Environmetrics 00, 1–26

DOI: 10.1002/env.XXXX

Adaptive cluster sampling for negatively correlated data

Stefano A. Gattone[a],* Esha Mohamed[b], Arthur L. Dryver[c] and A. Ralf Münnich[b]

Summary: Adaptive cluster sampling is a design specifically developed for rare and clustered populations. Under this sampling design, we consider the case when an auxiliary variable is available in addition to the variable of interest. The use of auxiliary information has been shown to improve efficiency of estimators though this results in biased estimators. The paper proposes two product estimators and their associated variance estimators for the adaptive cluster sampling design to be used when the study and the auxiliary variables are negatively correlated. For example, the population total of wildlife in a protected area can partly be influenced by such factors as diseases and pollution where presence of wildlife diseases or higher environmental pollution decreases population totals and distribution of wildlife. The exact expression of the bias together with the mean square error to the first degree of approximation have been obtained. We derive the conditions under which the suggested estimators provide more accurate estimation than the Horvitz-Thompson and Hansen-Hurwitz estimators with adaptive cluster sampling and the product estimator with simple random sampling. A simulation study is carried out to show the performance of the proposed estimators. Moreover, theoretical findings are supported by a numerical example with real data.

Keywords: Adaptive sampling; Product estimator; Coefficient of variation; Correlation coefficient.

[a] *DIGF, University of Tor Vergata, Rome, Italy*
[b] *Department of Social Statistics, University of Trier, Germany*
[c] *School of Business Administration, National Institute of Development Administration, Bangkok, Thailand*
[*] *Correspondence to: S.A. Gattone, DIGF, Tor Vergata university, Rome, Italy. E-mail: gattone@economia.uniroma2.it*

This paper has been submitted for consideration for publication in *Environmetrics*

1. INTRODUCTION

Adaptive cluster sampling (ACS) is an attractive sampling design since it can perform more efficiently than conventional designs for geographically rare and clustered populations (Thompson, 1990). Because of its flexibility, ACS designs have been developed under different settings by several authors. Thompson (1991) developed an ACS design in which the initial sample is selected by stratified sampling while Salehi and Seber (1997) introduced two stage adaptive cluster sampling. ACS in which the initial sample is selected by systematic sampling was considered by Dryver *et al.* (2012). Applications of ACS include Smith *et al.* (1995) who applied ACS to determine the density of Wintering Waterfowl population and Philippi (1995) who estimated abundance of some rare plants using ACS.

Under the ACS design, the estimators usually make use of only the information provided by the target variable. It is well known that the use of auxiliary information at the estimation stage helps in improving the efficiency of the estimators of the population parameters when the study variable y is highly correlated with the auxiliary variable x. Auxiliary information could be obtained from different sources such as previous census data, satellite images and sampling frames providing not only the unit identification labels but also the values of an auxiliary variable such as temperature, rainfall, soil classification, habitat and community attributes (Barabesi and Marcheselli, 2004; Salehi *et al.*, 2013).

Estimators in this context that make use of auxiliary information are ratio and product estimators. Ratio estimators are suitable when the correlation between y and x is highly positive. On the other hand, if the correlation is high but negative, product estimators can be used (Murthy, 1964).

Chao (2004) was the first to analyze the behavior of the ACS design when auxiliary information positively correlated with the variable of interest is available. He proposed a generalized ratio estimator based on the modified Horvitz-Thompson (HT) estimator under the ACS design. Dryver and Chao (2007) proposed another ratio estimator under

ACS based on the modified Hansen-Hurwitz (HH) estimator. Finally, Chao *et al.* (2011) and Lin and Chao (2014) completed the view by improving the ratio estimators under adaptive cluster sampling by making use of the Rao-Blackwell estimators.

In this paper we consider the case when the study and the auxiliary variables display a negative correlation. In particular we present modified Horvitz-Thompson and Hansen-Hurwitz product estimators for estimating finite population mean in adaptive cluster sampling. Their properties are analyzed and their variance estimators are derived and evaluated.

