right) &= \sum_{i\in U_N}\left(y_i - \frac{1}{N}\hat{t}_{r\pi}x_i\right)^2\frac{1-\pi_{iN}}{\pi_{iN}}\frac{I_{\{i\in s\}}}{\pi_{iN}} \\
&= \sum_{i\in U_N}\left[y_i - \bar{r}_{U_N}x_i - \left(\frac{1}{N}\hat{t}_{r\pi} - \bar{r}_{U_N}\right)x_i\right]^2\frac{1-\pi_{iN}}{\pi_{iN}}\frac{I_{\{i\in s\}}}{\pi_{iN}} \\
&= \sum_{i\in U_N}\left(y_i - \bar{r}_{U_N}x_i\right)^2\frac{1-\pi_{iN}}{\pi_{iN}}\frac{I_{\{i\in s\}}}{\pi_{iN}} \\
&\quad + \left(\frac{1}{N}\hat{t}_{r\pi} - \bar{r}_{U_N}\right)^2\sum_{i\in U_N}x_i^2\frac{1-\pi_{iN}}{\pi_{iN}}\frac{I_{\{i\in s\}}}{\pi_{iN}} \\
&\quad -2\left(\frac{1}{N}\hat{t}_{r\pi} - \bar{r}_{U_N}\right)\sum_{i\in U_N}\left(y_i - \bar{r}_{U_N}x_i\right)x_i\frac{1-\pi_{iN}}{\pi_{iN}}\frac{I_{\{i\in s\}}}{\pi_{iN}}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\widehat{AV}_{PO}\left(\hat{\theta}_{GHR}\right) - AV_{PO}\left(\hat{\theta}_{GHR}\right) &= \frac{1}{N^2\bar{x}_{U_N}^2}\sum_{i\in U_N}\left\{\left(y_i - \bar{r}_{U_N}x_i\right)^2\frac{1-\pi_{iN}}{\pi_{iN}}\left[\frac{I_{\{i\in s\}}}{\pi_{iN}} - 1\right]\right. \\
&\quad + \left(\frac{1}{N}\hat{t}_{r\pi} - \bar{r}_{U_N}\right)^2 x_i^2\frac{1-\pi_{iN}}{\pi_{iN}}\frac{I_{\{i\in s\}}}{\pi_{iN}} \\
&\quad - \left.2\left(\frac{1}{N}\hat{t}_{r\pi} - \bar{r}_{U_N}\right)\left(y_i - \bar{r}_{U_N}x_i\right)x_i\frac{1-\pi_{iN}}{\pi_{iN}}\frac{I_{\{i\in s\}}}{\pi_{iN}}\right\}.
\end{aligned}$$

Hence

$$\begin{aligned}
\left|\frac{\widehat{AV}_{PO}\left(\hat{\theta}_{GHR}\right)}{AV_{PO}\left(\hat{\theta}_{GHR}\right)} - 1\right| &\leq \left|\frac{\sum_{i\in U_N}\left(y_i - \bar{r}_{U_N}x_i\right)^2\frac{1-\pi_{iN}}{\pi_{iN}}\left[\frac{I_{\{i\in s\}}}{\pi_{iN}} - 1\right]}{\sum_{i\in U_N}\left(y_i - \bar{r}_{U_N}x_i\right)^2\frac{1-\pi_{iN}}{\pi_{iN}}}\right| \\
&\quad + \left(\frac{1}{N}\hat{t}_{r\pi} - \bar{r}_{U_N}\right)^2\frac{\sum_{i\in U_N}x_i^2\frac{1-\pi_{iN}}{\pi_{iN}}\frac{I_{\{i\in s\}}}{\pi_{iN}}}{\sum_{i\in U_N}\left(y_i - \bar{r}_{U_N}x_i\right)^2\frac{1-\pi_{iN}}{\pi_{iN}}} \\
&\quad + 2\left|\frac{1}{N}\hat{t}_{r\pi} - \bar{r}_{U_N}\right|\frac{\left|\sum_{i\in U_N}\left(y_i - \bar{r}_{U_N}x_i\right)x_i\frac{1-\pi_{iN}}{\pi_{iN}}\frac{I_{\{i\in s\}}}{\pi_{iN}}\right|}{\sum_{i\in U_N}\left(y_i - \bar{r}_{U_N}x_i\right)^2\frac{1-\pi_{iN}}{\pi_{iN}}} \\
&= A_N + B_N + C_N. \qquad (3.31)
\end{aligned}$$

Now,

$$\begin{aligned} E_p\left(A_N^2\right) &= \frac{var_p\left[\sum_{i\in U_N}\left(y_i-\bar{r}_{U_N}x_i\right)^2\frac{1-\pi_{iN}}{\pi_{iN}}\frac{I_{\{i\in s\}}}{\pi_{iN}}\right]}{\left[\sum_{i\in U_N}\left(y_i-\bar{r}_{U_N}x_i\right)^2\frac{1-\pi_{iN}}{\pi_{iN}}\right]^2} \\ &\leq \frac{1}{N\pi_{N*}^3\left(1-\pi_N^*\right)^2}\frac{\sum_{i\in U_N}\left(y_i-\bar{r}_{U_N}x_i\right)^4/N}{\left[\sum_{i\in U_N}\left(y_i-\bar{r}_{U_N}x_i\right)^2/N\right]^2} \\ &\to 0 \quad \text{as} \quad N\to\infty, \qquad \text{by } E_3 \end{aligned} \tag{3.32}$$

since $E_p\left(A_N^2\right)\to 0 \Rightarrow E_p\left|A_N\right|\to 0$ as $N\to\infty$. Next,

$$\begin{aligned} B_N &= \left(\frac{1}{N}\hat{t}_{r\pi}-\bar{r}_{U_N}\right)^2\frac{\sum_{i\in U_N}x_i^2\frac{1-\pi_{iN}}{\pi_{iN}}\frac{I_{\{i\in s\}}}{\pi_{iN}}}{\sum_{i\in U_N}\left(y_i-\bar{r}_{U_N}x_i\right)^2\frac{1-\pi_{iN}}{\pi_{iN}}} \\ &\leq \frac{1}{\pi_{N*}^2\left(1-\pi_N^*\right)}\frac{\sum_{i\in U_N}x_i^2/N}{\left[\sum_{i\in U_N}\left(y_i-\bar{r}_{U_N}x_i\right)^2/N\right]}\left(\frac{1}{N}\hat{t}_{r\pi}-\bar{r}_{U_N}\right)^2 \\ E_p\left(B_N\right) &\leq \frac{1}{\pi_{N*}^2\left(1-\pi_N^*\right)}\frac{\sum_{i\in U_N}x_i^2/N}{\left[\frac{1}{N}\sum_{i\in U_N}\left(y_i-\bar{r}_{U_N}x_i\right)^2\right]}var_p\left(\frac{1}{N}\hat{t}_{r\pi}\right) \\ &\to 0 \text{ as } N\to\infty, \end{aligned} \tag{3.33}$$

since

$$\begin{aligned} var_p\left(\frac{1}{N}\hat{t}_{r\pi}\right) &= \frac{1}{N^2}\sum_{ij\in U}\frac{r_i}{\pi_i}\frac{r_j}{\pi_j}\Delta_{ij} \\ &= \frac{1}{N^2}\sum_{i\in U}\left(\frac{r_i}{\pi_i}\right)^2\pi_i\left(1-\pi_i\right)+\frac{1}{N^2}\sum_{i\neq j\in U}\frac{r_i}{\pi_i}\frac{r_j}{\pi_j}\Delta_{ij} \\ &\leq \frac{1}{N\pi_{N*}}\left[\frac{1}{N}\sum_{i\in U}r_i^2\right]+0 \\ &= O\left(\frac{1}{N\pi_{N*}}\right), \end{aligned}$$

and

$$N\pi_{N*}^3\left(1-\pi_N^*\right)^2\leq N\pi_{N*}^3\left(1-\pi_N^*\right).$$

Since $E_p\left(A_N^2\right)\to 0$, $E_p\left(B_N\right)\to 0$, then by Cauchy-Schwarz inequality $E_p\left(C_N\right)\to 0$ as $N\to\infty$, and $E_p\left|\frac{\widehat{AV}_{PO}\left(\hat{\theta}_{GHR}\right)}{AV_{PO}\left(\hat{\theta}_{GHR}\right)}-1\right|\to 0$, hence $\frac{\widehat{AV}_{PO}\left(\hat{\theta}_{GHR}\right)}{AV_{PO}\left(\hat{\theta}_{GHR}\right)}-1\xrightarrow{p}0$ as $N\to\infty$.

■

Theorem 3.1.4 *Under conditions $E_1 - E_4$ and Poisson sampling design,*

$$\frac{\hat{\theta}_{GHR} - \theta}{\sqrt{\widehat{AV}_{PO}\left(\hat{\theta}_{GHR}\right)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,\ 1) \quad as\ N \to \infty.$$

Proof: Recall the definition of $\hat{\theta}_{GHR}$ under PO sampling design from equation (3.27); therefore,

$$\begin{aligned}
\hat{\theta}_{GHR} - \theta &= \frac{1}{\bar{x}_{U_N}} \frac{\hat{t}_{y\pi} - t_y}{N} - \frac{1}{N\bar{x}_{U_N}} \left(\frac{1}{N}\hat{t}_{y\pi}\right) + \frac{1}{\bar{x}_{U_N}} \left(\frac{1}{N}\hat{t}_{\check{y}\pi}\right) - \frac{1}{\bar{x}_{U_N}} \frac{\hat{t}_{r\pi} - t_r}{N} \frac{\hat{t}_{x\pi} - t_x}{N} \\
&\quad - \frac{\bar{r}_{U_N}}{\bar{x}_{U_N}} \frac{\hat{t}_{x\pi} - t_x}{N} \\
&= O_p\left(\frac{1}{N\pi_{N*}}\right) + \frac{1}{\bar{x}_{U_N}} \frac{\hat{t}_{y\pi} - t_y}{N} - \frac{\bar{r}_{U_N}}{\bar{x}_{U_N}} \frac{\hat{t}_{x\pi} - t_x}{N} \\
&= O_p\left(\frac{1}{N\pi_{N*}}\right) + \frac{1}{N\bar{x}_{U_N}} \sum_{i \in U} (y_i - \bar{r}_{U_N} x_i) \left(\frac{I_{\{i \in s\}}}{\pi_{iN}} - 1\right) \qquad (3.34)
\end{aligned}$$

Note that $(y_i - \bar{r}_{U_N} x_i)\left(I_{\{i \in s\}} \pi_{iN}^{-1} - 1\right)$ are independent with mean 0 and variance $(y_i - \bar{r}_{U_N} x_i)^2 (1 - \pi_{iN}) \pi_{iN}^{-1}$. Define

$$\begin{aligned}
C_N^2 &= var_p \left[\sum_{i \in U} (y_i - \bar{r}_{U_N} x_i) \left(\frac{I_{\{i \in s\}}}{\pi_{iN}} - 1\right)\right] \\
&= \sum_{i \in U} (y_i - \bar{r}_{U_N} x_i)^2 \frac{1 - \pi_{iN}}{\pi_{iN}} \\
&\geq \left(\liminf_{N \to \infty} \frac{1}{N} \sum_{k \in U_N} (y_k - \bar{r}_{U_N} x_k)^2\right) \frac{N(1 - \pi_N^*)}{1} \\
&\geq \left(\liminf_{N \to \infty} \frac{1}{N} \sum_{k \in U_N} (y_k - \bar{r}_{U_N} x_k)^2\right) N \pi_{N*}^3 (1 - \pi_N^*)^2 \\
&\to \infty \quad \text{as} \quad N \to \infty \qquad (3.35)
\end{aligned}$$

by E_3 and E_4. Now,

$$\begin{aligned}
\sum_{i \in U_N} E_p \left[(y_i - \bar{r}_{U_N} x_i) \left(\frac{I_{\{i \in s\}}}{\pi_{iN}} - 1\right)\right]^4 &= \sum_{i \in U_N} (y_i - \bar{r}_{U_N} x_i)^4 \left\{\left(\frac{1}{\pi_{iN}} - 1\right)^4 \pi_{iN} \right. \\
&\quad \left. + \left(\frac{0}{\pi_{iN}} - 1\right)^4 (1 - \pi_{iN})\right\}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i\in U_N} (y_i - \bar{r}_{U_N} x_i)^4 \left\{ \frac{(1-\pi_{iN})^4}{\pi_{iN}^3} \right. \\
&\quad \left. + (1-\pi_{iN}) \right\}, \\
&\leq \sum_{i\in U_N} (y_i - \bar{r}_{U_N} x_i)^4 \left\{ \frac{1}{\pi_{N*}^3} + 1 \right\} \\
&\leq \frac{2}{\pi_{N*}^3} \sum_{i\in U_N} (y_i - \bar{r}_{U_N} x_i)^4
\end{aligned}
$$

$$
\frac{1}{C_N^4} \leq \frac{1}{N^2 (1-\pi_N^*)^2 \left[\liminf_{N\to\infty} \frac{1}{N} \sum_{i\in U_N} (y_i - \bar{r}_{U_N} x_i)^2 / N \right]^2}
$$

so that

$$
\begin{aligned}
\frac{\sum_{i\in U_N} E_p \left[(y_i - \bar{r}_{U_N} x_i) \left(\frac{I_{\{i\in s\}}}{\pi_{iN}} - 1 \right) \right]^4}{C_N^4} &\leq \frac{1}{N\pi_{N*}^3 (1-\pi_N^*)^2} \\
&\times \frac{2 \sum_{i\in U_N} (y_i - \bar{r}_{U_N} x_i)^4 / N}{\left[\liminf_{N\to\infty} \sum_{i\in U_N} (y_i - \bar{r}_{U_N} x_i)^2 / N \right]^2} \\
&\to 0 \quad \text{as} \quad N \to \infty, \qquad (3.36)
\end{aligned}
$$

the Lyapunov condition holds for $\delta = 2$. Using Ash (2000) p.309, Lyapunov's condition implies that

$$
C_N^{-1} \sum_{i\in U_N} (y_i - \bar{r}_{U_N} x_i) \left(\frac{I_{\{i\in s\}}}{\pi_{iN}} - 1 \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \quad \text{as} \quad N \to \infty.
$$

From Lemma 3.1.2,

$$
\widehat{AV}_{PO} \left(\hat{\theta}_{GHR} \right) = \frac{1}{N^2 \bar{x}_{U_N}^2} \sum_{i\in U_N} (y_i - \bar{r}_{s_N} x_i)^2 \frac{1-\pi_{iN}}{\pi_{iN}} \frac{I_{\{i\in s\}}}{\pi_{iN}}
$$

is a consistent estimator for

$$
AV_{PO} \left(\hat{\theta}_{GHR} \right) = \frac{1}{N^2 \bar{x}_{U_N}^2} \sum_{i\in U_N} (y_i - \bar{r}_{U_N} x_i)^2 \frac{1-\pi_{iN}}{\pi_{iN}},
$$

and hence the theorem follows.

