Environ Ecol Stat (2011) 18:543–568
DOI 10.1007/s10651-010-0151-y

Leveraging the Rao–Blackwell theorem to improve ratio estimators in adaptive cluster sampling

Chang-Tai Chao · Arthur L. Dryver · Tzu-Ching Chiang

Received: 4 February 2010 / Revised: 22 June 2010 / Published online: 22 July 2010

Abstract Rao-Blackwellization is used to improve the unbiased Hansen–Hurwitz and Horvitz–Thompson unbiased estimators in Adaptive Cluster Sampling by finding the conditional expected value of the original unbiased estimators given the sufficient or minimal sufficient statistic. In principle, the same idea can be used to find better ratio estimators, however, the calculation of taking all the possible combinations into account can be extremely tedious in practice. The simplified analytical forms of such ratio estimators are not currently available. For practical interest, several improved ratio estimators in Adaptive Cluster Sampling are proposed in this article. The proposed ratio estimators are not the real Rao-Blackwellized versions of the original ones but make use of the Rao-Blackwellized univariate estimators. How to calculate the proposed estimators is illustrated, and their performance are evaluated by both of the Bivariate Poisson clustered process and a real data. The simulation result indicates that the proposed improved ratio estimators are able to provide considerably advantageous estimation results over the original ones.

Keywords Adaptive cluster sampling · Ratio estimator · Rao-Blackwellization · Sufficient statistic · Minimal sufficient statistic · Hansen–Hurwitz estimation · Horvitz–Thompson estimation

C.-T. Chao (✉) · T.-C. Chiang
Department of Statistics, School of Management, National Cheng-Kung University, Tainan 70101, Taiwan
e-mail: ctchao@stat.ncku.edu.tw

A. L. Dryver
Graduate School of Business Administration, National Institute of Development Administration, 118 Seri Thai Road, KlongChan, Bangkapi, Bangkok 10240, Thailand

1 Introduction

One of the longstanding issues in survey sampling is how to collect samples from a rare and cluster population in order to construct an inference regarding the population quantity of interest. The usual conventional sampling designs often fail to provide samples with meaningful information under such a population since it is highly possible that most of the sampled units would yield nothing. First proposed by Thompson (1990), Adaptive Cluster Sampling (ACS) is a data-driven sampling method which can provide higher sampling yields. In addition, ACS is often able to provide better inference than that provided by the comparable conventional sampling designs under rare and clustered populations. Various ACS designs associated with different initial designs have been developed in the past and the designs in the family of ACS have received considerable attention in different real-world survey situations because of their practical flexibility and ability to provide better estimation results. ACS designs have been widely used in different disciplines, such as ecological science (e.g. Acharya et al. 2000; Hanselman et al. 2003; Conners and Schwager 2002; Lo et al. 1997), environmental science (e.g. Vasudevan et al. 2001; Smith et al. 2003; Correll 2001), geographical science (Boomer et al. 2000), epidemiological study and social science (Thompson 1997; Thompson and Collins 2002).

Originally the estimators in ACS can be divided into two categories ; estimators of the modified Hansen and Hurwitz (1943) type and the modified Horvitz and Thompson Horvitz and Thompson (1952) type. Both of these estimators are design-unbiased; however still it is possible to further improve these estimators since they are not functions of the minimal sufficient statistic (Thompson and Seber 1996). In addition, the information obtained from the survey data is not fully utilized in the estimators described in Thompson (1990) since the data observed from the *edge units* which are not included in the initial sample are ignored at the inference stage. Thompson (1990) also described how to make use of the minimal sufficient statistic and Rao-Blackwellization (Rao 1945; Blackwell 1947) for better unbiased estimators. However, such procedure is not of practical interest because of the tedious calculation required. Salehi (1999) and Félix-Medina (2000) worked out the analytical form of the Rao-blackwellized unbiased estimators conditioning on the minimal sufficient statistic. For a simpler and more practical version, Dryver and Thompson (2005) proposed other Rao-Blackwellized unbiased estimators conditioning on a sufficient statistic.

All of the estimators described above are univariate estimators, which means that they use the information obtained from the population variable of primary interest, denoted as y, only. As in the usual survey sampling situations, other variables in addition to y often are available and naturally one would like to utilize these auxiliary information in the inference stage. Ratio estimation is a commonly used estimation method which makes use of the information obtained from a certain auxiliary variable, denoted as x, in order to take advantage of the correlation between y and x for better inference. Chao (2004) proposed the ratio estimator based on the original modified Horvitz–Thompson estimator and discussed the associated confidence intervals. Dryver and Chao (2007) proposed a ratio estimator based on the modified Hansen–Hurwitz estimator $\hat{\mu}_{r.hh}$ and modified Horvitz–Thompson estimator $\hat{\mu}_{r.ht}$, together with an unbiased ratio-type of estimator $\hat{\mu}_{r.hr}$ based on Hartley and Ross

(1954). While the unbiased $\hat{\mu}_{r.hr}$ is not recommended for its relatively high mean-squared error, the other two ratio estimators are considerably more efficient than other comparable estimators with the existence of certain correlation between x and y. Nevertheless, still the information of the edge units which are not initially selected is ignored. Hence, it is reasonable to assume that they can be improved if this ignored information can be utilized in the estimation. In this article, several improved ratio estimators which make use of the Rao-blackwellized univariate estimators described in Salehi (1999) and Dryver and Chao (2007) are proposed, and consequently the data in the edge units ignored before is utilized. In addition, the proposed improved ACS ratio estimators can be considered as valuable alternatives when ACS is the preferable sampling design but the ACS ratio estimators might not be appropriate (Dryver and Chao 2007). This article also investigates the appropriateness of the proposed improved ACS ratio estimators under such situations.

For readers that are not familiar with the general sampling procedure, terminologies, and associated estimation methods of ACS, Sect. 2 briefly describes ACS and reviews the typical unbiased estimators, Rao-Blackwellized unbiased estimators, and ratio estimators proposed in the past. The improved ratio estimators proposed in this article are described in Sect. 3, together with a numerical example in order to illustrate how they can be calculated. To evaluate the related performance of the proposed ratio estimators compared to the original ones, simulation results based on the pseudo populations generated by the Bivariate Poisson cluster process (Diggle 1983; Chao 2004) and a real population of Blue-winged teal data (Smith et al. 1995) are provided in Sect. 4. Section 5 gives a summary of this article.

2 Adaptive cluster sampling and the associated estimators

On contrast to the usual conventional sampling designs, the sampling procedure of ACS depends on the data obtained during the survey and consequently the sample cannot be determined before the survey. In general, an *initial sample* is selected by a conventional design, and the selection of the following sampling units depends on the data observed from the initial sample. The sampling selection goes on till certain condition of interest, which is a function of y and determined before the survey, can no longer be satisfied in the survey. Different ACS designs associated with different initial designs have been proposed in the past (e.g. Thompson 1991a,b; Chao 2003; Salehi and Seber 1997; Félix-Medina and Thompson 2004). The ACS design with simple random sampling without replacement (SRSWOR) is considered in this article. Nevertheless, the result can be extended straightforwardly to ACS with other initial sampling designs.

2.1 Sampling procedure and the associated technical notations

As the usual setup in a finite population sampling situation, the population size is considered as N and the units are labeled from $i = 1, \ldots, N$; hence the population is considered as a label set of $u = \{1, 2, \ldots, N\}$. The value of the population variable of interest associated with unit i is denoted as y_i. In ACS, a condition of interest C and the *neighborhood* of a unit are defined before the survey. The condition of interest C is a function of y, and usually C is an indicator function such as $C = I_{\{y_i > c\}}$, where c is a fixed constant. The neighborhood of i is a set of units such that if a unit is in the neighborhood of i, then i belongs to its neighborhood as well. The sampling begins with the selection of a set of initial sample s_0 by a conventional design. If any of the initially selected units satisfies C, the adaptive sampling procedure is performed to add new units according to the pre-defined neighborhood. If any of the newly added units satisfies C, then the same procedure goes on till no new unit satisfies C. The final sample s consists of the initial sample and all the units added afterwards (for the detailed sampling procedure of ACS, see Thompson 2002; Thompson and Seber 1996).

The set of all neighboring units that satisfy C is called a *network* and the adjacent units of a network that do not meet C are referred as *edge units*. Edge units and other units that do not meet the condition C can be considered as networks of size one. The collection of a network together with its associated edge units is called a *cluster*. The definition of a network assures that once a unit in the network is selected, all the other units in the same network will be included into the final sample (for an illustrative example of the network, cluster and edge unit, see Dryver and Chao 2007). The networks are disjoint sets determined by the definitions of C and the neighborhood. Hence the population can be partitioned into K networks, denoted as $\{\Psi_1, \ldots, \Psi_K\}$, according to the definitions of C and the neighborhood. Define m_k as the number of units belonging to the kth network Ψ_k, $\sum_{k=1}^{K} m_k = N$, and then the average y-value and x-value in the kth network can be defined as

$$w_{yk} = \frac{1}{m_k} \sum_{j \in \Psi_k} y_j$$

and

$$w_{xk} = \frac{1}{m_k} \sum_{j \in \Psi_k} x_j,$$

respectively.

2.2 Unbiased estimators under ACS

Thompson (1990) proposed two unbiased estimators for $\mu_y = \sum_{i=1}^{N} y_i / N$ in ACS, the modified Horvitz–Thompson estimator and the modified Hansen–Hurwitz estimator. The modified Horvitz–Thompson estimator is based on the initial inclusion probability of a network. The initial inclusion probability of kth network is

$$\alpha_k = 1 - \frac{\binom{N - m_k}{n_0}}{\binom{N}{n_0}}$$

under an initial SRSWOR design with an initial sample size n_0. Furthermore, let indicator variable I equal one if any unit of the kth network is in the initial sample, and zero otherwise. Also let n_k be the number of units selected from the kth network in the initial sample, $\sum_{k=1}^{K} n_k = n_0$.