Product estimators are indeed useful whenever a high negative relationship between the study and auxiliary variable is observed. In wildlife, negative correlation has been reported between Buffalo and distance to water sources (Hopcraft *et al.*, 2012; Bro-Jorgensen, 2003; Winnie *et al.*, 2008), small grazers (Grant's and Thompson's gazelle) and predators, middle sized grazers (Topi and Coke's Hartebeest) and food quality, Thompson's gazelle and high rainfall areas (Hopcraft *et al.*, 2012).

In Section 2, the classical ACS designs with its corresponding estimators are presented. Product estimation is developed in Subsection 3.1. Estimation under the ACS design for negative correlated data is described in Subsection 3.2. Conditions under which the suggested estimators have smaller mean square error than the HH and HT estimators under ACS and the product estimator under SRS are derived in Section 4. An illustrative example is provided in Section 5. A simulation study that compares the estimators under different settings is given in Section 6. A real application on animal populations and GIS-derived environmental variables is provided in Section 7. Discussion and conclusion are given in Section 8.

2. ACS WITH ITS DESIGN UNBIASED ESTIMATORS

Consider a finite population of N units in which a survey variable y takes values $y_1, y_2, \ldots, y_N$. The target parameter is $\mu_y = \frac{1}{N}\sum_{i=1}^{N} y_i$, the unknown population mean of the study variable y. The ACS design starts with an initial sample of n units selected from a population of size N by a conventional sampling design. If any of these initially selected units satisfies a condition of interest, C, neighboring units are added to the sample and observed. If any other unit in that neighborhood also satisfies C, further sampling is performed according to the defined neighborhood. The process of adding neighborhood units continues until there are no units in the neighborhood that satisfy C. The set of units satisfying condition C around the unit in the initial sample form a network. In environmental surveys, the condition for extra sampling is usually based on the count of a biological species while the neighborhood is usually defined by spatial proximity. Thompson (1990) derived two design unbiased estimators for estimating the population mean.

Consider a population partitioned into K distinct networks, the modified Horvitz-Thompson estimator is given by

$$\hat{\mu}_{HT} = \frac{1}{N}\sum_{k=1}^{K}\frac{z_k y_k^*}{\alpha_k} \tag{1}$$

where z_k is an indicator variable which equals one if any unit of the k-th network is in the initial sample, y_k^* is the sum of y values in the k-th network while α_k is the probability of including unit i in the k-th network given by

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

where m_k is the size of the k-th network. The design variance of $\hat{\mu}_{HT}$ is given by

$$V_{HT} = \frac{1}{N^2} \sum_{j=1}^{K} \sum_{k=1}^{K} y_j^* y_k^* \frac{(\alpha_{jk} - \alpha_j \alpha_k)}{(\alpha_j \alpha_k)} \tag{2}$$

where α_{jk} is the joint probability of networks j and k being intersected by the initial sample defined as

$$\alpha_{jk} = 1 - \frac{\binom{N-m_j}{n} + \binom{N-m_k}{n} - \binom{N-m_k-m_j}{n}}{\binom{N}{n}}.$$

The unbiased estimator of the variance of $\hat{\mu}_{HT}$ is given by

$$v_{HT} = \frac{1}{N^2} \sum_{j=1}^{K} \sum_{k=1}^{K} z_j z_k y_j^* y_k^* \frac{(\alpha_{jk} - \alpha_j \alpha_k)}{(\alpha_j \alpha_k \alpha_{jk})}. \tag{3}$$

An alternative estimator is the modified Hansen-Hurwitz (HH) estimator given by

$$\hat{\mu}_{HH} = \frac{1}{n} \sum_{k=1}^{n} w_{yk} \tag{4}$$

where summation is over units sampled and w_{yk} is the average of y values in the k-th network. The design variance of $\hat{\mu}_{HH}$ is given by

$$V_{HH} = \frac{N-n}{Nn} \sum_{k=1}^{N} \frac{(w_{yk} - \mu_y)^2}{N-1}. \tag{5}$$

The unbiased estimator of the variance of $\hat{\mu}_{HH}$ is given by

$$v_{HH} = \frac{N-n}{Nn} \sum_{k=1}^{n} \frac{(w_{yk} - \hat{\mu}_{HH})^2}{n-1}. \tag{6}$$