■

Result 3.1.5 *Under the model*

$$\xi : y_i = \beta x_i + \epsilon_i, \quad \textit{where} \quad \epsilon_i \quad \textit{are independent} \ \left(0,\ \sigma_i^2\right),$$

$E_\xi\left[AV_{PO}\left(\hat{\theta}_{GHR}\right)\right]$ *asymptotically attains the Godambe and Joshi (1965) lower bound.*

Proof: From (3.30), recall the definition of

$$AV_{PO}\left(\hat{\theta}_{GHR}\right) = \frac{1}{N^2\bar{x}_{U_N}^2}\sum_{i\in U_N}\left(y_i - \frac{1}{N}\hat{t}_{r\pi}x_i\right)^2\frac{1-\pi_{iN}}{\pi_{iN}}.$$

Since

$$E_\xi\left(y_i - \bar{r}_{U_N}x_i\right) = 0$$

then the model expectation of the approximate design variance is given by

$$\begin{aligned}
E_\xi\left[AV_{PO}\left(\hat{\theta}_{GHR}\right)\right] &= \frac{1}{t_x^2}\sum_{i\in U_N}\left[var_\xi\left(y_i - \bar{r}_{U_N}x_i\right)\right]\frac{1-\pi_{iN}}{\pi_{iN}} \\
&= \frac{1}{t_x^2}\sum_{i\in U_N}\frac{1-\pi_{iN}}{\pi_{iN}}\left\{\sigma_i^2 - \frac{2}{N}\sigma_i^2 + \frac{1}{N}x_i^2\left[\frac{1}{N}\sum_{j\in U_N}\frac{\sigma_j^2}{x_j^2}\right]\right\} \\
&= \frac{1}{t_x^2}\sum_{i\in U_N}\frac{1-\pi_{iN}}{\pi_{iN}}\sigma_i^2 - \frac{2}{N\bar{x}_{U_N}^2}\left[\frac{1}{N}\sum_{i\in U_N}\frac{1-\pi_{iN}}{N\pi_{iN}}\sigma_i^2\right] \\
&\quad + \frac{1}{N\bar{x}_{U_N}^2}\left[\frac{1}{N}\sum_{i\in U_N}\frac{1-\pi_{iN}}{N\pi_{iN}}x_i^2\right]\left[\frac{1}{N}\sum_{j\in U_N}\frac{\sigma_j^2}{x_j^2}\right] \\
&\leq GJLB + \frac{2}{N^2\pi_{N*}\bar{x}_{U_N}^2}\left[\frac{1}{N}\sum_{i\in U_N}\sigma_i^2\right] \\
&\quad + \frac{1}{N^2\pi_{N*}\bar{x}_{U_N}^2}\left[\frac{1}{N}\sum_{i\in U_N}x_i^2\right]\left[\frac{1}{N}\sum_{j\in U_N}\frac{\sigma_j^2}{x_j^2}\right] \\
&= GJLB + O\left(\frac{1}{N^2\pi_{N*}}\right). \qquad (3.37)
\end{aligned}$$

Since

$$GJLB = \frac{1}{t_x^2}\sum_{i\in U_N}\frac{1-\pi_{iN}}{\pi_{iN}}\sigma_i^2 \quad \text{has order} \quad O\left(\frac{1}{N\pi_{N*}}\right),$$

then

$$\frac{E_\xi\left[AV_{PO}\left(\hat{\theta}_{GHR}\right)\right]}{GJLB} = 1 + O\left(\frac{1}{N}\right) \to 1 \quad \text{as } N \to \infty.$$

■

From equation 3.26,

$$\begin{aligned}\hat{v}ar_{app}\left(\hat{\theta}_{GHR}\right) &= \frac{1}{N^2 t_x^2}\sum_{ij\in s}\frac{\hat{w}_i}{\pi_i}\frac{\hat{w}_j}{\pi_j}\frac{\Delta_{ij}}{\pi_{ij}} \\ &\overset{PO}{=} \frac{1}{N^4\bar{x}_{U_N}^2}\sum_{i\in s}\frac{1-\pi_i}{\pi_i^2}\hat{w}_i^2,\end{aligned}$$

where

$$\hat{w}_i = N\left[\left(y_i - \frac{1}{N}\hat{t}_{r\pi}x_i\right) + \frac{1-\pi_i}{N\pi_i}y_i\right].$$

Result 3.1.6 *Under* $E_1 - E_4$,

$$\hat{v}ar_{app}\left(\hat{\theta}_{GHR}\right) \overset{PO}{=} \widehat{AV}_{PO}\left(\hat{\theta}_{GHR}\right) + O\left(\frac{1}{N^2\pi_*^3}\right).$$

Proof: Since

$$\begin{aligned}\hat{v}ar_{app}\left(\hat{\theta}_{GHR}\right) &\overset{PO}{=} \frac{1}{N^4\bar{x}_{U_N}^2}\sum_{i\in s}\frac{1-\pi_i}{\pi_i^2}\hat{w}_i^2 \\ &= \frac{1}{N^2\bar{x}_{U_N}^2}\sum_{i\in s}\frac{1-\pi_i}{\pi_i^2}\left[\left(y_i - \frac{1}{N}\hat{t}_{r\pi}x_i\right) + \frac{1-\pi_i}{N\pi_i}y_i\right]^2 \\ &= \frac{1}{N^2\bar{x}_{U_N}^2}\sum_{i\in s}\frac{1-\pi_i}{\pi_i^2}\left(y_i - \frac{1}{N}\hat{t}_{r\pi}x_i\right)^2 + \frac{1}{N^2\bar{x}_{U_N}^2}\sum_{i\in s}\frac{(1-\pi_i)^3}{N^2\pi_i^4}y_i^2 \\ &\quad - \frac{2}{N^2\bar{x}_{U_N}^2}\sum_{i\in s}\frac{(1-\pi_i)^2}{N\pi_i^3}\left(y_i - \frac{1}{N}\hat{t}_{r\pi}x_i\right)y_i \\ &= \widehat{AV}_{PO}\left(\hat{\theta}_{GHR}\right) + \frac{1}{N^2\bar{x}_{U_N}^2}\sum_{i\in U_N}\frac{(1-\pi_i)^3}{N^2\pi_i^4}y_i^2 I_{\{i\in s\}} \\ &\quad - \frac{2}{N^2\bar{x}_{U_N}^2}\sum_{i\in U_N}\frac{(1-\pi_i)^2}{N\pi_i^3}\left(y_i - \frac{1}{N}\hat{t}_{r\pi}x_i\right)y_i I_{\{i\in s\}} \\ &\leq \widehat{AV}_{PO}\left(\hat{\theta}_{GHR}\right) + \frac{1}{N^2\bar{x}_{U_N}^2}\sum_{i\in U_N}\frac{(1-\pi_i)^3}{N^2\pi_i^4}y_i^2\end{aligned}$$

$$
\begin{aligned}
&+ \quad \frac{2}{N^2\bar{x}_{U_N}^2}\left|\sum_{i\in U_N}\frac{(1-\pi_i)^2}{N\pi_i^3}\left(y_i-\frac{1}{N}\hat{t}_{r\pi}x_i\right)y_i\right| \\
&\leq \quad \widehat{AV}_{PO}\left(\hat{\theta}_{GHR}\right)+\frac{1}{N^3\pi_*^4\bar{x}_{U_N}^2}\left[\frac{1}{N}\sum_{i\in U_N}y_i^2\right] \\
&+ \quad \frac{2}{N^2\pi_*^3\bar{x}_{U_N}^2}\left[\frac{1}{N}\left|\sum_{i\in U_N}\left(y_i-\frac{1}{N}\hat{t}_{r\pi}x_i\right)y_i\right|\right] \\
&= \quad \widehat{AV}_{PO}\left(\hat{\theta}_{GHR}\right)+O\left(\frac{1}{N^3\pi_*^4}\right)+O\left(\frac{1}{N^2\pi_*^3}\right) \\
&= \quad \underbrace{\widehat{AV}_{PO}\left(\hat{\theta}_{GHR}\right)}_{O\left(\frac{1}{N\pi_*}\right)}+O\left(\frac{1}{N^2\pi_*^3}\right).
\end{aligned}
$$

■

3.2 CLT of Separate GHR Estimator for Stratified Sampling Design

We consider the asymptotic distribution of $\hat{t}_{ySep,GHR}$. Consider a sequence of stratified finite populations $U_H = \{1,2,\ldots,N_H\} = \cup_{h=1}^{H} U_{hH}$, $|U_{hH}| = N_{hH} \geq 4$, $N_H = \sum_{h=1}^{H} N_{hH}$, where $H \to \infty$. From the hth population, a stratified probability sample $s_H = \cup_{h=1}^{H} s_{hH}$ is selected, where in stratum h, $s_{hH} \subset U_{hH}$ is selected via a probability sampling design $p_{hH}(\cdot)$, independently of the sample selected in any other stratum $\acute{h}$. Let $\{\pi_{iH}\}$, $\{\pi_{ijH}\}$, $\{\pi_{ijkH}\}$, and $\{\pi_{ijklH}\}$ be the first, second, third, and fourth-order inclusion probabilities respectively.

Consider the following assumptions:

$S_1 : N_{hH} \leq N_* < \infty \quad \text{for all} \quad h = 1,\ldots,H \quad \text{and all} \quad H.$

$S_2 : \pi_{iH} \geq \pi_{ijH} \geq \pi_* > 0 \quad \text{for all} \quad i,j \in U_H \quad \text{and for all} \quad H.$

$S_3 : \sum_{h=1}^{H} t_{xh}^2 var_p\left(\hat{\theta}_{GHR,Sep,h}\right) \to \infty, \text{ as } H \to \infty.$

$S_4 : 0 < l_x \leq x_j \leq u_x < \infty, \text{ and } |y_j| \leq u_y < \infty \ \forall\ j \in U_H.$

Theorem 3.2.1 *Under assumptions* $S_1 - S_4$,

$$
\frac{\hat{t}_{ySep,GHR}-t_y}{\sqrt{\sum_{h=1}^{H} t_{xh}^2 var_p\left(\hat{\theta}_{GHR,Sep,h}\right)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \quad as \quad H \to \infty.
$$

Remark 3.2.1 An unbiased estimator of $var_p\left(\hat{\theta}_{GHR,Sep,h}\right)$ is discussed in theorem 2.2.3.

Proof: Since $\hat{\theta}_{GHR,Sep,h}$ is applied to each stratum separately and sampling is independent across strata then the sequence $t_{x1}\hat{\theta}_{GHR,Sep,1}$, $t_{x2}\hat{\theta}_{GHR,Sep,2}$, $\ldots, t_{xH}\hat{\theta}_{GHR,Sep,H}$ is an independent sequence. For a measurable sampling design $t_{xh}\hat{\theta}_{GHR,Sep,h}$ is an unbiased estimator for $t_{xh}\theta_h$. By triangular inequality and from equation (2.22), we have for all $h = 1, \ldots, H$ that

$$\begin{aligned}\left|t_{xh}\hat{\theta}_{GHR,Sep,h}\right| &= |t_{xh}|\left|\hat{\theta}_{GHR,Sep,h}\right| \\ &\leq \frac{t_{xh}}{\pi_{*H}}\left\{\frac{u_y}{l_x}+\frac{u_y}{l_x}+\frac{u_y}{l_x}\frac{u_x}{l_x}\right\} \\ &\leq \frac{N_* u_x}{\pi_{*H}}\left\{\frac{u_y}{l_x}+\frac{u_y}{l_x}+\frac{u_y}{l_x}\frac{u_x}{l_x}\right\}, \end{aligned} \tag{3.38}$$

Using (Ash p.308), in this uniformly bounded case Lindeberg's condition holds, and the result follows.

■

Remark 3.2.2 The condition $\sum_{h=1}^{H} t_{xh}^2 var_p\left(\hat{\theta}_{GHR,Sep,h}\right) \to \infty$ excludes the case where $var_p\left(t_{xh}\hat{\theta}_{GHR,Sep,h}\right) = 0$ for infinitely many strata.

Example 3.2.1 Under stratified simple random sampling without replacement sampling design,

$$\frac{\sum_{h=1}^{H} t_{xh}\left(\hat{\theta}_{GHR,Sep,h}-\theta_h\right)}{\sqrt{\sum_{h=1}^{H} t_{xh}^2 var_p\left(\hat{\theta}_{GHR,Sep,h}\right)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \text{ as } H \to \infty.$$

Under this sampling design, $\pi_{ijH} \geq \pi_{*H} = \min_h \frac{n_h}{N_h}\frac{n_h-1}{N_h-1} > 0$, Also, the $var_p\left(t_{xh}\hat{\theta}_{GHR,Sep,h}\right)$ can be computed from Theorem 2.1.2, and we have to exclude the case where this variance is zero for infinitely many strata.

3.3 Central Limit Theory for $\tilde{\theta}_{GHR}$

The with-replacement GHR estimator $\tilde{\theta}_{GHR}$ was introduced in Section 2.3.1 and given by equation (2.26). In this section, we will discuss the asymptotic distribution of $\tilde{\theta}_{GHR}$ under two different cases where $\tilde{\theta}_{GHR}$ is asymptotically normal with a given mean and a given variance. The two cases are that the population size N is fixed and the number of independent draws, m, tends to infinity, and when both N and m tend to infinity. From equation (2.23) recall the definition of $Z_{Ni}(\cdot)$,

$$\left.\begin{array}{ll} Z_{Ni}(y) = \frac{1}{N}\sum_{k\in U}\frac{y_k}{p_k}I_{\{\kappa_i=k\}} & Z_{Ni}(x) = \frac{1}{N}\sum_{k\in U}\frac{x_k}{p_k}I_{\{\kappa_i=k\}} \\ Z_{Ni}(r) = \frac{1}{N}\sum_{k\in U}\frac{r_k}{p_k}I_{\{\kappa_i=k\}} & Z_{Ni}(\check{y}) = \frac{1}{N}\sum_{k\in U}\frac{\check{y}_k}{p_k}I_{\{\kappa_i=k\}} \end{array}\right\} \tag{3.39}$$

where $\check{y}_k = y_k/(Np_k)$ and $r_k = y_k/x_k$. Also, recall the definition of $\tilde{\theta}_{GHR}$ from equation (2.26),

$$\begin{aligned} \tilde{\theta}_{GHR} &= \bar{Z}_{Nm}(r) + \frac{1}{\bar{x}_{U_N}}\bar{Z}_{Nm}(y) + \frac{1}{(m-1)\,\bar{x}_{U_N}}\bar{Z}_{Nm}(\check{y}) \\ &\quad - \frac{m}{(m-1)\,\bar{x}_{U_N}}\bar{Z}_{Nm}(r)\,\bar{Z}_{Nm}(x). \end{aligned} \tag{3.40}$$

Consider the following conditions,

F_1 : $0 < l_x \le x_i \le u_x < \infty$ and $|y_i| \le u_y < \infty$,

F_2 : $0 < p_* \le p_k \le p^* < 1$ and $p_k = N^{-1} + c_k N^{-(1+\tau)}$, $c_k \in [-c, c]$,

for some $c, \tau > 0$, $\forall\ k \in U$,

F_3 : $\lim_{N\to\infty}\bar{r}_{U_N} = \mu_r$, $\lim_{N\to\infty}\bar{y}_{U_N} = \mu_y$, and $\lim_{N\to\infty}\bar{x}_{U_N} = \mu_x > 0$,

F_4 : $\liminf_{N\to\infty} S^2_{wU} > 0$, where $w_k = y_k - \bar{r}_{U_N}x_k$.