The modified Horvitz–Thompson estimator (Thompson 1990)

$$\hat{\mu}_{y \cdot ht} = \frac{1}{N} \sum_{k=1}^{K} \frac{m_k \cdot w_{yk}}{\alpha_k} \cdot I\,(n_k > 0) \tag{1}$$

is an unbiased estimator for μ_y. Similarly we can define the modified Horvitz–Thompson estimator for $\mu_x = \sum_{i=1}^{N} x_i / N$, the population mean of x, as

$$\hat{\mu}_{x \cdot ht} = \frac{1}{N} \sum_{k=1}^{K} \frac{m_k \cdot w_{xk}}{\alpha_k} \cdot I\,(n_k > 0) \tag{2}$$

On the other hand, the modified Hansen–Hurwitz estimator is based on the draw-by-draw selection probabilities that the network a unit belongs to is intersected by the initial sample (Thompson 1990). In fact, ACS can be considered as an SRSWOR from a transformed population consisting of N units labeled $1, 2, \ldots, N$ with associated values of interest $w_y^* = \left\{w_{yj_1}, w_{yj_2}, \ldots, w_{yj_N};\ \{j_1, j_2, \ldots, j_N\} \in \{1, 2, \ldots, K\}\right\}$ rather than the original y-values $\{y_1, y_2, \ldots, y_N\}$ (Thompson 2002; Dryver and Chao 2007). Similarly, $w_x^* = \left\{w_{xj_1}, w_{xj_2}, \ldots, w_{xj_N};\ \{j_1, j_2, \ldots, j_N\} \in \{1, 2, \ldots, K\}\right\}$ can be considered as the transformed associated values of the auxiliary variable x instead of $\{x_1, x_2, \ldots, x_N\}$. Based on this point of view, the modified Hansen–Hurwitz estimator of μ_y and μ_x are

$$\hat{\mu}_{y \cdot hh} = \frac{1}{n_0} \sum_{k=1}^{K} n_k w_{y_k} = \frac{1}{n_0} \sum_{i \in s_0} w_{yj_i} \tag{3}$$

and

$$\hat{\mu}_{x \cdot hh} = \frac{1}{n_0} \sum_{k=1}^{K} n_k w_{x_k} = \frac{1}{n_0} \sum_{i \in s_0} w_{xj_i}, \tag{4}$$

respectively. In other words, the modified Hansen–Hurwitz estimator can be considered as the sample mean from an SRSWOR under the transformed population (Thompson 2002; Dryver and Chao 2007).

It can be shown that there is no complete sufficient statistic available under a finite fixed population sampling (Thompson and Seber 1996). However, the minimal sufficient statistic is available and it is the collection of an unordered set of distinct units s_R and the associated observations $\mathbf{y}_{s_R}$, denoted as $d_R = \{s_R, \mathbf{y}_{SR}\}$. Neither $\hat{\mu}_{y \cdot ht}$ nor $\hat{\mu}_{y \cdot hh}$ is the function of d_R; hence it is possible to further improve these estimators by utilizing the Rao-Blackwellization technique to find their conditional expected values given d_R. Salehi (1999) and Félix-Medina (2000) derived the analytical expressions of the Rao-Blackwellized versions of the modified Horvitz–Thompson and Hansen–Hurwitz estimators conditioning on d_R. In this article, we use Salehi's method to construct the new ratio estimators under ACS.

We shall first introduce some important results described in Salehi (1999). The final sample is partitioned into three disjoint sets of networks; F_1, F_2, and F_3. F_1 is the set of all networks consisting of two or more units. F_2 consists of all the edge units, which are networks of size one in the final sample, and F_3 is composed of all other networks. Let $|\cdot|$ be the cardinality of a set, and denote $\eta = |F_1|$, $\phi = |F_2|$, $\zeta = |F_3|$. Additionally, define ν as the number of distinct units in the final sample, and C_k as a set of initial combinations, subsets of d_R, which contains no units from network k associated with F_1. We can then calculate the number of combinations of the initial sample which gives rise to d_R in an inclusion-exclusion formula

$$\begin{aligned}\xi &= \binom{\nu'}{n_0'} - \sum_{k\in F_1} |C_k| + \sum_{k,l\in F_1} |C_k \cup C_l| + \cdots + (-1)^{\eta} \left|\bigcap_{k\in F_1} C_k\right| \\ &= \binom{\nu'}{n_0'} - \sum_{k\in F_1} \binom{\nu' - m_k}{n_0'} + \sum_{k,l\in F_1} \binom{\nu' - m_k - m_l}{n_0'} \\ &\quad + \cdots + (-1)^{\eta} \binom{\nu' - \sum_{k\in F_1} m_k}{n_0'},\end{aligned} \tag{5}$$

where $\nu' = \nu - \zeta$ and $n_0' = n_0 - \zeta$ since all networks associated with F_3 must be intersected by an initial sample. Thus, a formula for the Horvitz–Thompson type of Rao-Blackwellized estimator given the minimal sufficient statistic is (Salehi 1999)

$$\hat{\mu}_{y\cdot ht(ms)} = \frac{1}{N}\left(\sum_{k\in F_1} \frac{m_k \cdot w_{yk}}{\alpha_k} + \frac{\xi'}{\xi} \cdot \sum_{k\in F_2} \frac{w_{yk}}{\alpha_k} + \sum_{k\in F_3} \frac{w_{yk}}{\alpha_k}\right), \tag{6}$$

where

$$\begin{aligned}\xi' &= \binom{\nu' - 1}{n_0' - 1} - \sum_{k\in F_1} \binom{\nu' - m_k - 1}{n_0' - 1} + \sum_{k,l\in F_1} \binom{\nu' - m_k - m_l - 1}{n_0' - 1} \\ &\quad + \cdots + (-1)^{\eta} \binom{\nu' - \sum_{k\in F_1} m_k - 1}{n_0' - 1}\end{aligned} \tag{7}$$

is the number of initial combinations which contain at least one unit from network k associated with F_2 since the edge units are not necessarily included in the initial sample. In other words, ξ'/ξ can be considered as a selected probability of edge units in the initial sample. Compared with the $\hat{\mu}_{y\cdot ht}$ in Eq. (1), it can be seen that the information observed from the edge units is utilized via Rao-Blackwellization in $\hat{\mu}_{y\cdot ht(ms)}$.

Since the modified Hansen–Hurwitz estimator is fundamentally based on the value of w_y^*, Salehi (1999) defined F_{1u} as a set of unit labels rather than network labels.

Additionally, F_2 and F_3 also can be considered as sets of unit labels. Then, the Rao-Blackwellized version of the Hansen–Hurwitz estimator conditioning on the minimal sufficient statistic is (Salehi 1999)

$$\hat{\mu}_{y\cdot hh(ms)} = \frac{1}{n_0}\left(\sum_{i\in F_{1u}} \frac{\xi_i}{\xi}\cdot w_{yj_i} + \frac{\xi'}{\xi}\sum_{i\in F_2} w_{yj_i} + \sum_{i\in F_3} w_{yj_i}\right), \tag{8}$$

where

$$\xi_{ji} = \binom{\nu'-1}{n_0'-1} - \sum_{l\in F_1; l\neq k}\binom{\nu'-m_l-1}{n_0'-1} + \sum_{\{l,h\}\in F_1;\{l,h\}\neq k}\binom{\nu'-m_l-m_h-1}{n_0'-1}$$
$$+\cdots+(-1)^{\eta-1}\binom{\nu' - \sum_{l\in F_1; l\neq k} m_l - 1}{n_0'-1} \tag{9}$$

for unit i is in network $k\in F_1$. Similar to the ξ_1, ξ_i/ξ can be considered as the selection probability of units i according to its network label in F_1. However, since the units in F_3 must be selected in an initial sample, its selection probability is one.

Although the Rao-Blackwellized estimators described above are at least as efficient as the original unbiased estimators, they can be arduous to calculate. For the easy-to-compute estimators Dryver and Thompson (2005) proposed other improved estimators by conditioning $\hat{\mu}_{hh}$ and $\hat{\mu}_{ht}$ on a sufficient statistic instead of the minimal sufficient statistic. They divided the final sample into two parts: a core part s_c and the remaining part $s_{\bar{c}}$. s_c is the set of all the distinct units for which condition C is satisfied, and $s_{\bar{c}}$ consists of the distinct units for which condition C is not met. For unit i, let f_i be the number of times that the network to which unit i belongs is intersected by the initial sample. Then, the statistic d^+

$$d^+ = \left\{(i, y_i, f_i), (j, y_j); i\in s_c, j\in s_{\bar{c}}\right\}. \tag{10}$$

is a sufficient statistic (for a detailed proof, please see Dryver and Thompson 2005). A new unbiased estimator which is the conditional expected value of $\hat{\mu}_{ht}$ given d^+ is

$$\hat{\mu}_{y\cdot ht(s)} = \mathrm{E}\left(\hat{\mu}_{y\cdot ht}|D^+ = d^+\right) = \frac{1}{N}\sum_{k=1}^{K}\frac{m_k\cdot w^+_{yk}}{\alpha_k}\cdot I(n_k>0), \tag{11}$$

where

$$w^+_{yk} = \begin{cases} w^+_e = \frac{1}{\phi}\sum\limits_{l\in F_2} w_{yl}, & \text{if } k\in F_2 \\ w_{yk}, & \text{if } k\notin F_2 \end{cases}. \tag{12}$$

In addition, the Rao-Blackwellized modified Hansen–Hurwitz estimator is

$$\hat{\mu}_{y\cdot hh(s)} = \mathrm{E}\left(\hat{\mu}_{y\cdot hh}|D^{+} = d^{+}\right) = \frac{1}{n_0}\sum_{k=1}^{K} n_k w_{yk}^{+}. \tag{13}$$

2.3 Ratio estimators under ACS

Dryver and Chao (2007) proposed a modified ratio estimator for μ_y viewing ACS from the perspective of a transformed population together with the traditional ratio estimation (e.g. Lohr 1999). This modified ratio estimator makes use of the modified Hansen–Hurwitz estimator to estimate μ_y and μ_x,

$$\hat{\mu}_{r\cdot hh} = \hat{\beta}_{hh}\mu_x = \frac{\sum_{i\in s_0} w_{yj_i}}{\sum_{i\in s_0} w_{xj_i}}\mu_x = \frac{\hat{\mu}_{y\cdot hh}}{\hat{\mu}_{x\cdot hh}}\mu_x, \tag{14}$$

where $\hat{\beta}_{hh}$ is the sample ratio between w_{yj_i} and w_{xj_i} under the transformed population. However, the population ratio between w_{yj_i} and w_{xj_i} in the transformed population is the same as the population ratio between y_i and x_i in the original population,

$$\beta_{hh} = \frac{\sum_{i=1}^{N} w_{yj_i}}{\sum_{i=1}^{N} w_{xj_i}} = \frac{\sum_{i=1}^{N} y_i}{\sum_{i=1}^{N} x_i}.$$

Thus, $\hat{\mu}_{r\cdot hh}$ can be viewed as the traditional ratio estimator in the transformed population and is still design-biased.