3. PRODUCT METHOD OF ESTIMATION

Consider a population divided into N units and for each value y_i of the study variable, there is a value of the auxiliary variable x_i, $i = 1, 2, \dots, N$. The aim is to estimate the population mean μ_y of the variable of interest y, given that the population mean μ_x of the auxiliary variable x is known. We first review the classical product estimator (Murthy, 1964) employed quite effectively with SRS in the presence of negative correlation between study variable y and auxiliary variable x. Then we propose two product estimators to be used with the ACS design.

3.1. Product estimator under SRS

For a simple random sample of size n drawn without replacement, let $\bar{y}$ and $\bar{x}$ be the sample mean estimators of μ_y and μ_x, respectively. The usual product estimator of the population mean μ_y is given by

$$\hat{\mu}_p = \frac{\bar{y}\bar{x}}{\mu_x} \tag{7}$$

whose bias and mean square error (MSE) are given by

$$B(\hat{\mu}_p) = \frac{N-n}{Nn}\frac{S_{xy}}{\mu_x}$$

and

$$MSE(\hat{\mu}_p) = \frac{N-n}{Nn}\left[S_y^2 + R^2 S_x^2 + 2RS_{xy}\right], \tag{8}$$

respectively, with $S_y^2 = \frac{1}{N-1}\sum_{i=1}^N (y_i - \mu_y)^2$, $S_x^2 = \frac{1}{N-1}\sum_{i=1}^N (x_i - \mu_x)^2$, $S_{xy} = \frac{1}{N-1}\sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y)$ and $R = \frac{\mu_y}{\mu_x}$.

By defining $C_y = \frac{S_y}{\mu_y}$ and $C_x = \frac{S_x}{\mu_x}$, the MSE of the product estimator can be written as

$$MSE(\hat{\mu}_p) = \mu_y^2(C_y^2 + C_x^2 + 2\rho_{xy}C_yC_x)$$

with ρ_{xy} denoting the correlation coefficient between y and x. Murthy (1964) analyzed the relative precision of the product estimator and showed that it is more efficient than the sample mean if

$$\rho_{xy} < -\frac{1}{2}\frac{C_x}{C_y}. \tag{9}$$

3.2. Product estimator under ACS

The product estimator $\hat{\mu}_p$ assumes simple random sampling without replacement. In many ACS survey studies, in addition to the variable of interest y, an auxiliary variable x may be collected to improve estimation accuracy and precision. We present two modified product estimators in order to extend the underlying idea of product estimation to the ACS sample selection method. We further provide their bias along with their variance estimators.

3.2.1. The HH product estimator

The product estimator in (7) can be modified by substituting the sample means $\bar{y}$ and $\bar{x}$ with the corresponding HH estimator given in (4). The proposed product estimator for ACS is given by

$$\hat{\mu}_{pHH} = \frac{\hat{\mu}_{HH_y}\hat{\mu}_{HH_x}}{\mu_x} \tag{10}$$

where $\hat{\mu}_{HH_y}$ and $\hat{\mu}_{HH_x}$ are the HH estimators of μ_y and μ_x, respectively.

To obtain the bias, the variance and MSE we write

$$e_0 = \frac{\hat{\mu}_{HH_y} - \mu_y}{\mu_y} \Rightarrow \hat{\mu}_{HH_y} = \mu_y(e_0 + 1)$$

$$e_1 = \frac{\hat{\mu}_{HH_x} - \mu_x}{\mu_x} \Rightarrow \hat{\mu}_{HH_x} = \mu_x(e_1 + 1)$$

then

$$\begin{aligned} E(e_0) = E(e_1) &= 0 \\ E(e_0^2) &= \frac{V_{HH_y}}{\mu_y^2} \\ E(e_1^2) &= \frac{V_{HH_x}}{\mu_x^2} \\ E(e_0 e_1) = \rho_{HH_{xy}} &\frac{\sqrt{V_{HH_y} V_{HH_x}}}{\mu_y \mu_x} \end{aligned}$$

where V_{HH_y} and V_{HH_x} are the variances of the HH estimators given in (5),

$$\rho_{HH_{xy}} = \frac{V_{HH_{xy}}}{\sqrt{V_{HH_y} V_{HH_x}}}$$

and

$$V_{HH_{xy}} = \frac{N-n}{Nn} \frac{1}{N-1} \sum_{k=1}^{N} (w_{yk} - \mu_y)(w_{xk} - \mu_x).$$