It follows under F_3 and by using Fuller (1996) Corollary 5.1.1.1 p. 220 that

$$\begin{aligned} \bar{Z}_{Nm}(x) - \bar{x}_{U_N} &= O_p\left(m^{-1/2}\right), \\ \bar{Z}_{Nm}(y) - \bar{y}_{U_N} &= O_p\left(m^{-1/2}\right), \\ \bar{Z}_{Nm}(r) - \bar{r}_{U_N} &= O_p\left(m^{-1/2}\right). \end{aligned} \tag{3.41}$$

Consider the following lemmas

Lemma 3.3.1 *Assume $F_1 - F_4$. For fixed N, as $m \to \infty$,*

$$\frac{1}{(m-1)}\bar{Z}_{Nm}(\check{y}) = O_p\left(\frac{1}{m}\right).$$

Proof: Since

$$\begin{aligned} E_p\left[\frac{1}{(m-1)}\bar{Z}_{Nm}(\check{y})\right] &= \frac{1}{(m-1)}\frac{1}{mN}E_p\left[\sum_{i=1}^{m}\sum_{k\in U}\frac{y_k}{Np_k}\frac{I_{\{\kappa_i=k\}}}{p_k}\right] \\ &= \frac{1}{(m-1)}\left[\frac{1}{N}\sum_{k\in U}\frac{y_k}{Np_k}\right] \\ &= O\left(\frac{1}{m}\right) \end{aligned} \tag{3.42}$$

and

$$\begin{aligned} var_p\left[\frac{1}{(m-1)}\bar{Z}_{Nm}(\check{y})\right] &= \frac{1}{(m-1)^2}\frac{1}{m^2N^2}var_p\left[\sum_{i=1}^{m}\sum_{k\in U}\frac{y_k}{Np_k}\frac{I_{\{\kappa_i=k\}}}{p_k}\right] \\ &= \frac{1}{(m-1)^2}\frac{1}{mN^2}\sum_{k\in U}\left(\frac{y_k}{Np_k}\right)^2\frac{1-p_k}{p_k} \\ &= \frac{1}{(m-1)^2}\frac{1}{m}\left[\frac{1}{N}\sum_{k\in U}\left(\frac{y_k}{Np_k}\right)^2\frac{1-p_k}{Np_k}\right] \\ &= O\left(\frac{1}{m^3}\right), \end{aligned} \tag{3.43}$$

it follows that

$$\begin{aligned} E_p\left[\frac{1}{(m-1)}\bar{Z}_{Nm}(\check{y})\right]^2 &= var_p\left[\frac{1}{(m-1)}\bar{Z}_{Nm}(\check{y})\right] + \left[E_p\left[\frac{1}{(m-1)}\bar{Z}_{Nm}(\check{y})\right]\right]^2 \\ &= O\left(\frac{1}{m^2}\right). \end{aligned} \tag{3.44}$$

The Lemma follows by using Fuller (1996) Corollary 5.1.1.1 p. 220.

■

Lemma 3.3.2 *Assume $F_1 - F_4$. As $N \to \infty$ and $m \to \infty$,*

$$\frac{1}{(m-1)}\bar{Z}_{Nm}(\check{y}) = O_p\left(\frac{1}{m}\right).$$

Proof: From equation (3.42), we have

$$E_p\left[\frac{1}{(m-1)}\bar{Z}_{Nm}(\check{y})\right] \leq \frac{1}{(m-1)}\frac{1}{Np_*}\left[\frac{1}{N}\sum_{k\in U_N} y_k\right] = O\left(\frac{1}{m}\right), \tag{3.45}$$

and from equation (3.43), we have

$$var_p\left[\frac{1}{(m-1)}\bar{Z}_{Nm}(\check{y})\right] \leq \frac{1}{(m-1)^2}\frac{1}{m}\frac{1}{(Np_*)^3}\left[\frac{1}{N}\sum_{k\in U_N} y_k^2\right] = O\left(\frac{1}{m^3}\right).$$

Therefore,

$$\begin{aligned} E_p\left[\frac{1}{(m-1)}\bar{Z}_{Nm}(\check{y})\right]^2 &= var_p\left[\frac{1}{(m-1)}\bar{Z}_{Nm}(\check{y})\right] + \left[E_p\left[\frac{1}{(m-1)}\bar{Z}_{Nm}(\check{y})\right]\right]^2 \\ &= O\left(\frac{1}{m^2}\right). \end{aligned} \tag{3.46}$$

The Lemma follows by using Fuller (1996) Corollary 5.1.1.1 p. 220.

■

Under conditions $F_1 - F_4$, rewrite $\tilde{\theta}_{GHR}$ as

$$\tilde{\theta}_{GHR} = O_p\left(\frac{1}{m}\right) + \bar{Z}_{Nm}(r) + \frac{1}{\bar{x}_{U_N}}\bar{Z}_{Nm}(y) - \frac{m-1+1}{(m-1)\bar{x}_{U_N}}\bar{Z}_{Nm}(r)\bar{Z}_{Nm}(x)$$

Therefore,

$$\begin{aligned} \tilde{\theta}_{GHR} - \theta &= O_p\left(\frac{1}{m}\right) - \frac{m}{(m-1)\bar{x}_{U_N}}\left(\bar{Z}_{Nm}(r) - \bar{r}_{U_N}\right)\left(\bar{Z}_{Nm}(x) - \bar{x}_{U_N}\right) \\ &\quad - \frac{1}{m-1}\bar{Z}_{Nm}(r) - \frac{m}{(m-1)\bar{x}_{U_N}}\bar{r}_{U_N}\bar{Z}_{Nm}(x) + \frac{m}{m-1}\bar{r}_{U_N} \\ &\quad + \frac{1}{\bar{x}_{U_N}}\bar{Z}_{Nm}(y) - \frac{\bar{y}_{U_N}}{\bar{x}_{U_N}} \\ &= O_p\left(\frac{1}{m}\right) - \frac{m}{(m-1)\bar{x}_{U_N}}\left(\bar{Z}_{Nm}(r) - \bar{r}_{U_N}\right)\left(\bar{Z}_{Nm}(x) - \bar{x}_{U_N}\right) \\ &\quad - \frac{1}{m-1}\bar{Z}_{Nm}(r) - \frac{m}{(m-1)\bar{x}_{U_N}}\bar{r}_{U_N}\left(\bar{Z}_{Nm}(x) - \bar{x}_{U_N}\right) \\ &\quad + \frac{1}{\bar{x}_{U_N}}\left(\bar{Z}_{Nm}(y) - \bar{y}_{U_N}\right) \\ &= O_p\left(\frac{1}{m}\right) + O_p\left(\frac{1}{m}\right) + O_p\left(\frac{1}{m^{3/2}}\right) + \frac{1}{\bar{x}_{U_N}}\left(\bar{Z}_{Nm}(y) - \bar{y}_{U_N}\right) \\ &\quad - \frac{m}{(m-1)\bar{x}_{U_N}}\bar{r}_{U_N}\left(\bar{Z}_{Nm}(x) - \bar{x}_{U_N}\right) \\ &= O_p\left(\frac{1}{m}\right) + \frac{1}{\bar{x}_{U_N}}\left(\bar{Z}_{Nm}(y) - \bar{y}_{U_N}\right) - \frac{1}{\bar{x}_{U_N}}\bar{r}_{U_N}\left(\bar{Z}_{Nm}(x) - \bar{x}_{U_N}\right) \end{aligned}$$

Hence,

$$\begin{aligned}\tilde{\theta}_{GHR}-\theta &= O_p\left(\frac{1}{m}\right)+\frac{1}{mN\bar{x}_{U_N}}\sum_{i=1}^{m}\sum_{k\in U}(y_k-\bar{r}_{U_N}x_k)\left[\frac{I_{\{\kappa_i=k\}}}{p_k}-1\right]\\ &= O_p\left(\frac{1}{m}\right)+\frac{1}{mN\bar{x}_{U_N}}\sum_{i=1}^{m}\sum_{k\in U}w_k\left[\frac{I_{\{\kappa_i=k\}}}{p_k}-1\right], \qquad (3.47)\end{aligned}$$

where $w_k=(y_k-\bar{r}_{U_N}x_k)$. Note that

$$\frac{1}{N}\sum_{k\in U}w_k\left[\frac{I_{\{\kappa_i=k\}}}{p_k}-1\right] \qquad (3.48)$$

are iid with mean 0 and variance

$$V_w=N^{-2}V_1 \qquad (3.49)$$

where

$$V_1=\sum_{k\in U_N}\left(\frac{w_k}{p_k}-t_w\right)^2 p_k=\sum_{k\in U_N}\frac{w_k^2}{p_k}-t_w^2. \qquad (3.50)$$

It follows that

$$\begin{aligned}\hat{V}_w &= \frac{1}{N^2}\hat{V}_1=\frac{1}{N^2}\frac{1}{m-1}\sum_{i\in s}\left[\sum_{k\in U_N}\frac{w_k}{p_k}I_{\{\kappa_i=k\}}-\hat{t}_w\right]^2\\ &= \frac{1}{m-1}\sum_{i\in s}\left[\sum_{k\in U_N}\frac{w_k}{Np_k}I_{\{\kappa_i=k\}}-\frac{1}{N}\hat{t}_w\right]^2 \qquad (3.51)\end{aligned}$$

is an unbiased estimator for V_w, where

$$\hat{t}_w=\frac{1}{m}\sum_{i\in s}\sum_{k\in U_N}\frac{w_k}{p_k}I_{\{\kappa_i=k\}}. \qquad (3.52)$$

Remark 3.3.1 Rewrite $\hat{V}_w$ as

$$\hat{V}_w=\frac{1}{m-1}\sum_{i\in s}\left(\psi_i-\bar{\psi}\right)^2=S_{\psi s}^2,$$

where $\psi_i=N^{-1}\sum_{k\in U_N}p_k^{-1}w_kI_{\{\kappa_i=k\}}$ and $\bar{\psi}=m^{-1}\sum_{i\in s}\psi_i$.

The rth mean, ν_r, of $N^{-1}\sum_{k\in U_N} p_k^{-1} w_k I_{\{\kappa_i=k\}}$ is given by

$$\nu_r = \frac{1}{N}\sum_{k\in U}\frac{w_k^r}{N^{r-1}p_k^{r-1}} \tag{3.53}$$

$$= O\left(\frac{1}{N^{r-1}p_*^{r-1}}\right). \tag{3.54}$$

Define $\mu_4 = E_p\left(\psi_1 - N^{-1}t_w\right)^4$; further,

$$\mu_4 = \nu_4 - 4\nu_3\frac{t_w}{N} + 6\nu_2\left(\frac{t_w}{N}\right)^2 - 4\nu_1\left(\frac{t_w}{N}\right)^3 + \left(\frac{t_w}{N}\right)^4. \tag{3.55}$$

For finite fourth moment iid random variables (Mood et al. (1974) Theorem 3.3 P. 11),

$$var_p\left(S_{\psi s}^2\right) = \frac{1}{m}\left(\mu_4 - \frac{m-3}{m-1}V_w^2\right). \tag{3.56}$$

Lemma 3.3.3 *For fixed population size N, and as $m \to \infty$ the estimator*

$$\widehat{AV}_{WR}\left(\tilde{\theta}_{GHR}\right) = \frac{1}{m\bar{x}_{U_N}^2}\hat{V}_w$$

is consistent for

$$AV_{WR}\left(\tilde{\theta}_{GHR}\right) = \frac{1}{m\bar{x}_{U_N}^2}V_w,$$

where V_w and $\hat{V}_w$ are given in equations (3.49) and (3.51) respectively.

Proof: From equation 3.55, μ_4 is finite when the population size N is fixed. Therefore,

$$E_p\left|\frac{\widehat{AV}_{WR}\left(\tilde{\theta}_{GHR}\right)}{AV_{WR}\left(\tilde{\theta}_{GHR}\right)} - 1\right|^2 = \frac{E_p\left(\hat{V}_w - V_w\right)^2}{V_w^2}$$

$$= \frac{var_p\left(\hat{V}_w\right)}{V_w^2}$$

$$= \frac{1}{V_w^2}O\left(\frac{1}{m}\right)$$

$$\to 0 \quad \text{as} \quad m \to \infty.$$

■

The following lemma covers the case when N, and $m \to \infty$.

Lemma 3.3.4 *Under conditions $F_1 - F_4$ and as m, $N \to \infty$ the estimator*

$$\widehat{AV}_{WR}\left(\tilde{\theta}_{GHR}\right) = \frac{1}{m^2 \bar{x}_{U_N}^2} \hat{V}_w$$

is consistent for

$$AV_{WR}\left(\tilde{\theta}_{GHR}\right) = \frac{1}{m\bar{x}_{U_N}^2} V_w.$$

Proof: From equation (3.55), $\mu_4 = O(1)$. Hence from equation (3.56), $var_p\left(\hat{V}_w\right) = O\left(\frac{1}{m}\right)$.