Chao (2004) proposed a generalized ratio estimator (e.g. Cassel et al. 1976) by replacing the inclusion probability with the initial selection probability, and this ratio estimator can be considered as one based on the modified Horvitz–Thompson estimator,

$$\hat{\mu}_{r\cdot ht} = \frac{\hat{\mu}_{y\cdot ht}}{\hat{\mu}_{x\cdot ht}}\cdot\mu_x. \tag{15}$$

The ratio estimators proposed by Chao (2004) and Dryver and Chao (2007) are usually more efficient than the original univariate unbiased estimators if x and y are correlated to a certain degree. They are also better than the conventional ratio estimators under comparable sample sizes. However, still the information obtained from the edge units are ignored in $\hat{\mu}_{r\cdot hh}$ and $\hat{\mu}_{r\cdot ht}$, and it is expected that better ratio estimators can be constructed by utilizing the edge information.

3 Improved ratio estimators

One way to carry out the Rao-Blackwellized version of $\hat{\mu}_{r\cdot hh}$ and $\hat{\mu}_{r\cdot ht}$ is to use the procedure described in Thompson (1990). However, most likely such a procedure cannot be applied in practice. Further, since the ACS types of ratio estimator described in

Sect. 2.3 are ratios of the univariate estimators of ACS multiplied by μ_x, the analytical forms of their Rao-Blackwellized versions conditioning on d_R and d^+ are complicated. In fact, the Rao-Blackwellization described in Sect. 2.2 cannot be applied since $\hat{\mu}_{r\cdot ht}$ and $\hat{\mu}_{r\cdot hh}$ cannot be partitioned into separated parts. For example,

$$\begin{aligned}
\hat{\mu}_{r\cdot ht} &= \frac{\sum_{k=1}^{K} \frac{m_k \cdot w_{yk}}{\alpha_k} \cdot I(n_k > 0)}{\sum_{k=1}^{K} \frac{m_k \cdot w_{xk}}{\alpha_k} \cdot I(n_k > 0)} \cdot \mu_x \\
&= \frac{\sum_{k\in F_1} \frac{m_k \cdot w_{yk}}{\alpha_k} \cdot I(n_k > 0) + \sum_{k\in F_2} \frac{m_k \cdot w_{yk}}{\alpha_k} \cdot I(n_k > 0) + \sum_{k\in F_3} \frac{m_k \cdot w_{yk}}{\alpha_k} \cdot I(n_k > 0)}{\sum_{k\in F_1} \frac{m_k \cdot w_{xk}}{\alpha_k} \cdot I(n_k > 0) + \sum_{k\in F_2} \frac{m_k \cdot w_{xk}}{\alpha_k} \cdot I(n_k > 0) + \sum_{k\in F_3} \frac{m_k \cdot w_{xk}}{\alpha_k} \cdot I(n_k > 0)} \cdot \mu_x \\
&\neq \left[\sum_{k\in F_1} \frac{w_{yk}}{w_{xk}} \cdot I(n_k > 0) + \sum_{k\in F_2} \frac{w_{yk}}{w_{xk}} \cdot I(n_k > 0) + \sum_{k\in F_3} \frac{w_{yk}}{w_{xk}} \cdot I(n_k > 0) \right] \cdot \mu_x .
\end{aligned} \tag{16}$$

Therefore, in this article we proposed other ratio estimators in which the Rao-Blackwellization technique is utilized in a rather straightforward and simple manner.

Similar to w_{yk}^{+} in Sect. 2.2, we define

$$w_{xk}^{+} = \begin{cases} w_{e\cdot x}^{+} = \frac{1}{\phi} \sum\limits_{l\in F_2} w_{xl}, & \text{if } k \in F_2 \\ w_{xk}, & \text{if } k \notin F_2 \end{cases} .$$

and for the ratio estimator of Horvitz–Thompson type, $\hat{\mu}_{r.ht}$, we proposed the following improved ratio estimators making use of the Rao-Blackwellized univariate estimators:

$$\hat{\mu}_{r\cdot ht(ms)} = \frac{\hat{\mu}_{y\cdot ht(ms)}}{\hat{\mu}_{x\cdot ht(ms)}} \mu_x \tag{17}$$

and

$$\hat{\mu}_{r\cdot ht(s)} = \frac{\hat{\mu}_{y\cdot ht(s)}}{\hat{\mu}_{x\cdot ht(s)}} \mu_x , \tag{18}$$

which are μ_x multiplied by that ratios of the univariate estimators of the Horvitz–Thompson type conditioning on a minimal sufficient or sufficient statistics, respectively, where

$$\hat{\mu}_{x\cdot ht(ms)} = \frac{1}{N} \left(\sum_{k\in F_1} \frac{m_k \cdot w_{xk}}{\alpha_k} + \frac{\xi'}{\xi} \cdot \sum_{k\in F_2} \frac{w_{xk}}{\alpha_k} + \sum_{k\in F_3} \frac{w_{xk}}{\alpha_k} \right)$$

and

$$\hat{\mu}_{x\cdot ht(s)} = \frac{1}{N} \sum_{k=1}^{K} \frac{m_k \cdot w_{xk}^{+}}{\alpha_k} \cdot I(n_k > 0).$$

For the ratio estimator obtained from the modified Hansen–Hurwitz estimator, we proposed the following improved ratio estimators:

$$\hat{\mu}_{r\cdot hh(ms)} = \frac{\hat{\mu}_{y\cdot hh(ms)}}{\hat{\mu}_{x\cdot hh(ms)}}\mu_x \tag{19}$$

and

$$\hat{\mu}_{r\cdot hh(s)} = \frac{\hat{\mu}_{y\cdot hh(s)}}{\hat{\mu}_{x\cdot hh(s)}}\mu_x, \tag{20}$$

where

$$\hat{\mu}_{x\cdot hh(ms)} = \frac{1}{n_0}\left(\sum_{i\in F_{1u}} \frac{\xi_{ji}}{\xi}\cdot w_{xj_i} + \frac{\xi'}{\xi}\sum_{i\in F_2} w_{xj_i} + \sum_{i\in F_3} w_{xj_i}\right)$$

and

$$\hat{\mu}_{x\cdot hh(s)} = \frac{1}{N}\sum_{k=1}^{K}\frac{m_k\cdot w_{xk}^{+}}{\alpha_k}\cdot I(n_k > 0).$$

Similar to $\hat{\mu}_{r\cdot ht(ms)}$ and $\hat{\mu}_{r\cdot ht(s)}$, $\hat{\mu}_{r\cdot hh(ms)}$ and $\hat{\mu}_{r\cdot hh(s)}$ are the ratios of the Rao-Blackwellized Hansen–Hurwitz univariate estimators for μ_y and μ_x proposed in Salehi (1999) and Dryver and Thompson (2005), respectively.

3.1 An illustrative example

Table 1 includes a small population which is used to illustrate how $\hat{\mu}_{r\cdot hh(s)}$, $\hat{\mu}_{r\cdot ht(s)}$, $\hat{\mu}_{r\cdot hh(ms)}$, and $\hat{\mu}_{r\cdot ht(ms)}$ can be calculated.

The population means of the example for y and x are 187.4 and 8.8, respectively. In this example the condition to adaptively add units is $C = \{y; y \geq 100\}$. Pretend an initial sample is taken by simple random sampling of size $n_0 = 6$ and that the units 1, 4, 5, 10, 12, and 15 are selected from the $N = 15$ population units. The inclusion probabilities for networks of different sizes can be calculated by:

$$m_k = 1,\quad \alpha_k = 1 - \frac{\binom{14}{6}}{\binom{15}{6}} = 0.4$$

$$m_k = 2,\quad \alpha_k = 1 - \frac{\binom{13}{6}}{\binom{15}{6}} = 0.6571$$

Table 1 An illustrative population

Unit i	y_i	w_{y_i}	x_i	w_{x_i}	Network k	m_k	α_k
1	150	150	10	10	1	1	0.4
2	3	3	0	0	2	1	0.4
3	2	2	0	0	3	1	0.4
4	7	7	8	8	4	1	0.4
5	1,000	600	30	20	5	3	0.8154
6	600	600	20	20	5	3	0.8154
7	200	600	10	20	5	3	0.8154
8	4	4	0	0	6	1	0.4
9	120	160	5	7.5	7	2	0.6571
10	200	160	10	7.5	7	2	0.6571
11	6	6	3	3	8	1	0.4
12	9	9	0	0	9	1	0.4
13	10	10	6	6	10	1	0.4
14	300	250	20	15	11	2	0.6571
15	200	250	10	15	11	2	0.6571

$$m_k = 3, \quad \alpha_k = 1 - \frac{\binom{12}{6}}{\binom{15}{6}} = 0.8154$$

The final sample consists of 14 units and the sampled units are 1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, and 15.