Hence, the bias is given by

$$\begin{aligned}
B(\hat{\mu}_{pHH}) &= E(\hat{\mu}_{pHH} - \mu_y) \\
&= \mu_y E\left[(1+e_0)(1+e_1) - 1\right] \\
&= \mu_y E(e_0 + e_1 + e_0 e_1) \\
&= \mu_y \rho_{HH_{xy}} \frac{\sqrt{V_{HH_y} V_{HH_x}}}{\mu_y \mu_x} \\
&= \frac{V_{HH_{xy}}}{\mu_x}. \qquad (11)
\end{aligned}$$

An unbiased estimator of the bias is given by

$$\hat{B}(\hat{\mu}_{pHH}) = \frac{v_{HH_{xy}}}{\mu_x}$$

where

$$v_{HH_{xy}} = \left(\frac{N-n}{Nn}\right) \frac{1}{n-1} \sum_{k=1}^{n} (w_{yk} - \hat{\mu}_{HH_y})(w_{xk} - \hat{\mu}_{HH_x}).$$

An approximation of the variance of $\hat{\mu}_{pHH}$ is given by

$$var(\hat{\mu}_{pHH}) \approx \left(\frac{N-n}{Nn}\right) \frac{1}{N-1} \sum_{k=1}^{N} (w_{yk} + R w_{xk})^2.$$

An estimator for this variance is obtained from

$$\widehat{var}(\hat{\mu}_{pHH}) \approx \left(\frac{N-n}{Nn}\right) \frac{1}{n-1} \sum_{k=1}^{n} (w_{yk} + \hat{R} w_{xk})^2. \qquad (12)$$

The MSE can be obtained by writing

$$
\begin{aligned}
MSE(\hat{\mu}_{pHH}) &= E(\hat{\mu}_{pHH} - \mu_y)^2 \\
&= \mu_y^2 E(e_0 + e_1 + e_0 e_1)^2 \\
&= \mu_y^2 E(e_0^2 + e_1^2 + 2e_0 e_1 + 2e_0^2 e_1 + 2e_0 e_1^2 + e_0^2 e_1^2).
\end{aligned}
$$

The MSE to the first order approximation is thus given by

$$
\begin{aligned}
MSE(\hat{\mu}_{pHH}) &\approx \mu_y^2 \left(\frac{V_{HH_y}}{\mu_y^2} + \frac{V_{HH_x}}{\mu_x^2} + 2\rho_{HH_{xy}} \frac{\sqrt{V_{HH_y} V_{HH_x}}}{\mu_y \mu_x} \right) \\
&\approx \mu_y^2 \left(\frac{V_{HH_y}}{\mu_y^2} + \frac{V_{HH_x}}{\mu_x^2} + 2\frac{V_{HH_{xy}}}{\mu_y \mu_x} \right) \qquad (13)
\end{aligned}
$$

which can be estimated by

$$
\widehat{MSE}(\hat{\mu}_{pHH}) \approx \bar{y}^2 \left(\frac{v_{HH_y}}{\hat{\mu}_{HH_y}^2} + \frac{v_{HH_x}}{\hat{\mu}_{HH_x}^2} + 2\frac{v_{HH_{xy}}}{\hat{\mu}_{HH_y} \hat{\mu}_{HH_x}} \right)
$$