From equation (3.50), recall the definition of V_1 and under the assumption F_2 we have

$$\begin{aligned} V_1 &= \sum_U \frac{w_k^2}{p_k} - t_w^2 \\ &= N \sum_U \frac{w_k^2}{1 + c_k/N^\tau} - N\frac{t_w^2}{N} \\ &= N\left(\sum_U w_k^2 - \frac{t_w^2}{N}\right) + N\sum_U w_k^2 O\left(\frac{1}{N^\tau}\right) \\ &= N(N-1)S_{wU}^2 + \frac{1}{N}\sum_U w_k^2 O\left(N^{2-\tau}\right) \end{aligned} \tag{3.57}$$

where the $O(\cdot)$ terms are uniform in $k \in U$. Hence

$$\frac{1}{N^2}\left(\sum_U \frac{w_k^2}{p_k} - t_w^2\right) \sim S_{wU}^2 + O\left(N^{-\tau}\right) \sim S_{wU}^2. \tag{3.58}$$

Therefore,

$$\begin{aligned} E_p\left|\frac{\widehat{AV}_{WR}\left(\tilde{\theta}_{GHR}\right)}{AV_{WR}\left(\tilde{\theta}_{GHR}\right)} - 1\right|^2 &= \frac{E_p\left(\hat{V}_w - V_w\right)^2}{\left[S_{wU}^2 + O\left(N^{-\tau}\right)\right]^2} \\ &= \frac{var_p\left(\hat{V}_w\right)}{\left[S_{wU}^2 + O\left(N^{-\tau}\right)\right]^2} \\ &= \frac{1}{\left[S_{wU}^2 + O\left(N^{-\tau}\right)\right]^2} O\left(\frac{1}{m}\right) \\ &\to 0 \quad \text{as} \quad m \to \infty. \end{aligned}$$

■

In the following theorems we will discuss the asymptotic distribution of $\tilde{\theta}_{GHR}$ when N is fixed and $m \to \infty$ and when m, $N \to \infty$.

Theorem 3.3.1 *For fixed population size N,*

$$\frac{\tilde{\theta}_{GHR} - \theta}{\sqrt{\widehat{AV}_{WR}\left(\tilde{\theta}_{GHR}\right)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1) \quad as \quad m \to \infty.$$

Proof: From equation (3.47), we have

$$\tilde{\theta}_{GHR} - \theta = O_p\left(\frac{1}{m}\right) + \frac{1}{m\bar{x}_{U_N}} \sum_{i=1}^{m} \sum_{k \in U} \frac{w_k}{N} \left[\frac{I_{\{\kappa_i = k\}}}{p_k} - 1\right] \tag{3.59}$$

Define

$$\begin{aligned} AV_{WR}\left(\tilde{\theta}_{GHR}\right) &= var_p \left[\frac{1}{m\bar{x}_{U_N}} \sum_{i=1}^{m} \underbrace{\sum_{k \in U} \frac{w_k}{N} \left[\frac{I_{\{\kappa_i = k\}}}{p_k} - 1\right]}_{\text{are iid } (0,\, V_w)} \right] \\ &= \frac{1}{m\bar{x}_{U_N}^2} V_w \end{aligned} \tag{3.60}$$

Therefore, by a standard CLT (Casella and Berger (2002) p.236)

$$\frac{\tilde{\theta}_{GHR} - \theta}{\sqrt{AV_{WR}\left(\tilde{\theta}_{GHR}\right)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1) \quad \text{as} \quad m \to \infty. \tag{3.61}$$

Since

$$\frac{\tilde{\theta}_{GHR} - \theta}{\sqrt{AV_{WR}\left(\tilde{\theta}_{GHR}\right)}} = \frac{\sqrt{\widehat{AV}_{WR}\left(\tilde{\theta}_{GHR}\right)}}{\sqrt{AV_{WR}\left(\tilde{\theta}_{GHR}\right)}} \frac{\tilde{\theta}_{GHR} - \theta}{\sqrt{\widehat{AV}_{WR}\left(\tilde{\theta}_{GHR}\right)}},$$

then the theorem follows from Lemma 3.3.3.

■

Theorem 3.3.2 *Under conditions* $F_1 - F_4$

$$\frac{\tilde{\theta}_{GHR} - \theta}{\sqrt{\widehat{AV}_{WR}\left(\tilde{\theta}_{GHR}\right)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,\, 1) \quad as \quad m,\, N \to \infty.$$

Proof: From equation (3.47), recall that

$$\tilde{\theta}_{GHR} - \theta = O_p\left(\frac{1}{m}\right) + \frac{1}{m\bar{x}_{U_N}} \sum_{i=1}^{m} \sum_{k \in U} \frac{w_k}{N} \left[\frac{I_{\{\kappa_i = k\}}}{p_k} - 1\right], \tag{3.62}$$

It follows from equation (3.55) that

$$\begin{aligned} E_p \left\{ \frac{1}{Nm} \sum_{k \in U} w_k \left[\frac{I_{\{\kappa_i = k\}}}{p_k} - 1\right] \right\}^4 &= \frac{\mu_4}{m^4} \\ &= O\left(\frac{1}{m^4}\right), \end{aligned} \tag{3.63}$$

and hence

$$\sum_{i=1}^{m} E_p \left\{ \frac{1}{Nm} \sum_{k \in U} w_k \left[\frac{I_{\{\kappa_i = k\}}}{p_k} - 1\right] \right\}^4 = O\left(\frac{1}{m^3}\right).$$

Define

$$\begin{aligned} C_N^2 &= var_p \left(\frac{1}{m} \sum_{i=1}^{m} \sum_{k \in U} \frac{w_k}{N} \left[\frac{I_{\{\kappa_i = k\}}}{p_k} - 1\right] \right) \\ &= \frac{1}{m} \frac{1}{N^2} \left(\sum_{U} \frac{w_k^2}{p_k} - t_w^2 \right) \qquad \text{using equation (3.58), we have} \\ &= \frac{1}{m} \left[S_{wU}^2 + O\left(N^{-\tau}\right) \right]. \end{aligned}$$

Note that

$$\begin{aligned} \frac{\sum_{i=1}^{m} E_p \left\{ \frac{1}{Nm} \sum_{k \in U} w_k \left[\frac{I_{\{\kappa_i = k\}}}{p_k} - 1\right] \right\}^4}{C_N^4} &= \frac{O\left(\frac{1}{m^3}\right)}{\frac{1}{m^2}\left[S_{wU}^2 + O\left(N^{-\tau}\right)\right]^2} \\ &= O\left(\frac{1}{m}\right) \frac{1}{\left[S_{wU}^2 + O\left(N^{-\tau}\right)\right]^2} \\ &\to 0 \quad \text{as} \quad N,\, m \to \infty, \end{aligned} \tag{3.64}$$

the Lyapunov's condition holds for $\delta = 2$. It follows that

$$C_N^{-1} \sum_{i=1}^{m} \sum_{k \in U} w_k \left[\frac{I_{\{\kappa_i = k\}}}{p_k} - 1\right] \xrightarrow{\mathcal{L}} \mathcal{N}(0,\, 1) \tag{3.65}$$

as $m,\ N \to \infty$. The theorem follows from Lemma 3.3.4, and Slutsky's theorem.

■

Result 3.3.1 *Under the model* $\xi:\ y_k = \beta x_k + \epsilon_k$ *where* ϵ_k *are independent* $(0,\ \sigma_k^2)$. *Then* $E_\xi\left[AV_{WR}\left(\tilde{\theta}_{GHR}\right)\right]$ *is not attained the Godambe and Joshi (1965) lower bound as* $m \to \infty$ *and attained the Godambe and Joshi (1965) lower bound as* $m,\ N \to \infty$.

Proof: Under the model ξ, we have

$$\begin{aligned} E_\xi(w_k) &= 0, \\ E_\xi(w_k)^2 &= var_\xi(w_k) \\ &= \sigma_k^2 - \frac{2}{N}\sigma_k^2 + \frac{1}{N^2}x_k^2 \sum_{l \in U_N} \frac{\sigma_l^2}{x_l^2}. \end{aligned}$$

$$\begin{aligned} t_w &= t_y - t_x \bar{r}_{U_N}, \\ E_\xi(t_w) &= 0, \\ E_\xi(t_w)^2 &= var_\xi(t_w) \\ &= \sum_{k \in U_N} \sigma_k^2 + \bar{x}_{U_N}^2 \sum_{k \in U_N} \frac{\sigma_k^2}{x_k^2} - 2\bar{x}_{U_N} \sum_{k \in U_N} \frac{\sigma_k^2}{x_k}. \end{aligned}$$

Therefore,

$$\begin{aligned} E_\xi(V_1) &= E_\xi\left(\sum_{k \in U_N} \frac{w_k^2}{p_k} - t_w^2\right) \\ &= \sum_{k \in U_N} \frac{1-p_k}{p_k}\sigma_k^2 - \frac{2}{N}\sum_{k \in U_N} \frac{\sigma_k^2}{p_k} + 2\bar{x}_{U_N} \sum_{k \in U_N} \frac{\sigma_k^2}{x_k} \\ &+ \left[\frac{1}{N^2}\sum_{k \in U_N} \frac{x_k^2}{p_k}\right]\left[\sum_{l \in U_N} \frac{\sigma_l^2}{x_l^2}\right] - \bar{x}_{U_N}^2\left[\sum_{l \in U_N} \frac{\sigma_l^2}{x_l^2}\right]. \end{aligned}$$

Hence,

$$E_\xi\left[AV_{WR}\left(\tilde{\theta}_{GHR}\right)\right] = E_\xi\left[\frac{1}{mN^2\bar{x}_{U_N}^2} V_1\right]$$

$$
\begin{aligned}
&= \frac{1}{mN^2\bar{x}_{U_N}^2}\sum_{k\in U_N}\frac{1-p_k}{p_k}\sigma_k^2 - \frac{2}{mN\bar{x}_{U_N}^2}\left[\frac{1}{N^2}\sum_{k\in U_N}\frac{\sigma_k^2}{p_k}\right] \\
&+ \frac{2}{mN\bar{x}_{U_N}}\left[\frac{1}{N}\sum_{k\in U_N}\frac{\sigma_k^2}{x_k}\right] \\
&+ \frac{1}{mN\bar{x}_{U_N}^2}\left[\frac{1}{N^2}\sum_{k\in U_N}\frac{x_k^2}{p_k}\right]\left[\frac{1}{N}\sum_{l\in U_N}\frac{\sigma_l^2}{x_l^2}\right] \\
&- \frac{1}{mN}\left[\frac{1}{N}\sum_{l\in U_N}\frac{\sigma_l^2}{x_l^2}\right] \\
&= \underbrace{\frac{1}{mN^2\bar{x}_{U_N}^2}\sum_{k\in U_N}\frac{1-p_k}{p_k}\sigma_k^2}_{GJLB=O\left(\frac{1}{m}\right)} + O\left(\frac{1}{mN}\right);
\end{aligned}
$$

therefore,

$$
\frac{E_\xi\left[AV_{WR}\left(\tilde{\theta}_{GHR}\right)\right]}{GJLB} = 1 + O_p\left(\frac{1}{N}\right).
$$

If the population size N is fixed, then the Godambe and Joshi (1965) lower bound is not attainable. However, if the population size $N \to \infty$, then the Godambe and Joshi (1965) lower bound is attainable.

■

3.4 Simulation Results

In this section we will compare various ratio estimators through simulations. The entire population consists of $(x_1, y_1), \ldots, (x_{1000}, y_{1000})$. Consider three different sampling designs, simple random sampling without replacement (SI), Poisson sampling (PO), and probability proportional to size without replacement (πps). This section will be divided into two subsections, covering the unstratified case and stratified case.

3.4.1 Simulation Results For Unstratified Sampling

Let x_i be independent and identically distributed (iid) as Gamma($\alpha = 3,\ \beta = 2$), with mean $\alpha\beta = 6$ and variance $\alpha\beta^2 = 12$.

Let $y_i = 3x_i + \epsilon_i$ with $\{\epsilon_i\}$ independent $N(0,\ 25x_i)$. (Under this particular models the bias of $\hat{\theta}$ is quite small, as mentioned in Remark 1.3.1.) Let z_i denote the size variable when πps sampling is used, and take

$$z_i = 10 + x_i + \eta_i, \qquad \{\eta_i\} \ iid\ N(0,\ 1)$$

independent of $\{\epsilon_i\}$. Under πps sampling, first and second order inclusion probabilities are obtained from the `out= JTPROBS` option in SAS `proc surveyselect`.

For PO sampling design, the first order inclusion probabilities are simulated by

$$\pi_i = b\frac{\exp(1.5x_i)}{1 + \exp(1.5x_i)}$$

and b is chosen such that $Nb\bar{E} = 10, 15, \ldots, 40, 75$ where $\bar{E} = N^{-1}\sum_U e_i$ and $e_i = \exp(1.5x_i)\left[1 + \exp(1.5x_i)\right]^{-1}$.

Furthermore, by independence the higher order inclusion probabilities are defined for PO sampling as follows:

$$\begin{array}{ll} \pi_{ij} = \pi_i\pi_j, & \text{for } ij \in s_{D_2}; \\ \pi_{ijk} = \pi_i\pi_j\pi_k, & \text{for } ijk \in s_{D_3}; \\ \pi_{ijkl} = \pi_i\pi_j\pi_k\pi_l, & \text{for } ijkl \in s_{D_4}. \end{array} \tag{3.66}$$

Define the mean square error (MSE) ratio $R = \frac{MSE_{(\hat{\theta}_{GHR})}}{MSE_{\hat{\theta}}}$, where the MSEs are empirical values based on 1500 simulated samples realization of the finite population. Values of R less than one favor $\hat{\theta}_{GHR}$. If the numerator and denominator of R were independent variance estimators, then R would have an approximate F distribution with 1499 numerator degrees of freedom and 1499 denominator degrees of freedom. The corresponding quantiles are $F(1499, 1499, 0.025) = 0.903676$ and $F(1499, 1499, 0.025) = 1.1065913$. The MSE's of $\hat{\theta}$ and $\hat{\theta}_{GHR}$ are dependent, however, and we would expect that the MSE ratio should have distribution more tightly

concentrated around one than the F-distribution when the two estimators are behaving equivalently. That is, if $R < 0.9$ or $R > 1.1$, then we can conclude that this is a significant difference in estimator performance, and is not due to chance in the 1500 replications of the simulation experiment.

Results for simple random sampling without replacement at various sample size are shown in Table 3.1. At all samples sizes, $\hat{\theta}_{GHR}$ and $\hat{\theta}$ are virtually indistinguishable in this case.

	sample size							
	10	15	20	25	30	35	40	75
$R = \frac{MSE(\hat{\theta})}{MSE(\hat{\theta}_{GHR})}$	0.986	0.992	0.991	1.000	1.000	0.996	0.993	1.003
$MSE\left(\hat{\theta}\right)$	0.450	0.287	0.216	0.178	0.137	0.121	0.108	0.047

Table 3.1: Empirical MSE ratios, each based on 1500 simulated SI samples. At various sample sizes from a fixed finite population.

Results for πps sampling without replacement, as implemented in SAS `proc surveyselect` option `METHOD=PPS`, are summarized in Table 3.2. In this case, the two estimators behave equivalently at small sample size ($n < 30$). For large n, $\hat{\theta}$ is essentially unbiased and has lower variance than $\hat{\theta}_{GHR}$, so the MSE ratios are significantly greater than one. This is consistent with the expectation that main advantage of $\hat{\theta}_{GHR}$ is its exact unbiasedness.