The networks selected in the initial sample that are not adaptively added are $k = 1, 4, 5, 7, 9, 11$, so $n_k > 0$ for only these networks in the final sample.

$$\begin{aligned}
\hat{\mu}_{y\cdot ht} &= \frac{1}{N}\sum_{k=1}^{K}\frac{m_k \cdot w_{yk}}{\alpha_k} \cdot I(n_k > 0) \\
&= \frac{1}{15}\left(\frac{1\cdot 150}{0.4} + \frac{1\cdot 7}{0.4} + \frac{3\cdot 600}{0.8154} + \frac{2\cdot 160}{0.6571} + \frac{1\cdot 9}{0.4} + \frac{2\cdot 250}{0.6571}\right) = 258.0275 \\
\hat{\mu}_{x\cdot ht} &= \frac{1}{15}\left(\frac{1\cdot 10}{0.4} + \frac{1\cdot 8}{0.4} + \frac{3\cdot 20}{0.8154} + \frac{2\cdot 7.5}{0.6571} + \frac{1\cdot 0}{0.4} + \frac{2\cdot 15}{0.6571}\right) = 12.4711 \\
\hat{\mu}_{y\cdot hh} &= \frac{1}{n_0}\sum_{k=1}^{K} n_k w_{yk} = \frac{1}{n_0}\sum_{i\in s_0} w_{yj_i} = \frac{1}{6}(150 + 7 + 600 + 160 + 9 + 250) = 196 \\
\hat{\mu}_{x\cdot hh} &= \frac{1}{6}(10 + 8 + 20 + 7.5 + 0 + 15) = 10.0833
\end{aligned}$$

The minimal sufficient statistic is

$$\begin{aligned}
d_R &= \{(i;\, y_i, x_i);\ i \in s_R\} \\
&= \{(1;\, 150, 10), (2;\, 3, 0), (4;\, 7, 8), (5;\, 1000, 30), (6;\, 600, 20), (7;\, 200, 10), \\
&\quad (8;\, 4, 0), (9;\, 120, 5), (10;\, 200, 10), (11;\, 6, 3), (12;\, 9, 0), (13;\, 10, 6), \\
&\quad (14;\, 300, 20), (15;\, 200, 10)\}
\end{aligned}$$

and

$$\begin{aligned}
F_1 &= \{\text{network 5, network 7, network 11}\} \\
&= \{(5;\, 1000, 30), (6;\, 600, 20), (7;\, 200, 10)\} \cup \{(9;\, 120, 5), (10;\, 200, 10)\} \\
&\quad \cup \{(14;\, 300, 20), (15;\, 200, 10)\} \\
F_2 &= \{\text{network 2, network 4, network 6, network 8, network 10}\} \\
&= \{(2;\, 3, 0)\} \cup \{(4;\, 7, 8)\} \cup \{(8;\, 4, 0)\} \cup \{(11;\, 6, 3)\} \cup \{(13;\, 10, 6)\} \\
F_3 &= \{\text{network 1, network 9}\} = \{(1;\, 150, 10)\} \cup \{(12;\, 9, 0)\} \\
\eta &= 3, \phi = 5, \zeta = 2 \\
\nu' &= \nu - \zeta = 14 - 2 = 12;\ n_0' = n_0 - \zeta = 6 - 2 = 4 \\
\xi &= \binom{12}{4} - \binom{12-3}{4} - 2 \cdot \binom{12-2}{4} + 2 \cdot \binom{12-3-2}{4} \\
&\quad + \binom{12-2-2}{4} - \binom{12-3-2-2}{4} \\
&= 84 \\
\xi' &= \binom{11}{3} - \binom{11-3}{3} - 2 \cdot \binom{11-2}{3} + 2 \cdot \binom{11-3-2}{3} \\
&\quad + \binom{11-2-2}{3} - \binom{11-3-2-2}{3} \\
&= 12
\end{aligned}$$

For network $5 \in F_1$, we have

$$\xi_i = \binom{11}{3} - 2 \cdot \binom{11-2}{3} + \binom{11-2-2}{3} = 32.$$

and for network 7 or network $11 \in F_1$,

$$\xi_i = \binom{11}{3} - \binom{11-2}{3} - \binom{11-3}{3} + \binom{11-2-3}{3} = 45.$$

Consequently, the Rao-Blackwellized estimators of μ_y and μ_x conditioning on d_R are calculated as following

$$\begin{aligned}
\hat{\mu}_{y\cdot ht(ms)} &= \frac{1}{N}\left(\sum_{k\in F_1}\frac{m_k\cdot w_{yk}}{\alpha_k}+\frac{\xi'}{\xi}\cdot\sum_{k\in F_2}\frac{w_{yk}}{\alpha_k}+\sum_{k\in F_3}\frac{w_{yk}}{\alpha_k}\right)\\
&= \frac{1}{15}\left[\left(\frac{3\cdot 600}{0.8154}+\frac{2\cdot 160}{0.6571}+\frac{2\cdot 250}{0.6571}\right)+\frac{12}{84}\cdot\left(\frac{3+7+4+6+10}{0.4}\right)\right.\\
&\quad\left.+\left(\frac{150}{0.4}+\frac{9}{0.4}\right)\right]\\
&= 257.5752\\
\hat{\mu}_{x\cdot ht(ms)} &= \frac{1}{15}\left[\left(\frac{3\cdot 20}{0.8154}+\frac{2\cdot 7.5}{0.6571}+\frac{2\cdot 15}{0.6571}\right)+\frac{12}{84}\cdot\left(\frac{0+8+0+3+6}{0.4}\right)\right.\\
&\quad\left.+\left(\frac{10}{0.4}+\frac{0}{0.4}\right)\right]\\
&= 11.5425\\
\hat{\mu}_{y\cdot hh(ms)} &= \frac{1}{n_0}\left(\sum_{i\in F_{1u}}\frac{\xi_i}{\xi}\cdot w_{yj_i}+\frac{\xi'}{\xi}\sum_{i\in F_2}w_{yj_i}+\sum_{i\in F_3}w_{yj_i}\right)\\
&= \frac{1}{6}\left[\left(\frac{32}{84}\cdot 3\cdot 600+\frac{45}{84}\cdot 2\cdot 160+\frac{45}{84}\cdot 2\cdot 250\right)\right.\\
&\quad\left.+\frac{12}{84}\,(3+7+4+6+10)+(150+9)\right]\\
&= 214.7143\\
\hat{\mu}_{x\cdot hh(ms)} &= \frac{1}{6}\left[\left(\frac{32}{84}\cdot 3\cdot 20+\frac{45}{84}\cdot 2\cdot 7.5+\frac{45}{84}\cdot 2\cdot 15\right)\right.\\
&\quad\left.+\frac{12}{84}(0+8+0+3+6)+(10+0)\right]\\
&= 9.8988
\end{aligned}$$

The sufficient statistic d^+ is

$$\begin{aligned}
d^+ &= \{(i;\, y_i, x_i, f_i);\, i\in s_c\}\cup\left\{\left(j;\, y_j, x_j\right);\, s_{\bar{c}}\right\}\\
&= \{(1;\,150,10,1),(5;\,1000,30,1),(6;\,600,20,1),(7;\,200,10,1),(9;\,120,5,1),\\
&\quad (10;\,200,10,1),(14;\,300,20,1),(15;\,200,10,1)\}\cup\{(2;\,3,0),(4;\,7,8),\\
&\quad (8;\,4,0),(11;\,6,3),(12;\,9,0),(13;\,10,6)\}
\end{aligned}$$

The Rao-Blackwellized estimators of Horvitz–Thompson type for μ_y conditioning on d^+ can be calculated as

$$\hat{\mu}_{y\cdot ht(s)} = \mathrm{E}\left(\hat{\mu}_{y\cdot ht}|D^+=d^+\right)=\frac{1}{N}\sum_{k=1}^{K}\frac{m_k\cdot w^+_{yk}}{\alpha_k}\cdot I\,(n_k>0),$$

where

$$w_{yk}^{+} = \begin{cases} w_e^{+} = \frac{1}{\phi} \sum\limits_{l \in F_2} w_{yl}, & \text{if } k \in F_2 \\ w_{yk}, & \text{if } k \notin F_2 \end{cases}.$$

In fact, $w_e^{+} = (3+7+4+6+10)/5 = 6$ and $w_{e \cdot x}^{+} = (0+8+0+3+6)/5 = 3.4$ is the averages of y values.

$$\hat{\mu}_{y \cdot ht(s)} = \frac{1}{15}\left(\frac{1 \cdot 150}{0.4} + \frac{1 \cdot 6}{0.4} + \frac{3 \cdot 600}{0.8154} + \frac{2 \cdot 160}{0.6571} + \frac{1 \cdot 9}{0.4} + \frac{2 \cdot 250}{0.6571}\right) = 257.8609$$

Similarly, we have

$$\hat{\mu}_{x \cdot ht(s)} = \frac{1}{15}\left(\frac{1 \cdot 10}{0.4} + \frac{1 \cdot 3.4}{0.4} + \frac{3 \cdot 20}{0.8154} + \frac{2 \cdot 7.5}{0.6571} + \frac{1 \cdot 0}{0.4} + \frac{2 \cdot 15}{0.6571}\right) = 11.7044$$

The Rao-Blackwellized estimators of Hansen–Hurwitz type for μ_y conditioning on d^{+} is calculated based on

$$\hat{\mu}_{y \cdot hh(s)} = \frac{1}{n_0} \sum_{k=1}^{K} n_k w_{yk}^{+}.$$

hence

$$\hat{\mu}_{y \cdot hh(s)} = \frac{1}{6}(150 + 6 + 600 + 160 + 9 + 250) = 195.8333.$$

Similarly

$$\hat{\mu}_{x \cdot hh(s)} = \frac{1}{6}(10 + 3.4 + 20 + 7.5 + 0 + 15) = 9.3167$$

Hence, the improved ratio estimators described in the previous section are

$$\hat{\mu}_{r \cdot ht} = \frac{\hat{\mu}_{y \cdot ht}}{\hat{\mu}_{x \cdot ht}} \cdot \mu_x = \frac{258.0275}{12.4711} \cdot 8.8 = 182.0723$$
$$\hat{\mu}_{r \cdot ht(s)} = \frac{\hat{\mu}_{y \cdot ht(s)}}{\hat{\mu}_{x \cdot ht(s)}} \cdot \mu_x = \frac{257.8609}{11.7044} \cdot 8.8 = 193.8738$$
$$\hat{\mu}_{r \cdot ht(ms)} = \frac{\hat{\mu}_{y \cdot ht(ms)}}{\hat{\mu}_{x \cdot ht(ms)}} \cdot \mu_x = \frac{257.5752}{11.5425} \cdot 8.8 = 196.3753$$
$$\hat{\mu}_{r \cdot hh} = \frac{\hat{\mu}_{y \cdot hh}}{\hat{\mu}_{x \cdot hh}} \cdot \mu_x = \frac{196}{10.0833} \cdot 8.8 = 171.0551$$
$$\hat{\mu}_{r \cdot hh(s)} = \frac{\hat{\mu}_{y \cdot hh(s)}}{\hat{\mu}_{x \cdot hh(s)}} \cdot \mu_x = \frac{195.8333}{9.3167} \cdot 8.8 = 184.9725$$

$$\hat{\mu}_{r\cdot hh(ms)} = \frac{\hat{\mu}_{y\cdot hh(ms)}}{\hat{\mu}_{x\cdot hh(ms)}} \cdot \mu_x = \frac{214.7143}{9.8988} \cdot 8.8 = 190.8803$$

4 Performance evaluation

As the Rao-Blackwellization technique is used in all the improved ratio estimators proposed in Eqs. (17)–(20), still we cannot guarantee that they would be always at least as efficient as the corresponding original ratio estimators. That is, we cannot assure the following inequalities

$$\begin{aligned} \text{MSE}\left(\hat{\mu}_{r\cdot ht(ms)}\right) &\le \text{MSE}\left(\hat{\mu}_{r\cdot ht(s)}\right) \le \text{MSE}\left(\hat{\mu}_{r\cdot ht}\right) \\ \text{MSE}\left(\hat{\mu}_{r\cdot hh(ms)}\right) &\le \text{MSE}\left(\hat{\mu}_{r\cdot hh(s)}\right) \le \text{MSE}\left(\hat{\mu}_{r\cdot hh}\right) \end{aligned}$$

since $\hat{\mu}_{r\cdot ht(s)}$, $\hat{\mu}_{r\cdot ht(ms)}$, $\hat{\mu}_{r\cdot hh(s)}$ and $\hat{\mu}_{r\cdot hh(ms)}$ are ratios of Rao-Blackwellized univariate estimators, but not the real Rao-Blackwellized version of the ratio estimators. Nevertheless, we expect that they would be better in most situations by separately improving the numerator and denominator of the original versions.