where v_{HH_y} and v_{HH_x} are the variance estimator given in (6). A useful alternative formulation for the MSE can be obtained by noting that ACS actually represents a sampling without replacement of n networks with inclusion probability $\frac{m_k}{N}$. Let $w_{y1}, w_{y2}, \ldots, w_{yN}$ be the population network means of the y-values with variance $S_{w_y}^2 = \sum_{i=1}^{N} \frac{(w_{yi} - \mu_y)^2}{N-1}$ and let $w_{x1}, w_{x2}, \ldots, w_{xN}$ be the population network means of the x-values with variance $S_{w_x}^2$. Furthermore, define the covariance between these two network populations as $S_{w_{xy}} = \frac{1}{N-1} \sum_{i=1}^{N} (w_{yi} - \mu_y)(w_{xi} - \mu_x)$. The HH estimators $\hat{\mu}_{HH_y}$ and $\hat{\mu}_{HH_x}$ simply correspond to the sample mean estimator applied to the transformed populations of y and x, respectively (Thompson and Seber, 1996; Dryver and Chao, 2007).

Then, we have that $V_{HH_y} = \frac{N-n}{Nn} S^2_{w_y}$, $V_{HH_x} = \frac{N-n}{Nn} S^2_{w_x}$, $V_{HH_{xy}} = \frac{N-n}{Nn} S_{w_{xy}}$ and by substitution in (13) we obtain

$$\begin{aligned} MSE(\hat{\mu}_{pHH}) &\approx \frac{N-n}{Nn}\mu_y^2 \left(\frac{S^2_{w_y}}{\mu_y^2} + \frac{S^2_{w_x}}{\mu_x^2} + 2\frac{S_{w_{xy}}}{\mu_y \mu_x} \right) \\ &\approx \frac{N-n}{Nn}\mu_y^2 \left(C^2_{w_y} + C^2_{w_x} + 2\rho_{w_{xy}} C_{w_y} C_{w_x} \right) \end{aligned} \quad (14)$$

where $C_{w_y} = \frac{S_{w_y}}{\mu_y}$, $C_{w_x} = \frac{S_{w_x}}{\mu_y}$ and $\rho_{w_{xy}}$ is the correlation at network level between y and x.

3.2.2. The HT product estimator

From the HT estimator in (1) we can derive another product estimator for ACS as follows:

$$\hat{\mu}_{pHT} = \frac{\hat{\mu}_{HT_y}\hat{\mu}_{HT_x}}{\mu_x}. \quad (15)$$

As for the HH product estimator we can derive the bias given by

$$B(\hat{\mu}_{pHT}) = \frac{V_{HT_{xy}}}{\mu_x} \quad (16)$$

and, to the first degree of approximation, the MSE given by

$$MSE(\hat{\mu}_{pHT}) \approx V_{HT_y} + R^2 V_{HT_x} + 2R V_{HT_{xy}} \quad (17)$$

where V_{HT_y} and V_{HT_x} are the variances of the HT estimators given in (2) and $V_{HT_{xy}}$ is the covariance between $\hat{\mu}_{HT_y}$ and $\hat{\mu}_{HT_x}$ equal to

$$V_{HT_{xy}} = \frac{1}{N^2}\sum_{j=1}^{K}\sum_{k=1}^{K} \Delta_{jk} \frac{y_j^*}{\alpha_j}\frac{x_k^*}{\alpha_k}$$

with $\Delta_{jk} = \alpha_{jk} - \alpha_j \alpha_k$.

With a straightforward use of notation, an alternative equivalent formulation can be given as follows

$$MSE(\hat{\mu}_{pHT}) \approx \mu_y^2 \left(C_{HT_y}^2 + C_{HT_x}^2 + 2\rho_{HT_{xy}} C_{HT_y} C_{HT_x} \right). \tag{18}$$