The sample size under PO sampling is random. Table 3.3 displays MSE ratios at various expected sample size, $E_p(n_s)$, ranging from 10 to 75. In contrast with Table 3.2, MSE ratios are high at small expected sample size favoring $\hat{\theta}$ over $\hat{\theta}_{GHR}$. This phenomenon is similar to that seen in comparing the weighed sample

	sample size							
	10	15	20	25	30	35	40	75
$R = \frac{MSE(\hat{\theta})}{MSE(\hat{\theta}_{GHR})}$	0.994	0.979	0.960	0.937	0.886	0.894	0.863	0.807
$MSE\left(\hat{\theta}\right)$	0.426	0.265	0.202	0.165	0.129	0.126	0.101	0.053

Table 3.2: Empirical MSE ratios, each based on 1500 simulated πps samples. At various sample sizes from a fixed finite population.

mean or Hajek estimator $\left(\sum_s y_k \pi_k^{-1} / \sum_s \pi_k^{-1}\right)$ to the Horvitz-Thompson estimator $\left(\sum_s y_k \pi_k^{-1} / N\right)$ in case of random sample size (see Särndal et al. (1992), p. 87). At higher expected sample sizes, the variation in sample size is less critical, and $\hat{\theta}$ and $\hat{\theta}_{GHR}$ are essentially equally efficient.

	$E_p(n_s)$							
	10	15	20	25	30	35	40	75
$R = \frac{MSE(\hat{\theta})}{MSE(\hat{\theta}_{GHR})}$	0.840	0.852	0.887	0.939	0.957	0.963	0.973	0.973
$MSE\left(\hat{\theta}\right)$	0.224	0.147	0.111	0.089	0.075	0.064	0.055	0.026

Table 3.3: Empirical MSE ratios, each based on 1500 simulated PO samples. at various sample sizes from a fixed finite population.

The overall conclusions from Tables 3.1 to 3.3 are that $\hat{\theta}$ and $\hat{\theta}_{GHR}$ perform similarly, with $\hat{\theta}_{GHR}$ having some advantages at small, fixed sample sizes due to its unbiasedness, and $\hat{\theta}$ having some advantages at small, random sample sizes due to its low variance.

We now consider confidence interval properties for the two estimators via simulation. The approximation methods described in Section 2.4 for estimating the variance of $\hat{\theta}_{GHR}$ will be used in the following simulation results. For each table, we compute the percent relative bias,

$$\%RelBias = \frac{(\text{simulation mean of estimator}) - \theta}{\theta} \cdot 100\%,$$

the percentages of nominal 95% confidence intervals that cover θ, and the average length of the confidence interval.

Table 3.4 gives results for SI. Both $\hat{\theta}$ and $\hat{\theta}_{GHR}$ have negligible empirical bias (less than 0.5% in absolute value) in all cases. The confidence intervals for $\hat{\theta}_{GHR}$ have slightly better coverage than those for $\hat{\theta}$, particularly at small sample sizes. This better coverage comes at the expense of slightly wider confidence intervals.

	Sample Size							
	10	15	20	25	30	35	40	75
%Rel Bias $(\hat{\theta})$	-0.482	-0.151	0.111	0.197	-0.442	0.042	-0.125	0.172
%Coverage $(\hat{\theta})$	92.3	92.7	93.1	93.3	94.9	94.0	94.0	96.3
Length of CI $(\hat{\theta})$	2.551	2.070	1.809	1.614	1.471	1.348	1.270	0.912
%Rel Bias $(\hat{\theta}_{GHR})$	-0.308	-0.060	0.289	0.275	-0.342	0.093	-0.058	0.189
%Coverage$_{app}$ $(\hat{\theta}_{GHR})$	93.3	94.2	93.9	94.5	95.5	94.3	94.5	96.5
%Coverage$_{pwr}$ $(\hat{\theta}_{GHR})$	93.5	94.3	94.3	94.8	95.8	94.8	95.0	96.9
Length of CI$_{app}$ $(\hat{\theta}_{GHR})$	2.849	2.240	1.915	1.687	1.531	1.398	1.306	0.925
Length of CI$_{pwr}$ $(\hat{\theta}_{GHR})$	2.864	2.258	1.934	1.709	1.555	1.423	1.333	0.961

Table 3.4: Percentage relative bias, percentage coverage of nominal 95% CIs. Average length of confidence intervals under 1500 simulated replicates of simple random sampling without replacement.

Table 3.5 summarizes the results for the πps sampling design. Results are similar to those under SI, with negligible empirical bias in all cases, slightly better coverage of confidence intervals associated with $\hat{\theta}_{GHR}$, at the expense of slightly wider confidence intervals.

Finally, results for Poisson sampling are summarized in Table 3.6. Once again, empirical bias is negligible in all cases, so the exact unbiasedness of $\hat{\theta}_{GHR}$ is bringing no real advantage. Further, the extra variability due to random sample sizes under PO negatively impacts $\hat{\theta}_{GHR}$, so that confidence interval coverage is no better than that of $\hat{\theta}$, at the expense of wider confidence intervals.

In these unstratified simulations, $\hat{\theta}_{GHR}$ performed overall very similar to $\hat{\theta}$, with slight advantages to one estimator or the other in certain cases. In the following subsections, we simulate the performance of $\hat{t}_{ySep,GHR}$ and $\hat{t}_{ySep,\hat{\theta}}$ under stratified

	Sample Size							
	10	15	20	25	30	35	40	75
%Rel Bias $\left(\hat{\theta}\right)$	-0.143	-0.228	-0.472	-0.882	0.276	-0.359	0.185	0.047
%Coverage $\left(\hat{\theta}\right)$	91.7	93.1	93.9	94.1	95.1	93.3	94.9	94.6
Length of CI $\left(\hat{\theta}\right)$	2.464	2.025	1.758	1.583	1.436	1.336	1.244	0.896
%Rel Bias $\left(\hat{\theta}_{GHR}\right)$	-0.159	-0.203	-0.439	-0.800	0.375	-0.279	0.232	0.107
%Coverage$_{app}$ $\left(\hat{\theta}_{GHR}\right)$	93.7	94.5	94.5	94.7	96.0	93.7	95.2	95.1
%Coverage$_{pwr}$ $\left(\hat{\theta}_{GHR}\right)$	93.5	94.6	94.5	94.9	96.1	94.3	95.1	93.9
Length of CI$_{app}$ $\left(\hat{\theta}_{GHR}\right)$	2.747	2.188	1.862	1.660	1.503	1.393	1.295	0.930
Length of CI$_{pwr}$ $\left(\hat{\theta}_{GHR}\right)$	2.758	2.198	1.877	1.675	1.511	1.400	1.304	0.939

Table 3.5: Percentage relative bias, percentage coverage of nominal 95% CIs. Average length of confidence intervals under 1500 simulated replicates of πps sampling without replacement.

sampling, to estimate the population total. In this setting, it is expected that bias will accumulate in $\hat{t}_{ySep,\hat{\theta}}$, while $\hat{t}_{ySep,GHR}$ is exactly unbiased, so GHR methodology may show some real advantages.

3.4.2 Simulation Results For Stratified Sampling

When the population is divided into different strata, $U = \cup_{h=1}^{H} U_h$, and each stratum is relatively homogenous, one can expect that stratification will improve the efficiency of the parameter estimate. In this subsection, we will concentrate on estimating the population total, t_y. Let $N_h = \sum_{k \in U_h} 1$, $t_{yh} = \sum_{k \in U_h} y_k$, $t_{xh} = \sum_{k \in U_h} x_k$ for $h = 1, \ldots, H$. The importance of stratification appears when the strata totals, t_{yh}, have big differences from stratum to stratum. To take advantage of the efficiency of stratification, we use $\hat{\theta}$ and $\hat{\theta}_{GHR}$ as components of separate ratio estimators, as discussed in Sections 1.4 and 2.2. Using either $\hat{\theta}$ or $\hat{\theta}_{GHR}$ as combined ratio estimators will ignore the efficiency afforded by stratification. Combined ratio estimators well only if the stratum ratios, θ_h, are not varying from stratum to stratum.

	En_S							
	10	15	20	25	30	35	40	75
%Rel Bias $\left(\hat{\theta}\right)$	-0.069	-0.159	-0.202	-0.155	-0.236	-0.225	-0.203	-0.073
%Coverage $\left(\hat{\theta}\right)$	92.1	93.3	93.8	93.7	93.9	93.7	93.1	94.3
Length of CI $\left(\hat{\theta}\right)$	1.773	1.459	1.259	1.125	1.021	0.942	0.876	0.620
%Rel Bias $\left(\hat{\theta}_{GHR}\right)$	-0.034	-0.026	-0.152	-0.125	-0.211	-0.201	-0.159	-0.060
%Coverage$_{app}$ $\left(\hat{\theta}_{GHR}\right)$	92.4	93.1	93.7	94.0	93.9	93.9	93.9	93.9
%Coverage$_{pwr}$ $\left(\hat{\theta}_{GHR}\right)$	91.6	93.3	93.6	94.1	94.5	94.5	94.7	95.4
Length of CI$_{app}$ $\left(\hat{\theta}_{GHR}\right)$	1.984	1.573	1.336	1.177	1.064	0.975	0.905	0.628
Length of CI$_{pwr}$ $\left(\hat{\theta}_{GHR}\right)$	1.948	1.566	1.342	1.192	1.086	1.001	0.935	0.679

Table 3.6: Percentage relative bias, percentage coverage of nominal 95% CIs. Average length of confidence intervals under 1500 simulated replicates of PO sampling without replacement.

As noted previously, the bias issue will arise when we are using $\hat{\theta}$ as a separate ratio estimator under stratified sampling, especially if the stratum sizes N_h are small. An exactly unbiased estimator $\hat{\theta}_{GHR}$ is crucial to overcome this problem.

From Section 1.4, recall of the definitions the separate ratio estimator and the combined ratio estimator:

$$\hat{t}_{ySep,\hat{\theta}} = \sum_{h=1}^{H} t_{xh} \frac{\sum_{i \in s_h} (y_i \,/\, \pi_i)}{\sum_{i \in s_h} (x_i \,/\, \pi_i)}, \tag{3.67}$$

$$\hat{t}_{yComb,\hat{\theta}} = t_x \frac{\sum_{h=1}^{H} \hat{t}_{yh}}{\sum_{h=1}^{H} \hat{t}_{xh}}, \tag{3.68}$$

where $\hat{t}_{yh}$ and $\hat{t}_{xh}$ are the HT estimator for t_y and t_x respectively for the h^{th} stratum.

Also, from Section 2.2 recall the definition of the GHR version of the separate ratio estimator: Estimate the population total, t_y, by

$$\hat{t}_{ySep,GHR} = \sum_{h=1}^{H} t_{xh} \hat{\theta}_{GHR,h}$$

where

$$\hat{\theta}_{GHR,h} = \frac{1}{N_h}\sum_{i\in s_h} r_i \frac{1}{\pi_i} + \frac{1}{N_h \bar{x}_{U_h}}\left[\sum_{i\in s_h}\frac{y_i}{\pi_i} - \frac{1}{N_h}\sum_{i\in s_h}\sum_{j\in s_h} r_i x_j \frac{1}{\pi_{ij}}\right]. \tag{3.69}$$

We simulate a stratified finite population as follows. For a given value of H (10, 20, 30, or 40) and a particular stratum $h \in \{1, 2, \ldots, H\}$, define $b_h = \pi/18 + 4\pi(h-1)/[9(H-1)]$. For $i \in U_h$, we simulate x_i independent $Uniform(0,\, b_h)$ and $y_i = h^{-1}\sin(x_i) + \epsilon_i$, ϵ_i independent $N\left(0,\, (0.001)^2 x_i\right)$. When πps is used, we use $z_i = 2 - \sqrt{x_i}$ as the size variable. Simulation results are based on $N_h = 50$ in every stratum, and 1500 iterations.

As noted earlier $\hat{\theta}$ has approximately bias zero if we are simulating from straight line passing through the origin. We will deviate little bit from this line by choosing $\sin(\cdot)$ which pass through the origin and for any choice of b_h we have $x_i \in (\pi/18,\ \pi/2)$, which means, in each stratum the values of $\sin(\cdot)$ can be approximated by a line. In this way we can produce bias for $\hat{\theta}$ and we can make it worst if N_h reduced to smaller numbers.

For this stratified population, the within-stratum ratios vary from stratum to stratum, so that the estimator $\hat{t}_{yComb,\hat{\theta}}$ is based on the wrong model. Therefore, we expect $\hat{t}_{yComb,\hat{\theta}}$ to have larger variance and larger confidence interval length than separate ratio estimators. Further, $\hat{t}_{ySep,\hat{\theta}}$ is based on the correct model and so is expected to have small variance, but potentially large bias due to the accumulation of bias from stratum to stratum.

The GHR-based separate ratio estimator is expected to have advantages in this context. Unlike $\hat{t}_{yComb,\hat{\theta}}$, $\hat{t}_{ySep,GHR}$ is based on the correct model, and should have relatively small variance. Unlike $\hat{t}_{ySep,\hat{\theta}}$, $\hat{t}_{ySep,GHR}$ is exactly unbiased, and so will not suffer from any accumulation of bias.

In our simulation study, we consider sample sizes of $n_h = 4$, 5, or 6 per stratum. (A minimum of $n_h = 4$ is needed to ensure fourth-order measurability, so

that an unbiased variance estimator can be computed.) Under Poisson sampling, the sample size is random and is zero in some stratum h with high probability when $E[n_s] = 4,\ 5,$ or 6. We therefore restrict attention to stratified simple random sampling without replacement (STSI) and stratified πps sampling (STπps) as implemented in SAS `proc surveyselect`.

For STSI, the exactly unbiased variance estimator is easily computed since fourth-order inclusion probabilities are readily available. For STπps, however, fourth-order inclusion probabilities are not available, and we use the first method of approximation as described in Section 2.4.

In stratified sampling design, the estimator $\hat{t}_{ycom}$ is wrong model of estimating the population total, t_y. Therefore, we expected $\hat{t}_{ycom}$ has a high variance and larger confidence interval length. In the same time, $\hat{t}_{ysep}$ is correct model for stratified sampling design, but has less variance and this is due to the accumulate bias from stratum to stratum. However, $\tilde{t}_{yGHR}$ is between the two estimators. $\tilde{t}_{yGHR}$ is exactly unbiased estimator for t_y and the correct model when the stratified sampling design.

Table 3.8 shows MSE ratios for $\hat{t}_{yComb,\hat{\theta}}$, and $\hat{t}_{ySep,\hat{\theta}}$ relative to $\hat{t}_{ySep,GHR}$. For $\hat{t}_{yComb,\hat{\theta}}$, the ratios are much larger than one, reflecting the large inefficiency of using common ratio model for this population. The ratios actually get worse with increasing stratum sample size or increasing number of strata, since the failure of the common ratio model becomes more evident in either case.