4.1 Simulation study

Adaptive cluster sampling is known to be an appropriate alternative to collect sample than the comparable conventional sampling designs since it can provide more efficient estimates together with higher sampling yields, especially for rare and clustered populations. Figure 1 illustrates a clustered primary population y together with an auxiliary population x. The illustrative population in Fig. 1 is generated by a linked pairs process (Diggle 1983, pp. 93–94) together with the Bivariate Poisson cluster process (Chao 2004), which is a generalization of Poisson cluster process (Diggle 1983, pp. 55–57). The Bivariate Poisson clustered process used in this article is that for each parent, there are two different types of siblings generated. The number of siblings, denoted as Y and X, are jointly distributed as a bivariate Poisson distribution $Poi(\theta_1, \theta_2, \theta)$. Marginally, Y and X are distributed as a Poisson distribution with parameter $\theta_1 + \theta$ and $\theta_2 + \theta$. The covariance and correlation between Y and X are θ and $\theta/(\theta_1 + \theta_2)$, respectively. In our study, the value of θ is selected so that the correlation between the two types of siblings is 0.8. The number of parents is distributed as $Poi(\theta_0)$.

The effect of different mean values of parents (θ_0) and siblings (θ_1 and θ_2) on the performance of the proposed improved ratio estimators are compared to those proposed by Dryver and Chao (2007) with respect to different initial sample sizes n_0 in the following simulation study. For each case of different $(\theta_0, \theta_1\theta_2)$, 50,000 populations are generated and an initial sample of size n_0 is selected by SRSWOR for each population. Different ratio estimators $\hat{\mu}_{r\cdot hh}$, $\hat{\mu}_{r\cdot ht}$, $\hat{\mu}_{r\cdot hh(s)}$, $\hat{\mu}_{r\cdot ht(s)}$, $\hat{\mu}_{r\cdot hh(ms)}$, and $\hat{\mu}_{r\cdot ht(ms)}$ are calculated based on the final sample obtained from the same initial sample. The neighboring units are defined as the adjacent units on the left, right, top, and bottom, and the condition of interest is chosen to be $C = \{y;\ y \ge 5\}$. In addition, the conventional ratio estimator $\hat{\mu}_r$ is calculated based an SRSWOR sample of size equal to the

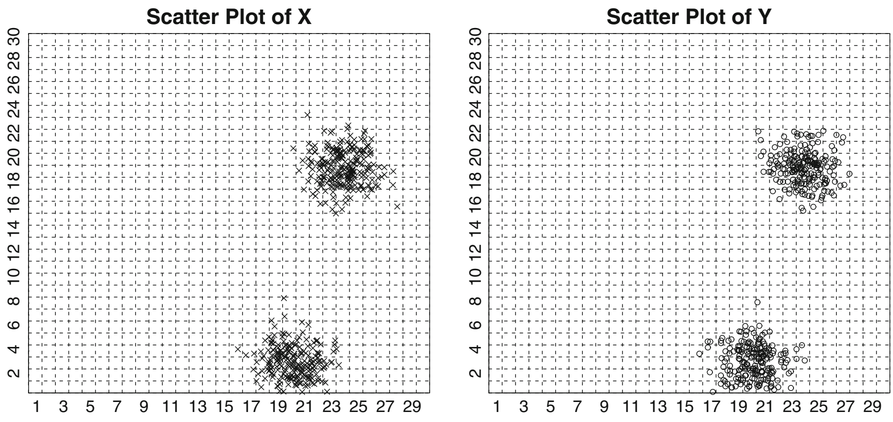
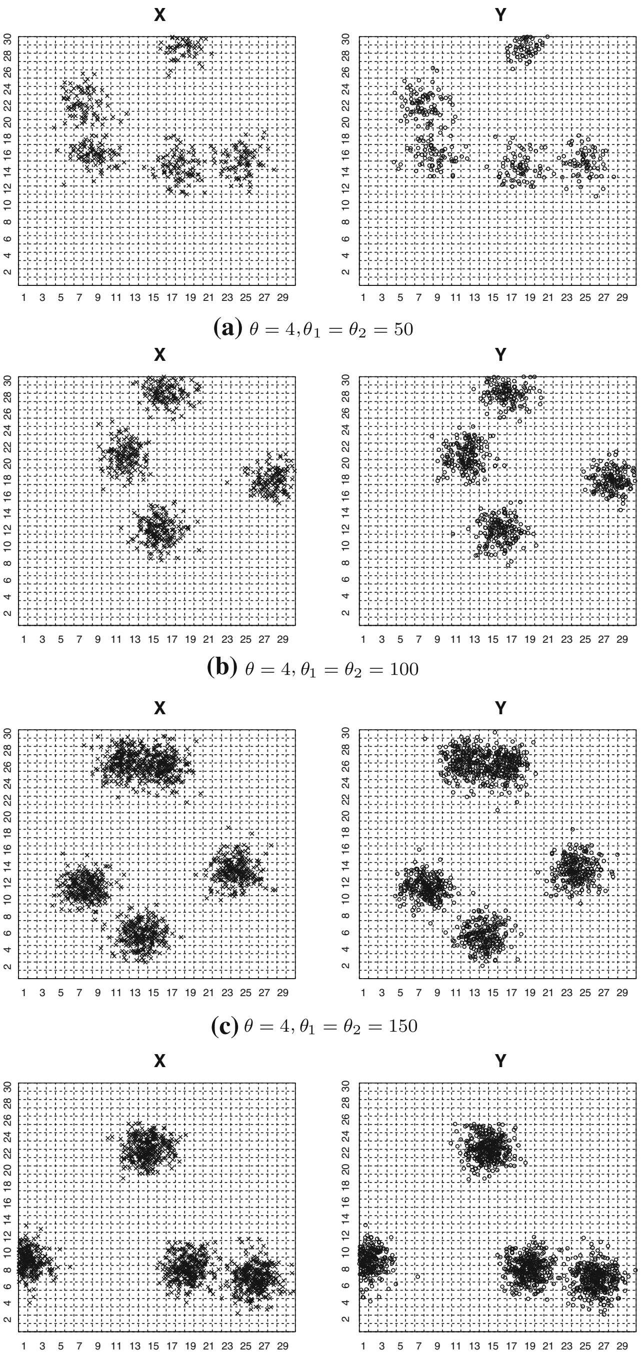


Fig. 1 The plots of y and x in a region partitioned into 900 square population units

final sample size n. $\hat{\mu}_r$ is used as a reference line to evaluate the performance of other estimators under ACS. The evaluation is based on the empirical relative efficiency of each $\hat{\mu}_{r\ldots}$ to $\hat{\mu}_r$ under the comparable final sample size n.

The empirical relative efficiency (E.R.E.) of $\hat{\mu}_{r\ldots}$ to $\hat{\mu}_r$ is defined as the ratio of the empirical mean-squared error of $\hat{\mu}_r$ to $\hat{\mu}_{r\ldots}$. Therefore, an E.R.E. of value greater than one indicates that the mean-squared error of $\hat{\mu}_{r\ldots}$ is smaller than $\hat{\mu}_r$ under comparable sample size. The greater the E.R.E. is, the better the performance of $\hat{\mu}_{r\ldots}$ is. The empirical mean-squared error is calculated by taking average of the squared error obtained from the estimates and the true value of μ under each of the 50,000 populations.

In the first part of the simulation study, the average number of parents θ_0 is fixed as four and the average number of offsprings is chosen to be 50,100,150, and 200, so that we can examine the effect of the number of offsprings on the performance of the proposed estimators. Figure 2 illustrates the examples of the population realizations generated when $\theta_0 = 4$ and $\theta_1 = \theta_2 = 50$, 100, 150, and 200. The population size is 900, and the initial sample size n_0 varies from 5 to 200. The simulation results are summarized in Fig. 3 and it shows that the six ratio estimators of A.C.S are usually better than $\hat{\mu}_r$ of SRSWOR under the comparable sample size. In addition, the improved ratio estimators $\hat{\mu}_{r\cdot hh(\cdot)}$ and $\hat{\mu}_{r\cdot ht(\cdot)}$ proposed in this article are considerably better than their original versions proposed in Dryver and Chao (2007) as the E.R.E can be further improved. While making use of sufficient statistic d^+, the improved ratio estimator based on Horvitz–Thompson estimation $\left(\hat{\mu}_{rht(s)}, \bullet\right)$ is always better than that based on Hansen–Hurwitz estimation $\left(\hat{\mu}_{rhh(s)}, \circ\right)$. The performance of the improved ratio estimator based on the minimal sufficient statistic d_R, $\hat{\mu}_{r.ht(ms)}$ ($\blacktriangle$) and $\hat{\mu}_{r.hh(ms)}(\triangle)$ are similar to each other, and both of them are better than $\hat{\mu}_{r\cdot\cdot(s)}$ as one would expect. Notice that the y-axes (E.R.E) of all the graphics in Fig. 3 are set to be the same for the purpose of a fair visual comparison, hence the simulation result when $\theta_1 = \theta_2 = 50$ is not as clear as others. The pattern, however, is the same as described above. Also notice that, as the number of parents is fixed as four, the sampling situation is more appropriate for ACS as the average number of offsprings increases, since accordingly the population is relatively more clustered.