The MSE can be estimated by

$$\widehat{MSE}(\hat{\mu}_{pHT}) \approx v_{HT_y} + \hat{R}^2 v_{HT_x} + 2\hat{R} v_{HT_{xy}}$$

where v_{HT_y} and v_{HT_x} are the variance estimators given in (3) and $v_{HT_{xy}}$ is the unbiased estimator of $V_{HT_{xy}}$ given by

$$v_{HT_{xy}} = \frac{1}{N^2} \sum_{j=1}^{K} \sum_{k=1}^{K} \check{\Delta}_{jk} \frac{z_j y_j^*}{\alpha_j} \frac{z_k x_k^*}{\alpha_k}$$

with $\check{\Delta}_{jk} = \frac{\alpha_{jk} - \alpha_j \alpha_k}{\alpha_{jk}}$.
An approximation of the variance of $\hat{\mu}_{pHT}$ is given by

$$var(\hat{\mu}_{pHT}) \approx \frac{1}{N^2} \sum_{j=1}^{K} \sum_{k=1}^{K} \Delta_{jk} \frac{(y_j^* + R x_j^*)}{\alpha_j} \frac{(y_k^* + R x_k^*)}{\alpha_k}$$

with variance estimate

$$\widehat{var}(\hat{\mu}_{pHT}) \approx \frac{1}{N^2} \sum_{j=1}^{K} \sum_{k=1}^{K} \check{\Delta}_{jk} \frac{z_j (y_j^* + \hat{R} x_j^*)}{\alpha_j} \frac{z_k (y_k^* + \hat{R} x_k^*)}{\alpha_k}. \tag{19}$$

4. EFFICIENCY COMPARISONS OF THE PRODUCT ESTIMATOR

The proposed product estimators for ACS will now be theoretically compared in terms of efficiency with respect to

- standard ACS estimators with no use of auxiliary information

- product estimator under SRS.

Since comparisons are based on the first order approximation of the MSE of the suggested estimators, a simulation study will be provided in the next section to support the theoretical findings.

4.1. Comparisons with ACS

Thompson (1990) derived the condition under which ACS is more efficient than SRS. We can now analyze the condition under which it turns useful to use auxiliary information with ACS. By means of (14) we can easily verify that $MSE(\hat{\mu}_{pHH}) < V_{HH}$, *i.e.* the product estimator $\hat{\mu}_{pHH}$ is more efficient than the direct estimator $\hat{\mu}_{HH}$ when

$$\frac{N-n}{Nn}\mu_y^2\left(C_{w_y}^2 + C_{w_x}^2 + 2\rho_{w_{xy}}C_{w_y}C_{w_x}\right) < \frac{N-n}{Nn}S_{w_y}^2$$

which simplifies to

$$\rho_{w_{xy}} < -\frac{1}{2}\frac{C_{w_x}}{C_{w_y}}. \tag{20}$$

Analogously, by means of (18), we have that the product HT-type estimator $\hat{\mu}_{pHT}$ is more efficient than $\hat{\mu}_{HT}$, *i.e.* $MSE(\hat{\mu}_{pHT}) < V_{HT}$ if

$$\rho_{HT_{xy}} < -\frac{1}{2}\frac{C_{HT_x}}{C_{HT_y}}. \tag{21}$$

Both relations are similar to condition (9) which compares the product estimator with the sample mean under SRS. However, conditions (20) and (9) are more easily interpretable since they are linked to the population values of y and x. Indeed, $\rho_{w_{xy}}$ measures the correlation between y and x at the network level while ρ_{xy} is the correlation between y and x at unit level. Similarly, C_{w_x} and C_{w_y} are the coefficient of variations of the population network values

w_{xi} and w_{yi} while C_x and C_y are the coefficient of variations of the population values x and y.

For the product HT estimator, the correlation coefficient $\rho_{HT_{xy}}$ and the coefficient of variations C_{HT_y} and C_{HT_x} depend on the network totals y_j^* and inclusion probabilities α_j. Hence, condition (21) also relates somehow to the correlation and the variation of x and y at the network level.

We will now focus on $\hat{\mu}_{HH}$ but the same considerations hold also for $\hat{\mu}_{HT}$.

If at the network level, y and x have the same coefficient of variation then the product estimator is better than the direct estimator if $\rho_{w_{xy}} \leq -0.5$. As the variability of the x variable increases, more negative correlation $\rho_{w_{xy}}$ is necessary to have the product estimator more efficient than the direct one. As the variability of the x variable decreases, less negative correlation $\rho_{w_{xy}}$ is necessary to have the product estimator more efficient than the direct one. When, for example, $C_{w_x} = 2C_{w_y}$ then whatever the correlation $\rho_{w_{xy}}$, the product estimator will always be less efficient than the direct estimator. When, for example, $C_{w_y} = 2C_{w_x}$, $\rho_{w_{xy}}$ has to be less than -0.25 in order to have the product estimator more efficient than the direct one.