For $\hat{t}_{ySep,\hat{\theta}}$, the MSE ratios in Table 3.8 show that the separate ratio estimator performs better with respect to the separate GHR estimator as the sample size increases within strata, but worse for fixed sample size as the number of strata increases. This is due to the accumulation of bias in $\hat{t}_{ySep,\hat{\theta}}$ across the strata.

Coverage of the nominal 95% confidence intervals for $\hat{t}_{yComb,\hat{\theta}}$ is close to 95% for all cases considered, with better coverage at larger sample sizes, but the average length of these confidence intervals is much greater than those of $\hat{t}_{ySep,GHR}$ or $\hat{t}_{ySep,\hat{\theta}}$.

Average length of confidence interval for $\hat{t}_{ySep,\hat{\theta}}$ are smallest among those considered, but coverage is far less than the nominal 95% in all cases with $\hat{t}_{ySep,\hat{\theta}}$. Coverage increases with increasing sample size within strata, but decreases with increasing number of strata, due to accumulation of bias across strata. Average length of the confidence intervals for $\hat{t}_{ySep,GHR}$ are somewhat larger than for $\hat{t}_{ySep,\hat{\theta}}$ but much less than $\hat{t}_{yComb,\hat{\theta}}$. Still, the coverage of the confidence intervals is close to the nominal 95% in all cases, improves with increased sample sizes within strata, and is not adversely affected by increasing the number of strata, unlike $\hat{t}_{ySep,\hat{\theta}}$.

H	Est	%Coverage $n_h=4$	%Coverage $n_h=5$	%Coverage $n_h=6$	CI Length $n_h=4$	CI Length $n_h=5$	CI Length $n_h=6$	%RVB $n_h=4$	%RVB $n_h=5$	%RVB $n_h=6$
10	$\hat{t}_{ySep,\hat{\theta}}$	80.1	82.3	86.5	1.258	1.155	1.059	-31.69	-29.33	-20.76
	$\hat{t}_{ySep,GHR}$	86.9	89.7	91.9	1.776	1.618	1.448	-2.45	-3.59	1.78
	$\hat{t}_{yComb,\hat{\theta}}$	92.0	93.4	93.5	8.677	7.695	6.946	-5.92	-3.32	0.857
20	$\hat{t}_{ySep,\hat{\theta}}$	78.9	82.7	86.9	0.917	0.83	0.766	-27.36	-23.01	-13.53
	$\hat{t}_{ySep,GHR}$	91.7	93.3	94.1	1.311	1.139	1.029	1.188	-3.4	5.033
	$\hat{t}_{yComb,\hat{\theta}}$	94.0	94.5	94.7	8.095	7.154	6.444	2.717	5.826	5.484
30	$\hat{t}_{ySep,\hat{\theta}}$	76.3	79.7	83.3	0.771	0.696	0.639	-26.01	-20.79	-22.13
	$\hat{t}_{ySep,GHR}$	93.7	94.5	93.9	1.086	0.943	0.846	3.328	0.908	-4.48
	$\hat{t}_{yComb,\hat{\theta}}$	93.8	94.4	94.0	7.837	6.928	6.246	4.383	4.155	-0.321
40	$\hat{t}_{ySep,\hat{\theta}}$	73.9	80.3	83.0	0.725	0.654	0.599	-26.55	-24.17	-22.85
	$\hat{t}_{ySep,GHR}$	94.1	93.1	93.7	1.003	0.871	0.781	0.022	-4.78	-4.91
	$\hat{t}_{yComb,\hat{\theta}}$	93.2	93.8	93.9	7.754	6.855	6.184	1.074	-3.36	-2.67

Table 3.7: Percentages of Confidence Intervals Covering t_y under $STSI$ sampling design.

H		$n_h=4$	$n_h=5$	$n_h=6$
10	$MSE\left(\hat{t}_{ySep,\hat{\theta}}\right)/MSE\left(\hat{t}_{ySep,GHR}\right)$	0.832	0.816	0.774
	$MSE\left(\hat{t}_{yComb,\hat{\theta}}\right)/MSE\left(\hat{t}_{ySep,GHR}\right)$	22.167	21.436	22.346
20	$MSE\left(\hat{t}_{ySep,\hat{\theta}}\right)/MSE\left(\hat{t}_{ySep,GHR}\right)$	1.014	0.897	0.854
	$MSE\left(\hat{t}_{yComb,\hat{\theta}}\right)/MSE\left(\hat{t}_{ySep,GHR}\right)$	36.112	35.312	38.573
30	$MSE\left(\hat{t}_{ySep,\hat{\theta}}\right)/MSE\left(\hat{t}_{ySep,GHR}\right)$	1.222	1.109	0.991
	$MSE\left(\hat{t}_{yComb,\hat{\theta}}\right)/MSE\left(\hat{t}_{ySep,GHR}\right)$	50.545	51.786	51.921
40	$MSE\left(\hat{t}_{ySep,\hat{\theta}}\right)/MSE\left(\hat{t}_{ySep,GHR}\right)$	1.344	1.175	1.0825
	$MSE\left(\hat{t}_{yComb,\hat{\theta}}\right)/MSE\left(\hat{t}_{ySep,GHR}\right)$	58.519	60.838	61.142

Table 3.8: MSE Ratios under STSI design based on 1500 replicates.

Table 3.10 shows MSE ratios for $\hat{t}_{yComb,\hat{\theta}}$, and $\hat{t}_{ySep,\hat{\theta}}$ relative to $\hat{t}_{ySep,GHR}$. For $\hat{t}_{yComb,\hat{\theta}}$, the ratios are much larger than one, reflecting the large inefficiency of

H	Est	%Coverage $n_h=4$	%Coverage $n_h=5$	%Coverage $n_h=6$	CI Length $n_h=4$	CI Length $n_h=5$	CI Length $n_h=6$	%RVB $n_h=4$	%RVB $n_h=5$	%RVB $n_h=6$
10	$\hat{t}_{ySep,\hat{\theta}}$	69.9	74.2	74.8	1.278	1.2	1.116	-48	-38.637	-41.86
	$\hat{t}_{ySep,GHR}$	97.9	95.3	93.5	3.463	2.396	1.812	77.506	6.331	-21.103
	$\hat{t}_{yComb,\hat{\theta}}$	91.2	92.5	92.5	11.224	10.032	9.12	-7.415	-6.305	-9.852
20	$\hat{t}_{ySep,\hat{\theta}}$	62.8	70.9	76.2	0.924	0.862	0.808	-44.436	-39.079	-32.197
	$\hat{t}_{ySep,GHR}$	99.3	98.2	96.5	3.011	2.071	1.547	162.725	70.306	25.564
	$\hat{t}_{yComb,\hat{\theta}}$	94.0	94.1	92.7	10.067	8.986	8.182	0.934	-6.738	-7.214
30	$\hat{t}_{ySep,\hat{\theta}}$	59.3	65.4	71.3	0.779	0.725	0.676	-39.392	-34.517	-33.536
	$\hat{t}_{ySep,GHR}$	99.7	99.6	97.9	2.759	1.899	1.418	204.748	142.712	42.375
	$\hat{t}_{yComb,\hat{\theta}}$	94.0	94.3	94.4	9.566	8.536	7.774	3.681	-0.53	2.803
40	$\hat{t}_{ySep,\hat{\theta}}$	54.2	61.3	68.1	0.734	0.678	0.631	-41.927	-35.648	-32.044
	$\hat{t}_{ySep,GHR}$	99.9	99.7	99.1	2.61	1.815	1.362	257.24	157.061	78.992
	$\hat{t}_{yComb,\hat{\theta}}$	93.9	93.3	94.4	9.416	8.41	7.667	1.067	2.11	0.152

Table 3.9: Percentages of Confidence Intervals Covering t_y under $ST\pi ps$ Sampling Design.

H		$n_h=4$	$n_h=5$	$n_h=6$
10	$MSE\left(\hat{t}_{ySep,\hat{\theta}}\right)/MSE\left(\hat{t}_{ySep,GHR}\right)$	0.766	0.648	0.729
	$MSE\left(\hat{t}_{yComb,\hat{\theta}}\right)/MSE\left(\hat{t}_{ySep,GHR}\right)$	20.506	20.054	22.308
20	$MSE\left(\hat{t}_{ySep,\hat{\theta}}\right)/MSE\left(\hat{t}_{ySep,GHR}\right)$	0.931	0.899	0.832
	$MSE\left(\hat{t}_{yComb,\hat{\theta}}\right)/MSE\left(\hat{t}_{ySep,GHR}\right)$	29.252	34.476	37.837
30	$MSE\left(\hat{t}_{ySep,\hat{\theta}}\right)/MSE\left(\hat{t}_{ySep,GHR}\right)$	1.031	1.224	0.959
	$MSE\left(\hat{t}_{yComb,\hat{\theta}}\right)/MSE\left(\hat{t}_{ySep,GHR}\right)$	35.456	49.38	41.65
40	$MSE\left(\hat{t}_{ySep,\hat{\theta}}\right)/MSE\left(\hat{t}_{ySep,GHR}\right)$	1.465	1.442	1.269
	$MSE\left(\hat{t}_{yComb,\hat{\theta}}\right)/MSE\left(\hat{t}_{ySep,GHR}\right)$	46.148	54.141	56.774

Table 3.10: MSE Ratios under STπps design based on 1500 replicates

using common ratio model for this population. The ratios actually get worse with increasing number of strata, since the failure of the common ratio model becomes more evident in either case.

For $\hat{t}_{ySep,\hat{\theta}}$, the MSE ratios in Table 3.10 show that the separate ratio estimator performs better with respect to the separate GHR estimator as the sample size increases within strata, but worse for fixed sample size as the number of strata increases. This is due to the accumulation of bias in $\hat{t}_{ySep,\hat{\theta}}$ across the strata.

Coverage of the nominal 95% confidence intervals for $\hat{t}_{yComb,\hat{\theta}}$ is close to 95% for all cases considered, but the average length of these confidence intervals is much greater than those of $\hat{t}_{ySep,GHR}$ or $\hat{t}_{ySep,\hat{\theta}}$. Average length of confidence interval for

$\hat{t}_{ySep,\hat{\theta}}$ are smallest among those considered, but coverage is far less than the nominal 95% in all cases with $\hat{t}_{ySep,\hat{\theta}}$. Coverage increases with increasing sample size within strata, but decreases with increasing number of strata, due to accumulation of bias across strata. Average length of the confidence intervals for $\hat{t}_{ySep,GHR}$ are somewhat larger than for $\hat{t}_{ySep,\hat{\theta}}$ but much less than $\hat{t}_{yComb,\hat{\theta}}$. Still, the coverage of the confidence intervals is close to the nominal 95% in all cases, improves with increased sample sizes within strata, and is not adversely affected by increasing the number of strata, unlike $\hat{t}_{ySep,\hat{\theta}}$.

As a final comment, the bias affect the percentage of coverage of $\hat{t}_{ysep}$ and keep it far a way from the nominal value %95, and $\tilde{t}_{yGHR}$ works better than $\hat{t}_{ysep}$. However, the percentage coverage of $\hat{t}_{ysep}$ can be worst than the given simulation results by decreasing the strata size (in the given simulation are fix and equal to 50) and by increasing number of strata. In the same time, the percentage coverage of $\tilde{t}_{yGHR}$ will not affected by decreasing the strata size or increasing number of strata.

Chapter 4

CONCLUSIONS

In this work, the Hartley and Ross (1954) estimator has been generalized to unequal-probability sampling designs, under the condition of measurability (strictly positive second-order inclusion probabilities). This results in generalized Hartley and Ross (GHR) estimation. Two distinct versions were considered, ane building on the Horvitz and Thompson (1952) estimator and one building on the Hansen and Hurwitz (1943) estimator for with-replacement sampling.

In Horvitz-Thompson based GHR estimator is of interest because it is unbiased and an exact variance and an unbiased estimator for the exact variance were obtained. The computations for the exact variance and the unbiased variance estimator of the GHR require higher-order inclusion probabilities (up to fourth order), which are not easily obtained in general. To overcome this problem, two methods of approximation were given. Simulation results for SI, PO, and WR sampling indicate that these are useful approximations.

The GHR estimator was shown to be mean square consistent under mild conditions. These conditions are met by simple random sampling without replacement, simple random cluster sampling, and stratified sampling designs.

Central limit theorems (CLTs) were established for GHR under the SI design and under the Poisson sampling (PO) design. The asymptotic variance and a consistent estimator for the asymptotic variance were given under the two sampling designs SI and PO.

The GHR was evaluated under a super-population model, and it was shown that the Godambe and Joshi (1965) lower bound is attainable for GHR under SI and PO sampling designs. The GHR is thus asymptotically equivalent to the standard, simple ratio estimator, and superior to the naive ratio estimator. The GHR can be combined with simple estimator to produce other estimators that have approved in the literature including Hartley and Ross (1954), Murthy and Nanjamma (1959) and Nieto de Pascual (1961). The GHR was compared to other estimators analytically and via simulation.

The version of GHR derived using a Hansen and Hurwitz (1943) type estimator for with-replacement sampling was shown to be unbiased. This estimator was discussed under two different asymptotic scenarios, when the population size N is fixed and number of independent draws m tends to infinity and when both m and N tend to infinity. Under each of the two cases, a CLT was established and the asymptotic variance and a consistent estimator for the variance were given. The Godambe and Joshi (1965) lower bound was shown to be almost attainable under the first case and attainable for the second case.

An important problem is estimation of the population total t_y under a stratified sampling design when stratum x-totals are known, particularly in the case of small stratum sizes. If biased estimators are used to estimate within-stratum population y-totals, the bias may accumulate across strata. The unbiased GHR estimators were adapted to deal with such situations by redefining the GHR as a separate GHR estimator, analogous to the classic separate ratio estimator of survey statistics. A CLT was proven for the separate GHR estimator under a stratified sampling design when the stratum sizes are fixed and the number of strata tends to infinity. Simulation results showed that GHR under different sampling designs gives excellent results compared to other almost unbiased estimators proposed in the literature, even when the number of strata is not large.

The work in this dissertation can be extended in a number of directions. Work on GHR-type estimators could be pursued in at least five directions. First, simulation results indicate that the two methods of approximate variance estimation behave very similarly. The first approximate variance estimator was shown to be consistent for the true variance. Under what conditions is the second approximate variance estimator, based on a with-replacement approximation, consistent for the true variance? Second, it would be of interest to weaken the conditions used in establishing central limit theory for GHR, while still giving conditions that are relatively easy to check in practice. Third, we have shown asymptotic attainment of the Godambe-Joshi lower bound under SI, PO, and WR sampling designs. Does the GJLB always hold for the GHR estimators, and if not, under what designs or models does it fail? Fourth, simulation results for both unstratified and stratified designs showed good performance of GHR relative to standard estimators under particular simulated finite populations. These simulations were quite limited, however, and it would be of interest to conduct a broader simulation study to give guidance to users on appropriate choices of ratio estimators. Fifth, and finally, it would be of interest to apply GHR to real data.