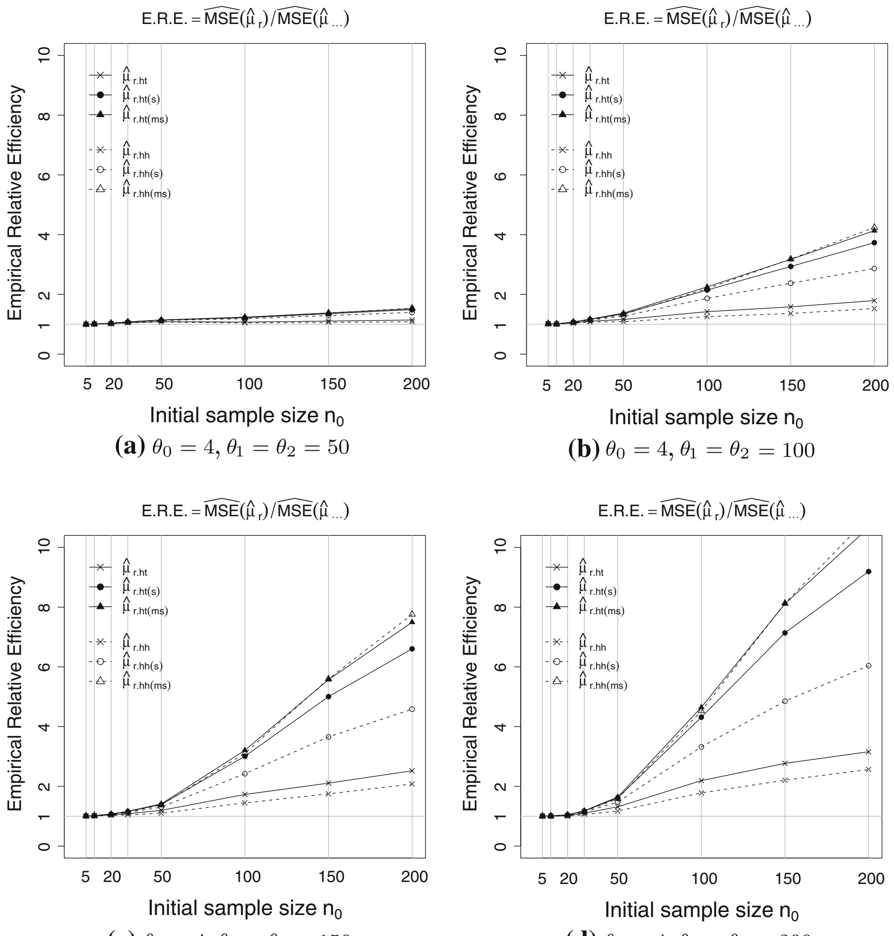
(a) $\theta = 4, \theta_1 = \theta_2 = 50$

(b) $\theta = 4, \theta_1 = \theta_2 = 100$

(c) $\theta = 4, \theta_1 = \theta_2 = 150$

(d) $\theta = 4, \theta_1 = \theta_2 = 200$

Fig. 2 Illustrative clustered populations, $\theta_0 = 4$

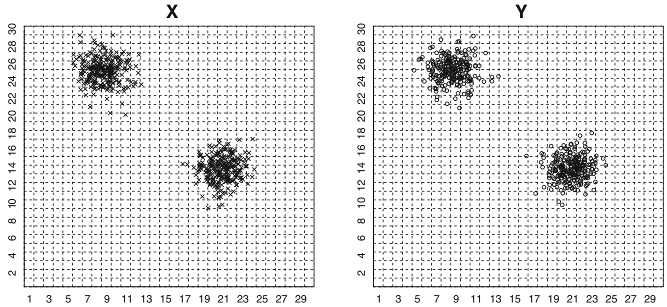


(a) $\theta_0 = 4, \theta_1 = \theta_2 = 50$
(b) $\theta_0 = 4, \theta_1 = \theta_2 = 100$
(c) $\theta_0 = 4, \theta_1 = \theta_2 = 150$
(d) $\theta_0 = 4, \theta_1 = \theta_2 = 200$

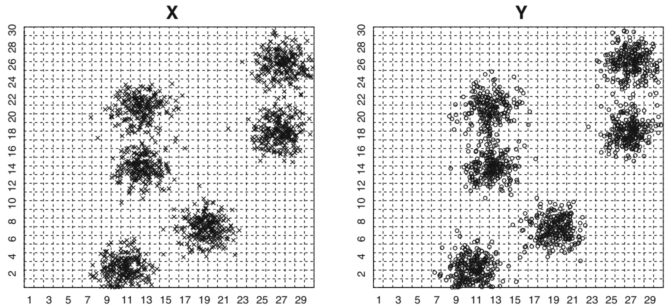
Fig. 3 Empirical relative efficiency of ratio estimators under adaptive cluster sampling to the conventional ratio estimator under SRSWOR with comparable sample size. $\theta_0 = 4, \theta_1 = \theta_2 = 50, 100, 150, 200$

In the second part of the simulation study, the average offspring number is fixed as $\theta_1 = \theta_2 = 150$, and the average numbers of parents are chosen to be 2, 6, 8, 10. Figure 4 illustrates the examples of realizations generated when $\theta_1 = \theta_2 = 150$ and $\theta_0 = 2$, 6, 8, and 10.

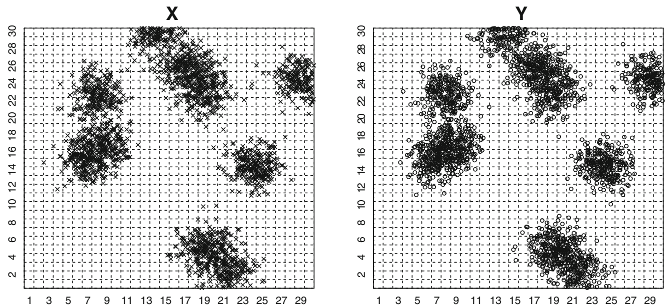
Despite the similar pattern as seen in Fig. 3 among the unimproved/improved ratio estimators under ACS, still there are other phenomenons different from the results in Fig. 3. In fact, when θ_0 is large, for instance $\theta_0 = 8$ or 10, clearly it is not be as advantageous to apply ACS as in other cases, such as when $\theta_0 = 2$ or 4. The reason is that the populations can no longer be be regarded a rare one as Fig. 4 shows. For such populations, $\hat{\mu}_{r.hh}$ and $\hat{\mu}_{r.ht}$ perform worse than the conventional $\hat{\mu}_r$ when the sample size is not large enough. For example they can be better than $\hat{\mu}_r$ only when the initial sample size is greater than 100 or more (Fig. 5c, d). The improved ratio estimators



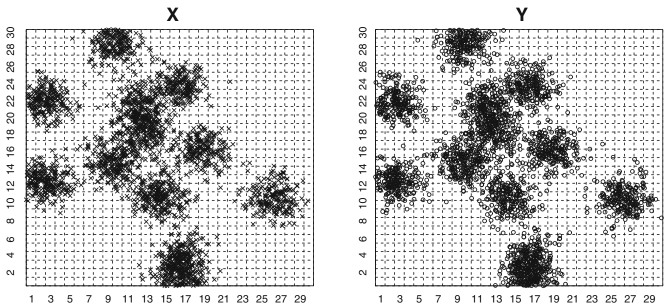
(a) $\theta_0 = 2,\ \theta_1 = \theta_2 = 150$



(b) $\theta_0 = 6,\ \theta_1 = \theta_2 = 150$

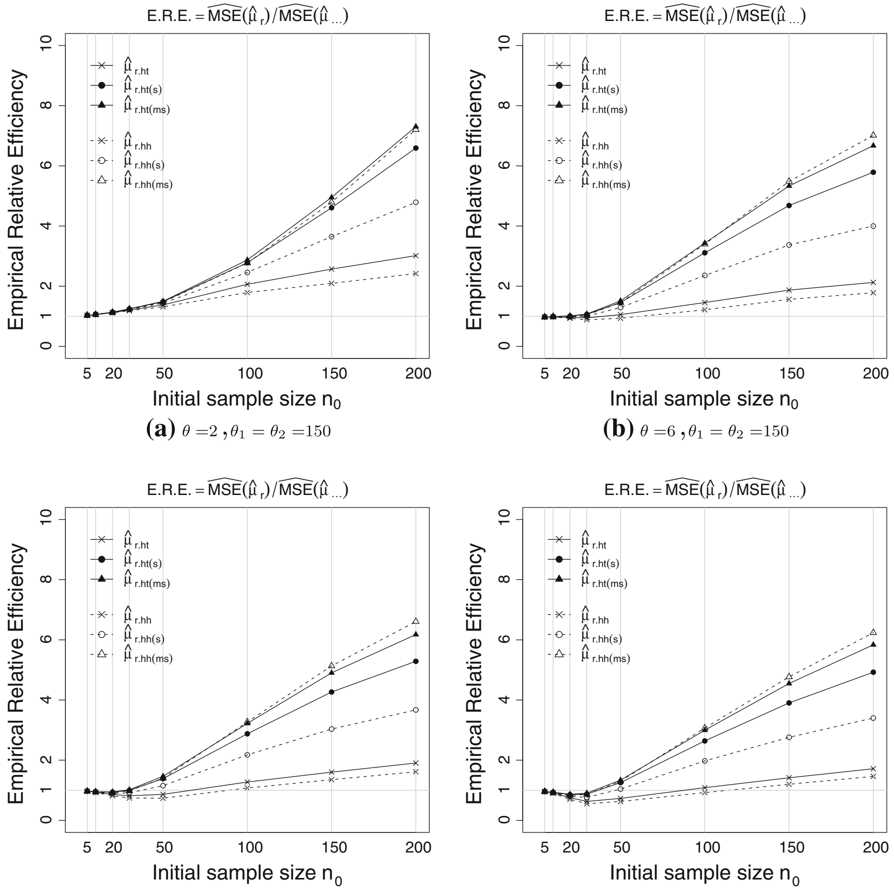


(c) $\theta_0 = 8,\ \theta_1 = \theta_2 = 150$



(d) $\theta_0 = 10,\ \theta_1 = \theta_2 = 150$

Fig. 4 Illustrative clustered populations, $\theta_1 = \theta_2 = 150$



(a) $\theta = 2, \theta_1 = \theta_2 = 150$ **(b)** $\theta = 6, \theta_1 = \theta_2 = 150$

(c) $\theta = 8, \theta_1 = \theta_2 = 150$ **(d)** $\theta = 10, \theta_1 = \theta_2 = 150$

Fig. 5 Empirical relative efficiency of ratio estimators under adaptive cluster sampling to the conventional ratio estimator under SRSWOR with comparable sample size. $\theta_0 = 2, 6, 8, 10, \theta_1 = \theta_2 = 150$

under ACS proposed in this article also can not always be superior to $\hat{\mu}_r$. However, they do not need sample size as large as that required by $\hat{\mu}_{r.hh}$ and $\hat{\mu}_{r.ht}$ to provide better estimation results than $\hat{\mu}_r$. As Fig. 5c, d) indicate, the improved ratio estimators perform better than the conventional $\hat{\mu}_r$ when the initial sample size is greater than 30 or more.