The use of the auxiliary variable would be recommended when the correlation is present at the network level because this will transfer to the value of $\rho_{w_{xy}}$. For rare and clustered populations it would be difficult to find a situation in which $C_{w_y} < C_{w_x}$. The use of the product estimator is recommended as long as it has a negative correlation with the y variable and the y values are rare and clustered.

4.2. Comparison with product under SRS

Another useful comparison is with the product estimator under SRS. Let us write the MSE of $\hat{\mu}_p$ in (8) in terms of the variances of $\bar{y}$ and $\bar{x}$ and the covariance between them. With a straightforward use of notation we have that

$$MSE(\hat{\mu}_p) = V_{\bar{y}} + R^2 V_{\bar{x}} + 2RV_{\bar{x}\bar{y}}.$$

By recalling (17) we have that $MSE(\hat{\mu}_{pHT}) < MSE(\hat{\mu}_p)$, *i.e.* $\hat{\mu}_{pHT}$ is more efficient than $\hat{\mu}_p$ when

$$V_{HTy} + R^2 V_{HTx} + 2RV_{HTxy} < V_{\bar{y}} + R^2 V_{\bar{x}} + 2RV_{\bar{x}\bar{y}}. \tag{22}$$

or, equivalently when

$$C^2_{HTy} + C^2_{HTx} + 2\rho_{HTxy} C_{HTy} C_{HTx} < C^2_{\bar{y}} + C^2_{\bar{x}} + 2\rho_{xy} C_{\bar{x}} C_{\bar{y}}. \tag{23}$$

A similar condition can be obtained for the HH product estimator $\hat{\mu}_{pHH}$. By recalling (13) we have that $MSE(\hat{\mu}_{pHH}) < MSE(\hat{\mu}_p)$, *i.e.* $\hat{\mu}_{pHH}$ is more efficient than $\hat{\mu}_p$ when

$$V_{HHy} + R^2 V_{HHx} + 2RV_{HHxy} < V_{\bar{y}} + R^2 V_{\bar{x}} + 2RV_{\bar{x}\bar{y}} \tag{24}$$

or, equivalently, by noting that $\rho_{w_{xy}} = \rho_{HHxy}$ when

$$C^2_{HHy} + C^2_{HHx} + 2\rho_{w_{xy}} C_{HHy} C_{HHx} < C^2_{\bar{y}} + C^2_{\bar{x}} + 2\rho_{xy} C_{\bar{x}} C_{\bar{y}}. \tag{25}$$

The efficiency of the product estimators under ACS depends by the conjoint effects of three factors:

- efficiency of ACS vs SRS for the y variable (V_{HTy} vs $V_{\bar{y}}$ or V_{HHy} vs $V_{\bar{y}}$)
- efficiency of ACS vs SRS for the x variable (V_{HTx} vs $V_{\bar{x}}$ or V_{HHx} vs $V_{\bar{x}}$)
- size of the correlation at the network level, ρ_{HTxy} or $\rho_{w_{xy}}$ and of the correlation at unit level, ρ_{xy}.

These three factors interact with each other. Consider the situation in which the population

y values are rare and clustered so to have ACS more efficient than SRS. In such a case, we would have $V_{HTy} < V_{\bar{y}}$ or, equivalently $C^2_{HTy} < C^2_{\bar{y}}$. In presence of a negative correlation between y and x, the spread of the auxiliary variable may be such that the HT estimator $\hat{\mu}_{HTx}$ is less efficient than $\bar{x}$, i.e. $V_{HTx} > V_{\bar{x}}$. The final effect on the efficiency of $\hat{\mu}_{pHT}$ over $\hat{\mu}_p$ will be positive if the efficiency gain of ACS over SRS for the y variable will overcome the efficiency loss of ACS over SRS for the x variable.