In addition to extending results for GHR, it would also be of interest in future work to apply the theoretical methods used in this dissertation to other ratio estimators, to produce a thorough, systematic study of the properties of the Hartley and Ross (1954), Murthy and Nanjamma (1959), and Nieto de Pascual (1961) estimators in unstratified and stratified cases. Asymptotic results could be derived and simulation studies undertaken. Further, the unbiased estimation techniques used to develop GHR in the presence of auxiliary information might be extended from ratios to other kinds of population parameters, including variances and regression coefficients. These would require additional auxiliary information.

Appendix A

MEAN SQUARE CONSISTENCY EXAMPLES

In this section we will give details that were omitted in previous chapters.

Example A.0.1 Assume $n \sim O\left(N^{\delta}\right)$, for $\frac{5}{6} < \delta \leq 1$. Then $\hat{\theta}_{GHR}$ is mean square consistent estimator for θ under SI sampling design. If $\delta = 1$, then the finite population corrections $\left(fpc = 1 - \frac{n}{N}\right)$ cannot be ignored and we can ignore it if $\delta < 1$.

Solution For SI sampling design,

$$\begin{aligned}
\pi_i &= \frac{n}{N}, \\
\pi_{ij} &= \frac{n}{N}\frac{n-1}{N-1} \qquad \text{for} \quad ij \in U_{D_2}, \\
\pi_{N*} &= \frac{n}{N}\frac{n-1}{N-1}, \\
&\text{and} \\
\pi_{ijkl} &= \frac{n}{N}\frac{n-1}{N-1}\frac{n-2}{N-2}\frac{n-3}{N-3} \qquad \text{for} \quad ijkl \in U_{D_4}.
\end{aligned}$$

Therefore,

$$\sum_{i \in U} \pi_i^2 = O\left(N^{2\delta-1}\right),$$

and

$$\begin{aligned}
\sum_{ij \in U_{D_2}} \left(\pi_{ij} - \pi_i \pi_j\right)^2 &= \sum_{ij \in U_{D_2}} \left[\frac{n}{N}\frac{n-1}{N-1} - \left(\frac{n}{N}\right)^2\right]^2 \\
&= N(N-1)\left(\frac{n}{N}\right)^2 \left(\frac{N(n-1) - n(N-1)}{N(N-1)}\right)^2 \\
&= N(N-1)\left(\frac{-n}{N}\right)^2 \left(1 - \frac{n}{N}\right)^2 \frac{1}{(N-1)^2} \\
&= O\left(N^{2\delta-2}\right).
\end{aligned}$$

Hence, $\eta^* = \max\{2\delta - 1, 2\delta - 2\} = 2\delta - 1 < 2$.

To find η, consider the following two cases:

$$\begin{aligned}\sum_{ij\in U_{D_2}} N\pi_{ij}^2 &= \sum_{ij\in U_{D_2}} N\left(\frac{n}{N}\frac{n-1}{N-1}\right)^2 \\ &= O\left(N^{4\delta-1}\right),\end{aligned}$$

and

$$\begin{aligned}\sum_{ijkl\in U_{D_4}} [\pi_{ijkl} - \pi_{ij}\pi_{kl}]^2 &= O\left(N^4\right)\left(\frac{n}{N}\frac{n-1}{N-1}\right)^2\left[\frac{n-2}{N-2}\frac{n-3}{N-3} - \frac{n}{N}\frac{n-1}{N-1}\right]^2 \\ &= O\left(N^{4\delta}\right)\left[\frac{-4nN^2 + 4n^2N + 6\left(N^2 - n^2\right) - 6\left(N - n\right)}{N\left(N-1\right)\left(N-2\right)\left(N-3\right)}\right]^2 \\ &= O\left(N^{6\delta-4}\right).\end{aligned}$$

Therefore, $\eta = \max\{6\delta - 4, 4\delta - 1\} = 4\delta - 1 < 4$. Since

$$min\left\{1, \frac{2-\eta^*}{4}, \frac{4-\eta}{4}\right\} = min\left\{1, \frac{3-2\delta}{4}, \frac{5-4\delta}{4}\right\} = \frac{3-2\delta}{4},$$

then

$$N^{min\left\{1, \frac{2-\eta^*}{4}, \frac{4-\eta}{4}\right\}}\pi_{N*} = O\left(N^{\frac{6\delta-5}{4}}\right) \to \infty, \text{ as } N \to \infty \text{ and for } \frac{5}{6} < \delta \leq 1.$$

Hence the SI sampling design is a mean square consistent. If $\delta = 1$, then the finite population corrections $\left(fpc = 1 - \frac{n}{N}\right)$ can not be ignored and we can ignore it if $\delta < 1$.

Example A.0.2 Consider simple random cluster sampling design (SIC). Under this design, M is the number of clusters, C is the cluster size, $N = MC$ is the population size, and draw m clusters from the M clusters via SI design and observe all elements in each selected cluster. Assume $m \sim O\left(M^\delta\right)$, for $\frac{5}{6} < \delta \leq 1$, then $\hat{\theta}_{GHR}$ is mean square consistent estimator for θ.

Solution. Let U_h be the h^{th} cluster, $h = 1, 2, \ldots, M$. Now, for $i \in U_h$, we have

$$\pi_i = \frac{m}{M},$$

$$\begin{aligned} \sum_{h=1}^{M} \sum_{i \in U_h} \pi_i^2 &= MC \left(\frac{m}{M}\right)^2 \\ &= O\left(M^{2\delta-1}\right), \end{aligned}$$

$$\pi_{ij} = \pi_i = \pi_j = \frac{m}{M} \qquad \forall i \neq j \in U_h,$$

and

$$\begin{aligned} \sum_{h=1}^{M} \sum_{i \neq j \in U_h} \left(\pi_{ij} - \pi_i \pi_j\right)^2 &= MC\,(c-1) \left(\frac{m}{M}\right)^2 \left(1 - \frac{m}{M}\right)^2 \\ &= O\left(M^{2\delta-1}\right). \end{aligned}$$

If $h \neq \acute{h}$ and $i \in U_h, j \in U_{\acute{h}}$ then

$$\begin{aligned} \pi_i &= \pi_j = \frac{m}{M}, \\ \pi_{ij} &= \frac{m}{M} \frac{m-1}{M-1}, \end{aligned}$$

and

$$\begin{aligned} \sum_{h \neq \acute{h}} \sum_{i \in U_h, j \in U_{\acute{h}}} \left(\pi_{ij} - \pi_i \pi_j\right)^2 &= M\,(M-1)\,C^2 \left(\frac{m}{M}\right)^2 \left[\frac{M\,(m-1) - m\,(M-1)}{M\,(M-1)}\right]^2 \\ &= O\left(M^{2\delta-2}\right). \end{aligned}$$

Therefore, $\eta^* = \max\{2\delta - 1, 2\delta - 2\} = 2\delta - 1 < 2$.

If $i \neq j \in U_h$ then $\pi_{ij} = \frac{m}{M}$; then

$$\sum_{h} \sum_{i \neq j \in U_h} MC\pi_{ij}^2 = O\left(M^{2\delta}\right),$$

and for $h \neq \acute{h}$ and $i \in U_h, j \in U_{\acute{h}}$,

$$\pi_{ij} = \frac{m}{M} \frac{m-1}{M-1},$$

and

$$\sum_{h \neq \acute{h}} \sum_{ij} MC\pi_{ij}^2 = O\left(M^{4\delta-1}\right).$$

- If $ijkl \in U_h$ and $ijkl \in D_4$ then

$$\pi_{ijkl} = \pi_{ij} = \pi_{kl} = \frac{m}{M},$$

and then

$$\begin{aligned}\sum_{h=1}^{M} \sum_{ijkl \in U_h} \Delta_{ijkl}^2 &= MC(C-1)(C-2)(C-3)\left(\frac{m}{M}\right)^2\left(1-\frac{m}{M}\right)^2 \\ &= O\left(M^{2\delta-1}\right).\end{aligned}$$

- Two different clusters and two elements in each cluster: U_h, $U_{\acute{h}}$, and $h, \acute{h} \in D_2$.

  - Case(1): $ij \in U_h$ and $kl \in U_{\acute{h}}$.

$$\begin{aligned}\pi_{ij} &= \pi_{kl} = \frac{m}{M}, \\ \pi_{ijkl} &= \frac{m}{M}\frac{m-1}{M-1}, \\ \Delta_{ijkl} &= -\frac{m}{M}\left(1-\frac{m}{M}\right)\frac{1}{M-1},\end{aligned}$$

and

$$\sum_{h,\acute{h} \in D_2} \sum_{ij \in U_h, kl \in U_{\acute{h}}} \Delta_{ijkl}^2 = O\left(M^{2\delta-2}\right).$$

  - Case(2): $ik \in U_h$ and $jl \in U_{\acute{h}}$.

$$\begin{aligned}\pi_{ijkl} &= \pi_{ij} = \pi_{kl} = \frac{m}{M}\frac{m-1}{M-1}, \\ \Delta_{ijkl} &= \frac{m}{M}\frac{m-1}{M-1}\left(1-\frac{m}{M}\frac{m-1}{M-1}\right),\end{aligned}$$

and

$$\sum_{h,\acute{h} \in D_2} \sum_{ik \in U_h, jl \in U_{\acute{h}}} \Delta_{ijkl}^2 = O\left(M^{4\delta-2}\right).$$

  - Case(3): $il \in U_h$ and $jk \in U_{\acute{h}}$.

$$\pi_{ijkl} = \pi_{ij} = \pi_{kl} = \frac{m}{M}\frac{m-1}{M-1},$$

and

$$\sum_{h,\acute{h} \in D_2} \sum_{il \in U_h, jk \in U_{\acute{h}}} \Delta_{ijkl}^2 = O\left(M^{4\delta-2}\right).$$

- Two different clusters and three elements in one cluster and one element in the other one:

  - Case(1): $ijk \in U_h$ and $l \in U_{\acute{h}}$.

$$\begin{aligned} \pi_{ij} &= \frac{m}{M}, \\ \pi_{ijkl} &= \pi_{kl} = \frac{m}{M}\frac{m-1}{M-1}, \\ \Delta_{ijkl} &= \frac{m}{M}\frac{m-1}{M-1}\left[1-\frac{m}{M}\right], \end{aligned}$$

and

$$\sum_{h,\acute{h}\in D_2}\sum_{ijk\in U_h, l\in U_{\acute{h}}} \Delta_{ijkl}^2 = O\left(M^{4\delta-2}\right).$$

  - Case(2): $ijl \in U_h$ and $k \in U_{\acute{h}}$

$$\begin{aligned} \pi_{ij} &= \frac{m}{M}, \\ \pi_{ijkl} &= \pi_{kl} = \frac{m}{M}\frac{m-1}{M-1}, \end{aligned}$$

and

$$\sum_{h,\acute{h}\in D_2}\sum_{ijl\in U_h, k\in U_{\acute{h}}} \Delta_{ijkl}^2 = O\left(M^{4\delta-2}\right).$$

  - Case(3): $ikl \in U_h$ and $j \in U_{\acute{h}}$.

$$\begin{aligned} \pi_{kl} &= \frac{m}{M}, \\ \pi_{ijkl} &= \pi_{ij} = \frac{m}{M}\frac{m-1}{M-1}, \end{aligned}$$

and

$$\sum_{h,\acute{h}\in D_2}\sum_{ikl\in U_h, j\in U_{\acute{h}}} \Delta_{ijkl}^2 = O\left(M^{4\delta-2}\right).$$

  - Case(4): $jkl \in U_h$ and $i \in U_{\acute{h}}$.

$$\begin{aligned} \pi_{kl} &= \frac{m}{M}, \\ \pi_{ijkl} &= \pi_{ij} = \frac{m}{M}\frac{m-1}{M-1}, \end{aligned}$$

and

$$\sum_{h,\acute{h}\in D_2}\sum_{jkl\in U_h, i\in U_{\acute{h}}}\Delta_{ijkl}^2 \;=\; O\left(M^{4\delta-2}\right).$$

- Three different clusters: U_h; $U_{\acute{h}}$; $U_{\acute{\acute{h}}}$, and $h,\acute{h},\acute{\acute{h}}\in D_3$.

  - Case(1): $ij\in U_h$ and $k\in U_{\acute{h}}$, and $l\in U_{\acute{\acute{h}}}$.

$$\begin{aligned}
\pi_{ij} &= \frac{m}{M},\\
\pi_{kl} &= \frac{m}{M}\frac{m-1}{M-1},\\
\pi_{ijkl} &= \frac{m}{M}\frac{m-1}{M-1}\frac{m-2}{M-2},\\
\Delta_{ijkl} &= -2\left(\frac{m}{M}\right)\left(\frac{m-1}{M-1}\right)\left(1-\frac{m}{M}\right)\frac{1}{(M-2)},
\end{aligned}$$

and

$$\sum_{h,\acute{h},\acute{\acute{h}}\in D_3}\sum_{ij\in U_h, k\in U_{\acute{h}}, l\in U_{\acute{\acute{h}}}}\Delta_{ijkl}^2 \;=\; O\left(M^{4\delta-3}\right).$$

  - Case(2): $ik\in U_h$ and $j\in U_{\acute{h}}$, and $l\in U_{\acute{\acute{h}}}$.

$$\begin{aligned}
\pi_{ij} &= \pi_{kl} = \frac{m}{M}\frac{m-1}{M-1},\\
\pi_{ijkl} &= \frac{m}{M}\frac{m-1}{M-1}\frac{m-2}{M-2},\\
\Delta_{ijkl} &= O\left(M^{3\delta-3}\right),
\end{aligned}$$

and

$$\sum_{h,\acute{h},\acute{\acute{h}}\in D_3}\sum_{ik\in U_h, j\in U_{\acute{h}}, l\in U_{\acute{\acute{h}}}}\Delta_{ijkl}^2 \;=\; O\left(M^{6\delta-3}\right).$$

  - Case(3): $il\in U_h$, $j\in U_{\acute{h}}$ and $k\in U_{\acute{\acute{h}}}$.

$$\begin{aligned}
\pi_{ij} &= \pi_{kl} = \frac{m}{M}\frac{m-1}{M-1},\\
\pi_{ijkl} &= \frac{m}{M}\frac{m-1}{M-1}\frac{m-2}{M-2},
\end{aligned}$$

and

$$\sum_{h,\acute{h},\acute{\acute{h}}\in D_3}\sum_{il\in U_h, j\in U_{\acute{h}}, k\in U_{\acute{\acute{h}}}}\Delta_{ijkl}^2 \;=\; O\left(M^{6\delta-3}\right).$$

– Case(4): $jk \in U_h$, $i \in U_{\acute{h}}$ and $l \in U_{\acute{\acute{h}}}$.