4.2 Blue-winged teal data

Dryver and Chao (2007) used the blue-winged teal data (Smith et al. 1995, Table 2) to examine the performance of $\hat{\mu}_{r.hh}$ and $\hat{\mu}_{r.ht}$. The original data are used as x_i and w_{x_i} in order to generate the population of interest y_i by two different models:

Table 2 Blue-winged teal data (Smith et al. 1995)

0	0	3	5	0	0	0	0	0	0
0	0	0	24	14	0	0	10	103	0
0	0	0	0	2	3	2	0	13, 639	1
0	0	0	0	0	0	0	0	14	122
0	0	0	0	0	0	2	0	0	177

Table 3 E.R.E to sample mean with respect to comparable sample sizes under model 1

n_0	$\hat{\mu}_r$	$\hat{\mu}_{r \cdot hh}$	$\hat{\mu}_{r \cdot hh(s)}$	$\hat{\mu}_{r \cdot hh(ms)}$	$\hat{\mu}_{r \cdot ht}$	$\hat{\mu}_{r \cdot ht(s)}$	$\hat{\mu}_{r \cdot ht(ms)}$
2	3.2920	3.3230	3.3230	3.3230	3.3231	3.3231	3.3231
5	8.5475	8.5792	8.5792	8.5794	8.5795	8.5795	8.5795
8	21.4089	21.3145	21.3145	21.3184	21.3192	21.3192	21.3192
11	52.6807	51.8981	51.8981	51.9189	51.9198	51.9198	51.9198
14	205.1740	201.4697	201.4697	201.7672	201.7563	201.7563	201.7563
17	800.2225	795.6666	795.6666	799.3839	798.7072	798.7072	798.7072
20	2,875.6417	3,602.1326	3,602.1326	3,628.2504	3,620.9540	3,620.9540	3,620.9540

Model 1

$$y_i = 4x_i + \epsilon_i,$$

where $\epsilon_i \sim N(0, x_i)$, and

Model 2

$$y_i = 4w_{xi} + \epsilon_i,$$

where $\epsilon_i \sim N(0, w_{xi})$. The condition of interest is $C = \{y; y > 0\}$. We use the same models to simulate y together with the the same condition of interest to evaluate of the improved ratio estimators.

Notice that y would be zero if x is zero under Model 1 and 2 used in Dryver and Chao (2007). Such models together with the condition of $C = \{y; y > 0\}$ lead both of y and x to be zero for all edge units in the simulated populations. However, it can be shown that $\hat{\mu}_{ht} = \hat{\mu}_{ht(s)} = \hat{\mu}_{ht(ms)}$ and $\hat{\mu}_{hh} = \hat{\mu}_{hh(s)}$ when all the edge units have zero y and x values. Hence, $\hat{\mu}_{r.ht} = \hat{\mu}_{r.ht(s)} = \hat{\mu}_{r.ht(ms)}$ and $\hat{\mu}_{r.hh} = \hat{\mu}_{r.hh(s)}$, and in practice only $\hat{\mu}_{r.hh(ms)}$ needs to be calculated to improve the ratio estimation in ACS under such a situation.

Tables 3 and 4 are the empirical relative efficiencies of all conventional and ACS ratio estimators considered in this article to the sample mean $\bar{y}$ under comparable final sample sizes with respect to different initial sample sizes, hence the larger value indicates less mean-squared error than $\bar{y}$. The results in these two tables show that $\hat{\mu}_{r.ht} = \hat{\mu}_{r.ht(s)} = \hat{\mu}_{r.ht(ms)}$ and $\hat{\mu}_{r.hh} = \hat{\mu}_{r.hh(s)}$ as expected. However, $\hat{\mu}_{r.hh(ms)}$ does provide a slightly better result than $\hat{\mu}_{r.hh}$ under both models.

Table 4 E.R.E. to sample mean with respect to comparable sample sizes under model 2

n_0	$\hat{\mu}_r$	$\hat{\mu}_{r\cdot hh}$	$\hat{\mu}_{r\cdot hh(s)}$	$\hat{\mu}_{r\cdot hh(ms)}$	$\hat{\mu}_{r\cdot ht}$	$\hat{\mu}_{r\cdot ht(s)}$	$\hat{\mu}_{r\cdot ht(ms)}$
2	0.9542×10^{-4}	1.3132	1.3132	1.3132	1.3132	1.3132	1.3132
5	0.8719×10^{-4}	2.0033	2.0033	2.0034	2.0034	2.0034	2.0034
8	1.1753×10^{-4}	3.6307	3.6307	3.6313	3.6314	3.6314	3.6314
11	1.5347×10^{-4}	8.3010	8.3010	8.3049	8.3054	8.3054	8.3054
14	1.9193×10^{-4}	22.3088	22.3088	22.3322	22.3302	22.3302	22.3302
17	2.0901×10^{-4}	72.2126	72.2126	72.4671	72.4500	72.4500	72.4500
20	2.2147×10^{-4}	595.2213	595.2213	602.4089	602.4088	602.4088	602.4088

Table 5 E.R.E to sample mean with respect to comparable sample sizes under model 3

n_0	n	$\hat{\mu}_r$	$\hat{\mu}_{r\cdot hh}$	$\hat{\mu}_{r\cdot hh(s)}$	$\hat{\mu}_{r\cdot hh(ms)}$	$\hat{\mu}_{r\cdot ht}$	$\hat{\mu}_{r\cdot ht(s)}$	$\hat{\mu}_{r\cdot ht(ms)}$
2	9.3	3.2799	3.3853	3.3857	3.3859	3.3857	3.3861	3.3861
5	19.1	7.2950	7.3082	7.3219	7.3673	7.3588	7.3702	7.3726
8	25.6	12.6336	10.4838	10.5181	10.8206	10.8086	10.8360	10.8407
11	30.1	20.8829	14.2995	14.3628	15.3210	15.2856	15.3405	15.3516
14	33.2	32.8508	20.6350	20.7029	22.9652	22.7607	22.8159	22.8497
17	35.6	100.1819	37.4424	37.7197	43.2162	42.2963	42.5732	42.6161
20	37.4	180.6239	78.4095	79.3721	96.2167	92.0900	93.1439	93.2702

For avoiding the situation that both of the y and x values of the edge units' are zero in Model 1 and 2, we use the same data and slightly modified models to generate the population variable of interest y. These models are

Model 3

$$y_i = \begin{cases} 4x_i + \epsilon_i & \text{if } x_i > 0 \\ y_i \sim Poi(1) & \text{Otherwise} \end{cases}$$

where $\epsilon_i \sim N(0, x_i)$, and

Model 4

$$y_i = \begin{cases} 4w_{xi} + \epsilon_i & \text{if } x_i > 0 \\ y_i \sim Poi(1) & \text{Otherwise} \end{cases}$$

where $\epsilon_i \sim N(0, w_{xi})$. We choose $C = \{y; y \geq 5\}$ so that the simulation situation would be similar to that in Sect. 4.1. The results are summarized in Tables 5 and 6. The final sample sizes with respect to different initial sample sizes in Model 1 and 2 are the same as what given in Dryver and Chao (2007), but they are slightly different in Model 3 and 4. The final sample sizes are given in Tables 5 and 6, and none of the final sample sizes is out of control under the new models. The similar pattern as in Sect. 4.1 among the ratio estimators in ACS also appears in Tables 5 and 6.

Table 6 E.R.E to sample mean with respect to comparable sample sizes under model 4

n_0	n	$\hat{\mu}_r$	$\hat{\mu}_{r\cdot hh}$	$\hat{\mu}_{r\cdot hh(s)}$	$\hat{\mu}_{r\cdot hh(ms)}$	$\hat{\mu}_{r\cdot ht}$	$\hat{\mu}_{r\cdot ht(s)}$	$\hat{\mu}_{r\cdot ht(ms)}$
2	9.6	0.9815×10^{-4}	1.3023	1.3025	1.3026	1.3024	1.3026	1.3027
5	19.7	0.8624×10^{-4}	1.8579	1.8606	1.8726	1.8716	1.8738	1.8743
8	26.1	1.0617×10^{-4}	2.5863	2.5951	2.6999	2.6998	2.7057	2.7071
11	30.5	1.4318×10^{-4}	3.7010	3.7282	4.1962	4.1582	4.1803	4.1906
14	33.5	2.0015×10^{-4}	5.9476	5.9847	7.1707	7.0048	7.0458	7.0572
17	35.8	2.1271×10^{-4}	11.6287	11.7063	15.0994	14.5517	14.6236	14.6383
20	37.6	2.2055×10^{-4}	34.8380	35.2818	52.0117	45.8888	46.7231	46.8794

5 Discussion

Rao-Blackwellization is a well-known method for constructing better estimators for population quantity of interest. In ACS different initial samples can yield the same final sample, but do not always yield the same estimate for the estimators $\hat{\mu}_{hh}$, $\hat{\mu}_{ht}$, $\hat{\mu}_{r\cdot hh}$, and $\hat{\mu}_{r\cdot ht}$. Conditioning on the minimal sufficient statistic this source of sampling variability from ACS is eliminated. Conditioning on a sufficient statistic that is not the minimal sufficient statistic can eliminate some to most of this source of sampling variability. Simplified analytical forms of Rao-Blackwellized estimators based on a sufficient statistic d^+ and the minimal sufficient statistic d_R are described in Dryver and Thompson (2005) and Salehi (1999) respectively for the univariate case.