$$\begin{aligned} \pi_{ij} &= \pi_{kl} = \frac{m}{M}\frac{m-1}{M-1}, \\ \pi_{ijkl} &= \frac{m}{M}\frac{m-1}{M-1}\frac{m-2}{M-2}, \end{aligned}$$

and

$$\sum_{h,\acute{h},\acute{\acute{h}} \in D_3} \sum_{jk \in U_h, i \in U_{\acute{h}}, l \in U_{\acute{\acute{h}}}} \Delta^2_{ijkl} = O\left(M^{6\delta-3}\right).$$

– Case(5): $jl \in U_h$, $i \in U_{\acute{h}}$ and $k \in U_{\acute{\acute{h}}}$.

$$\begin{aligned} \pi_{ij} &= \pi_{kl} = \frac{m}{M}\frac{m-1}{M-1}, \\ \pi_{ijkl} &= \frac{m}{M}\frac{m-1}{M-1}\frac{m-2}{M-2}, \end{aligned}$$

and

$$\sum_{h,\acute{h},\acute{\acute{h}} \in D_3} \sum_{jl \in U_h, i \in U_{\acute{h}}, k \in U_{\acute{\acute{h}}}} \Delta^2_{ijkl} = O\left(M^{6\delta-3}\right).$$

– Case(6): $kl \in U_h$, $i \in U_{\acute{h}}$ and $j \in U_{\acute{\acute{h}}}$.

$$\begin{aligned} \pi_{kl} &= \frac{m}{M}, \\ \pi_{ij} &= \frac{m}{M}\frac{m-1}{M-1}, \\ \pi_{ijkl} &= \frac{m}{M}\frac{m-1}{M-1}\frac{m-2}{M-2}, \end{aligned}$$

and

$$\sum_{h,\acute{h},\acute{\acute{h}} \in D_3} \sum_{kl \in U_h, i \in U_{\acute{h}}, j \in U_{\acute{\acute{h}}}} \Delta^2_{ijkl} = O\left(M^{4\delta-3}\right).$$

- Four different clusters: $i \in U_h$, $j \in U_{\acute{h}}$, $k \in U_{\acute{\acute{h}}}$, $l \in U_{\acute{\acute{\acute{h}}}}$ and $\left(h, \acute{h}, \acute{\acute{h}}, \acute{\acute{\acute{h}}}\right) \in D_4$.

$$\begin{aligned} \pi_{ij} &= \pi_{kl} = \frac{m}{M}\frac{m-1}{M-1}, \\ \pi_{ijkl} &= \frac{m}{M}\frac{m-1}{M-1}\frac{m-2}{M-2}\frac{m-3}{M-3}, \end{aligned}$$

$$\Delta_{ijkl} = \frac{m}{M}\frac{m-1}{M-1}\left(1-\frac{m}{M}\right)\frac{6M+6m-6-4mM}{(M-1)(M-2)(M-3)},$$

and

$$= O\left(M^{3\delta-4}\right)$$

$$\sum_{h,\acute{h},\acute{\acute{h}},\acute{\acute{\acute{h}}}\in D_4}\sum_{ijkl}\Delta_{ijkl}^2 = O\left(M^{6\delta-4}\right).$$

Therefore,

$$\eta = \max\{2\delta, 4\delta-1, 2\delta-1, 2\delta-2, 4\delta-2, 4\delta-3, 6\delta-3, 6\delta-4\}$$
$$= 4\delta-1 \qquad \text{for} \quad \frac{5}{6} < \delta \leq 1.$$

Since

$$min\left\{1, \frac{2-\eta^*}{4}, \frac{4-\eta}{4}\right\} = min\left\{1, \frac{3-2\delta}{4}, \frac{5-4\delta}{4}\right\} = \frac{3-2\delta}{4},$$

and under this sampling design, $\pi_{M*} = \frac{m}{M}\frac{m-1}{M-1}$ Then,

$$M^{min\left\{1, \frac{2-\eta^*}{4}, \frac{4-\eta}{4}\right\}}\pi_{M*} = O\left(M^{\frac{6\delta-5}{4}}\right) \to \infty, \text{ as } M \to \infty \text{ and for } \frac{5}{6} < \delta \leq 1.$$

Example A.0.3 For $0 < \delta \leq 1$, $\hat{\theta}_{GHR}$ is mean square consistent estimator for θ under ST (stratified sampling) design. Let $H_N \sim O\left(N^{\delta}\right)$ be the number of strata, $N_h \sim O\left(N^{1-\delta}\right)$ be the h^{th} stratum size.

Solution

Let U_h be the h^{th} stratum and $h = 1, 2, \ldots, H_N$. Now, for $i \in U_h$, we have

$$\sum_{h=1}^{H_N}\sum_{i\in U_h}\pi_i^2 \leq O(N).$$

If $i \neq j \in U_h$ then

$$\sum_{h=1}^{H_N}\sum_{i\neq j\in U_h}(\pi_{ij}-\pi_i\pi_j)^2 \leq O\left(N^{2-\delta}\right).$$

If $h \neq \acute{h},\ i \in U_h, j \in U_{\acute{h}}$ then $\pi_{ij} = \pi_i \pi_j$. Therefore,

$\sum_{h \neq \acute{h}} \sum_{i \in U_h, j \in U_{\acute{h}}} (\pi_{ij} - \pi_i \pi_j)^2 = O(1)$.

Hence,

$$\eta^* = \max\{1, 0, 2 - \delta\} = 2 - \delta, \text{ for } 0 < a < \delta \leq 1.$$

- Find the order of $\sum_{ij \in U_{D_2}} N \pi_{ij}^2$.

  If $i \neq j \in U_h$ then

  $$\sum_h \sum_{i \neq j \in U_h} N \pi_{ij}^2 \ \leq \ O\left(N^{3-\delta}\right),$$

  and for $h \neq \acute{h}$, $i \in U_h$, and $j \in U_{\acute{h}}$ then

  $$\sum_{h,\acute{h}} \sum_{ij} N \pi_{ij}^2 \ \leq \ O\left(N^3\right).$$

- Find the order of $\sum_{ijkl \in U_{D_4}} \Delta_{ijkl}^2$, where $\Delta_{ijkl} = \pi_{ijkl} - \pi_{ij}\pi_{kl}$.

  - For $ijkl \in U_{D_4}$ and $ijkl \in U_h$ then

    $$\sum_{h=1}^{H_N} \sum_{ijkl \in U_h} \Delta_{ijkl}^2 \ \leq \ O\left(N^{4-3\delta}\right).$$

  - Two different Strata, $U_h, U_{\acute{h}},\ h \neq \acute{h}$; and two elements in each strata

    * Case(1): $ij \in U_h$ and $kl \in U_{\acute{h}}$

      $$\sum_{h,\acute{h}} \sum_{ijkl} \Delta_{ijkl}^2 \ = \ O(1).$$

    * Case(2): $ik \in U_h$ and $jl \in U_{\acute{h}}$

      $$\sum_{h,\acute{h}} \sum_{ijkl} \Delta_{ijkl}^2 \ \leq \ O\left(N^{4-2\delta}\right).$$

      For other cases, they have the same order.

  - Three different Strata $U_h, U_{\acute{h}}, U_{\acute{\acute{h}}}$ and $h, \acute{h}, \acute{\acute{h}} \in D_3$.

* Case(1): $i \neq j \in U_h$ and $k \in U_{\acute{h}}$ and $l \in \acute{\acute{h}}$.

$$\sum_{h,\acute{h}} \sum_{ijkl} \Delta_{ijkl}^2 \;=\; O(1).$$

* Case(2): $i \neq k \in U_h$ and $j \in U_{\acute{h}}$ and $l \in \acute{\acute{h}}$.

$$\sum_{h,\acute{h},\acute{\acute{h}}} \sum_{ijkl} \Delta_{ijkl}^2 \;\leq\; O\left(N^{4-\delta}\right).$$

Other cases are either of order $O(1)$ or $O\left(N^{4-\delta}\right)$.

– Four different clusters: $i \in U_h$ and $j \in U_{h_1}$ and $k \in U_{h_2}$ and $l \in U_{h_3}$ and $h,\, h_1,\, h_2,\, h_3 \in D_4$.

$$\sum_{h,\acute{h},\acute{\acute{h}},\dddot{h}} \sum_{ijkl} \Delta_{ijkl}^2 \;=\; O(1).$$

Therefore, $\eta = \max\{3, 3-\delta, 4-3\delta, 0, 4-2\delta, 4-\delta\} = 4-\delta$. Since $\min\left\{1, \frac{2-\eta^*}{4}, \frac{4-\eta}{4}\right\} = \frac{\delta}{4}$, then choose δ such that $\delta > 0$ and $N^{\frac{\delta}{4}}\pi_{N*} \to \infty$ as $N \to \infty$. Hence this design is a mean square consistent.

Appendix B

NOTATION

The following notations are used in this work.

$U_N = \{1, 2, \ldots, N\}$	Finite Population
U_h	h^{th} Stratum or Cluster
$t_y = \sum_{i \in U} y_i$	Population total t_y
$\theta = \frac{t_y}{t_x}$	Population ratio totals(means) of the variables y and x
β	Superpopulation parameter
s	Probability sample
n_S	Sample size
$p(s)$	Probability of drawing the sample s from U_N
$p(\cdot)$	Fix measurable sampling design
π_i	First order inclusion probability of element i
π_{ij}	Second order inclusion probability of elements i and j
π_{ijk}	Third order inclusion probability of elements i, j, and k
π_{ijkl}	Fourth order inclusion probability of elements i, j, k, and l
I_k	Sample membership indicators
$\Delta_{ij} = \pi_{ij} - \pi_i \pi_j$	Covariance of I_i and I_j
$\Delta_{ijk} = \pi_{ijk} - \pi_i \pi_{jk}$	Covariance of I_i and I_{jk}
$\Delta_{ijkl} = \pi_{ijkl} - \pi_{ij} \pi_{kl}$	Covariance of I_{ij} and I_{kl}
p_k	Probability of drawing element k with replacement sampling
SI	Simple random sampling without replacement
WR	Random sampling with replacement
ST	Stratified sampling
$STSI$	Stratified sampling with SI in each stratum
$ST\pi ps$	Stratified sampling with πps in each stratum
SIC	simple random cluster sampling
PO	Poisson sampling
πps	Probability proportional to size sampling without replacement design
pps	Probability proportional to size sampling with replacement design
θ	Population ratio
$\hat{\theta}$	An estimator of θ
$\hat{\theta}_{GHR}$	Generalized Hartley and Ross estimator for θ
$E_p\left(\hat{\theta}\right)$	Expected value of an estimator of θ under sampling design p

$E_\xi\left(\hat{\theta}\right)$	Expected value of an estimator of θ under model design ξ
$var_p\left(\hat{\theta}\right)$	Variance of an estimator of θ under sampling design p
$var_\xi\left(\hat{\theta}\right)$	Variance of an estimator of θ under model design ξ
$\hat{v}ar_p\left(\hat{\theta}\right)$	Estimate of $var\left(\hat{\theta}\right)$
MSE	Mean square error
$\hat{t}_{y\pi} = \sum_{i\in s} \frac{y_i}{\pi_i}$	π estimator of the population total t_y
$\bar{r}_s = \frac{1}{N}\sum_{i\in s}\left(\frac{y_i}{x_i}\right)/\pi_i$	Sample mean of the ratios
D_t	Set of all distinct t_tuples $(i_1, i_2, \ldots, i_t)$
S_{D_t}	Set of all distinct t_tuples $(i_1, i_2, \ldots, i_t)$ from s
U_{D_t}	Set of all distinct t_tuples $(i_1, i_2, \ldots, i_t)$ from U

Bibliography

Ash, R. B. (2000). *Probability and Measure Theory* (2nd ed.). Massachusetts: Harcourt Academic Press.

Bickel, P. J. and D. A. Freedman (1984). Asymptotic normality and the bootstrap in stratified sampling. *Annals of Statistics 12*, 470–482.

Breidt, F. J. (2002). National resources inventory (NRI). *Encyclopedia of Environmetrics 3*, 135–1356.

Breidt, F. J. and J. D. Opsomer (2000). Local polynomial regression estimators in survey sampling. *Annals of Statistics 28*, 1026–1053.

Brewer, K. R. W. and M. Hanif (1983). *Sampling with Unequal Probabilities*. New York: Springer-Verlag.

Casella, G. and R. L. Berger (2002). *Statistical Inference* (2nd ed.). California: Duxbury.

Fuller, W. A. (1996). *Introduction to Statistical Time Series* (2 ed.). New York, NY: John Wiley & Sons.

Godambe, V. P. and V. M. Joshi (1965). Admissibility and bayes etimation in sampling finite population. *Annals of Mathematical Statistics 36*, 1707–1722.

Hájek, J. (1960). Limiting distributions in simple random sampling from finite population. *Publications of the Mathematical Institute of the Hungarian Academy 5*, 361–374.

Hansen, M. H. and W. N. Hurwitz (1943). On the theory of sampling from finite populations. *Annals of Mathematical Statistics 14*, 333–362.

Hanurav, T. V. (1967). Optimum utilization of auxiliary information: πps sampling of two units from a stratum. *Journal of the Royal Statistical Society 29*, 374–391–685.

Hartley, H. O. and A. Ross (1954). Unbiased ratio estimates. *Nature 174*, 270–271.

Horvitz, D. G. and D. J. Thompson (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association 47*, 663–685.

Isaki, C. and W. Fuller (1982). Survey design under the regression superpopulation model. *Journal of the American Statistical Association 77*, 89–96.

Krewski, D. and J. N. K. Rao (1981). Inference from stratified samples: Properties of the linearization, jacknife and balanced repeated replication methods. *Annals of Statistics 9*, 1010–1019.

Lahiri, D. B. (1951). A method for sample selection providing unbiased ratio estimates. *Bulletin of the International Statistical Institute 33*, 133–140.

Mickey, M. R. (1959). Some finite population unbiased ratio and regression estimators. *Journal of the American Statistical Association 54*, 594–612.

Mood, A. M., F. A. Graybill, and D. C. Boes (1974). *Introduction to the theory of statistics* (3rd ed.). New York: McGraw-Hill.

Murthy, M. and N. Nanjamma (1959). Almost unbiased ratio estimates based on interpenetrating subsample estimate. *Sankhyā 21*, 381–392.

Nieto de Pascual, J. (1961). Unbiased ratio estimators in stratified sampling. *Journal of the American Statistical Association 56*, 70–87.

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

Stuart, A. and J. K. Ord (1987). *Kendall's Advaced Theorey of Statistics* (5th ed.), Volume 1. New York: Oxford University Press.

Vijayan, K. (1968). An exact πps sampling scheme-generalization of a metod of hanurav. *Journal of the Royal Statistical Society 30*, 556–566.