In principle, the Rao-Blackwellized ratio estimation under ACS can be calculated by taking all the possible combinations into account (Thompson 2002). The calculation can be, however, extremely arduous. The analytical forms of the ratio estimators under ACS are currently unavailable. Instead of deriving the simple analytical forms of the Rao-Blaskwellized ratio estimators, which is probably an impossible task, in this article we propose various improved ratio estimators by intuitively taking ratio of the Rao-Blackwellized estimators proposed by Dryver and Thompson (2005) or Salehi (1999). The ratio of the Rao-Blackwellized estimators given the minimal sufficient statistic yields the same estimate for the same final sample as the Rao-Blackwellized ratio estimators conditioning on the minimal sufficient statistic would. In this sense although the approaches are different they conceptually achieve the same goal of eliminating the sampling variability arising from different initial samples that yield the same final sample. For this reason we believe the improved ratio estimators $\hat{\mu}_{r\cdot hh(ms)}$ and $\hat{\mu}_{r\cdot hh(ms)}$ proposed in this article capture the majority of the benefits, if not all, that the real Rao-Blackwellized ratio estimators based on the minimal sufficient statistic would produce. On the other hand, Rao-Blackwellized estimators conditioning on d^+ described by Dryver and Thompson (2005) also can be used to construct better ratio estimators by improving the denominator and numerator of the original ACS ratio estimators separately as $\hat{\mu}_{r\cdot hh(s)}$ and $\hat{\mu}_{r\cdot ht(s)}$ described in this article. In fact, it can be concluded that the proposed ratio estimators, though not the real Rao-Blackwellized ones, are preferable alternatives compared to the original versions from the simulation results. The simulation results described in Sect. 4 indicate that the proposed ratio estimators are always better than the ACS ratio estimators proposed in Dryver and Chao (2007).

The ACS ratio estimators perform better than the typical ratio estimator when there is a high correlation on a network level and a low correlation on the unit level (Dryver and Chao 2007). The same is true for the improved ACS ratio estimators according to the simulation results in Sect. 4.2. The simulation study using Model 3 illustrates that ACS might not be an appropriate sampling design when ratio estimation is possible even if the population is rare and clustered as the E.R.E's of the ratio estimators of ACS are in general worse than the conventional $\hat{\mu}_r$. Regardless, the proposed improved ratio estimators are better than their associated original versions. This indicates that the improved ratio estimators proposed in this article can be an appropriate alternative estimation tool under such situations. The ACS ratio estimators for Model 4 are always better than $\hat{\mu}_r$, and the improved ratio estimators are better than their original versions. Also, a similar pattern among these proposed ratio estimators as in Sect. 4.1 also appears in Tables 6. In addition, When the the edge units are equal to zero then the improved Horvitz–Thompson type ACS ratio estimators $\hat{\mu}_{r.ht(.)}$ equals the Horvitz–Thompson type ACS estimator $\hat{\mu}_{r.ht}$ as the Horvitz–Thompson type univariate estimators are equal as well. Thus given that the improved ACS ratio estimators perform at least as well as the ACS ratio estimators in terms of efficiency, it can be concluded that the proposed improved ratio estimators are preferable alternatives to the ACS ratio estimators.

On the other hand, The ACS designs can provide better estimation results than the comparable conventional designs when the population is a clustered and rare one. Otherwise, the estimation provided by the unimproved ACS estimator can be worse, as the illustrative populations in Fig. 4c, d and the associated E.R.E.'s in Fig. 5c, d demonstrate. However, often such population information might not be clear before the survey. Nevertheless, other than the statistical perspective the field investigators might want to use ACS for other practical advantages, such as higher sampling yield, even under certain risk of less efficient estimates. The proposed improved ratio estimators can also provide better chance for the investigators to protect themselves from such risk when the population is in fact not a rare one.

Acknowledgments Support for this research is provided by the National Science Council, Taiwan, NSC 94-2118-M-030-002-. The authors would like to express their sincere appreciation to the anonymous referees for the valuable comments.

References

Acharya B, Bhattarai G, de Gier A, Stein A (2000) Systematic adaptive cluster sampling for the assessment of rare tree species in Nepal. For Ecol Manage 137(1–3):65–73

Blackwell D (1947) Conditional expectation and unbiased sequential estimation. Ann Math Stat 18:105–110

Boomer K, Werner C, Brantley S (2000) CO_2 emissions related to the yellowstone volcanic system 1. Developing a stratified adaptive cluster sampling plan. J Geophys Res Solid Earth 105(B5):10817–10830

Cassel CM, Särndal CE, Wretman JH (1976) Some results on generalized difference estimation and generalized regression estimation for finite population. Biometrics 63:615–620

Chao CT (2003) Adaptive cluster sampling on stratified populations. J Chin Stat Assoc 41:141–180

Chao CT (2004) Ratio estimation on adaptive cluster sampling. J Chin Stat Assoc 42(3):307–327
Conners ME, Schwager SJ (2002) The use of adaptive cluster sampling for hydroacoustic surveys. Ices J Marine Sci 59(6):1314–1325
Correll RL (2001) The use of composite sampling in contaminated sites- a case study. Environ Ecol Stat 8(3):185–200
Diggle PJ (1983) Statistical analysis of spatial point patterns. Academic Press, London
Dryver AL, Chao CT (2007) Ratio estimators in adaptive cluster sampling. Environmetrics 18:607–620
Dryver AL, Thompson SK (2005) Improved unbiased estimators in adaptive cluster sampling. J R Stat Soc B67(1):157–166
Félix-Medina MH (2000) Analytical expressions for Rao–Blackwell estimators in adaptive cluster sampling. J Stat Plan Inference 84:221–236
Félix-Medina MH, Thompson SK (2004) Adaptive cluster double sampling. Biometrika 91:877–891
Hansen MM, Hurwitz WN (1943) On the theory of sampling from finite population. Ann Math Stat 14:333–362
Hanselman DH, Quinn TJ, Lunsford D, Heifetz J, Clausen D (2003) Applications in adaptive cluster sampling of Gulf of Alaska rockfish. Fish Bull 101(3):501–513
Hartley HO, Ross A (1954) Unbiased ratio estimators. Nature 174:270–271
Horvitz DG, Thompson DJ (1952) A generalization of sampling without replacement from a finite universe. J Am Stat Assoc 47:663–685
Lo NCH, Griffith D, Hunter JR (1997) Using a restricted adaptive cluster sampling to estimate Pacific hake larval abundance. Calif Coop Ocean Fish Invest Rep 38:103–113
Lohr SL (1999) Sampling: design and analysis. Duxbury Press, Pacific Grove
Rao CR (1945) Information and accuracy attainable in estimation of statistical parameters. Bull Calcutta Math Soc 37:81–91
Salehi MM (1999) Rao–Blackwell versions of the Horvitz–Thompson and Hansen–Hurwitz estimators in adaptive cluster sampling. J Environ Ecol Stat 6:183–195
Salehi M, Seber GAF (1997) Two-stage adaptive cluster sampling. Biometrics 53(3):959–970
Smith DR, Conroy MJ, Brakhage DH (1995) Efficiency of adaptive cluster sampling for estimating density of wintering waterfowl. Biometrics 51:777–788
Smith DR, Villella RF, Lemarie DP (2003) Applications of adaptive cluster sampling to low-density populations of freshwater mussels. Environ Ecol Stat 10:7–15
Thompson SK (1990) Adaptive cluster sampling. J Am Stat Assoc 85:1050–1059
Thompson SK (1997) Adaptive sampling in behavioral surveys. In: Harrison L, Hughes A (eds) The validity of self-reported drug use: improving the accuracy of survey estimates. NIDA research monograph 167. National Institute of Drug Abuse, Rockville pp 296–319
Thompson SK (1991a) Adaptive cluster sampling: design with primary and secondary units. Biometrics 47:1103–1115
Thompson SK (1991b) Stratified adaptive cluster sampling. Biometrika 78:389–397
Thompson SK (2002) Sampling, 2nd edn. Wiley, New York
Thompson SK, Collins LA (2002) Adaptive sampling in research on risk-related behaviors. Drug Alcohol Depend 68:S57–S67
Thompson SK, Seber GAF (1996) Adaptive sampling. Wiley, New York
Vasudevan K, Kumar A, Chellam R (2001) Structure and composition of rainforest floor amphibian communities in Kalakad-Mundanthurai Tiger Reserve. Curr Sci 80(3):406–412

Author Biographies

Chang-Tai Chao received his Ph.D. in statistics from The Pennsylvania State University, State College, PA, USA, in 1999 under the advisement of Dr. Steven K. Thompson. Upon he finished his graduate study; he served as an assistant professor in the Department of Statistics, School of Management, National Cheng Kung University, Tainan City, Taiwan. He is currently serving as an associate professor in the same department. His main research interests include optimal sampling strategy, spatial sampling methods, adaptive sampling and related sampling survey issues.

Arthur L. Dryver completed his undergraduate in Mathematical Sciences/Statistics from Rice University, Houston, TX, USA, in 1993. He completed his Ph.D. in Statistics from The Pennsylvania State University, State College, PA, USA, in 1999. He worked as a consultant (PricewaterhouseCoopers, AnaBus,

and Experian) in the United States from 1999 to 2003. He became a lecturer at the National Institute of Development Administration (NIDA), Bangkok, Thailand since 2003. He is presently an Associate Professor in the Graduate School of Business Administration, NIDA, Bangkok, Thailand. His main area of research is in adaptive sampling strategies.

Tzu-Ching Chiang completed her bachelor degree majored in Statistics from the Department of Statistics, National Cheng Kung University, Tainan City, Taiwan. She received her master degree in the same department under the advisement of Dr. Chang-Tai Chao. She became a research assistant in the Institute of Statistical Science, Academia Sinica, Taiwan. She is currently a freelancer working in statistical consulting. Her research interests include statistical consulting and sampling survey.