Finally, one has to consider the values of ρ_{HTxy} and ρ_{xy}. Keeping fixed the previous two factors, a correlation at the network level higher than a correlation at unit level will cause a positive effect in terms of efficiency of $\hat{\mu}_{pHT}$ over $\hat{\mu}_p$.

5. ILLUSTRATIVE EXAMPLE

To shed light on the computations, we consider two small populations of size $N = 5$ as shown in Table A.1 where the condition to adaptively add adjacent units is $C = (y : y \geq 5)$. The y values are from Thompson (1990) while the x values are such that the correlation at unit level is equal to $\rho_{xy} = -0.09$ and the correlation at network level is equal to $\rho_{w_{xy}} = -0.95$. The population means for y and x are 202.6 and 5.4 respectively.

[Table 1 about here.]

In Table A.2 are listed the $\binom{5}{2} = 10$ possible samples of ACS when an initial sample of size $n = 2$ is selected by simple random sampling without replacement. In Table A.2, for each sample we compute the estimators $\bar{y}$, $\bar{x}$, $\hat{\mu}_{HHy}$, $\hat{\mu}_{HHx}$, $\hat{\mu}_{HTy}$, $\hat{\mu}_{HTx}$, $\hat{\mu}_{pHH}$ and $\hat{\mu}_{pHT}$.

[Table 2 about here.]

The population values are such that condition (9) is not satisfied since $\rho_{xy} = -0.09$ is greater than $-\frac{1}{2}\frac{C_x}{C_y} = -0.11$. Product estimator is not suitable for this population under SRS. On the other hand, due to the high value of correlation at network level product estimation turns out useful under ACS. Both conditions (20) and (21) are satisfied since $\rho_{w_{xy}} = -0.95$ is less than $-\frac{1}{2}\frac{C_w x}{C_w y} = -0.16$ and $\rho_{HT_{xy}} = -0.69 < -\frac{1}{2}\frac{C_{HT_x}}{C_{HT_y}} = -0.22$. For this small example, the most efficient estimator of μ_y is the product estimator $\hat{\mu}_{HHp}$ with an MSE equal to 13579.58.

6. SIMULATION STUDY

In this section we perform a simulation study in order to describe the situations in which the use of an auxiliary variable together with an adaptive cluster design works most efficiently. Artificial populations of the y variable were generated using the Poisson Cluster Process (Diggle, 1983) with parameters λ_1 and λ_2. They represent the number of parents and the number of offsprings, respectively. The offsprings are uniformly and independently allocated around their parents at a radial distance selected from an exponential distribution with mean equal to 0.5. The auxiliary variable x was simulated so to have a negative correlation with the y variable at network level. The simulation setting is as follows:

- $\lambda_1 = 5, 10$
- $\lambda_2 = 5, 10, 15, 50, 100$
- correlation at network level: low ($-0.1 < \rho_{w_{xy}} \leq 0$), intermediate ($-0.4 \leq \rho_{w_{xy}} \leq -0.6$), high ($\rho_{w_{xy}} \leq -0.8$)

The combinations of the population parameters cover different degrees of rarity and aggregation, from very rare and clustered populations ($\lambda_1 = 5,\ \lambda_2 = 5$) to more sparse and less clustered populations ($\lambda_1 = 10,\ \lambda_2 = 100$) and different levels of network correlation between x and y. This enabled the study of the suggested estimators to populations where ACS or product estimation are known to perform better than the SRS mean estimator and where ACS or product estimation are known to be less efficient relative to the SRS mean estimator.

ACS was set up with different values of the initial sample size $n = 5, 10, 20$ and 50 while for SRS, the size of the sample was set to the empirical expectations of the sample size under the adaptive designs. For each population we drew 50000 samples with the corresponding sampling designs from which the values of the estimators $\hat{\mu}_p$, $\hat{\mu}_{HT}$, $\hat{\mu}_{HH}$, $\hat{\mu}_{pHH}$ and $\hat{\mu}_{pHT}$ are computed.

Detailed results of the simulation are presented in the Appendix. In particular, the relative efficiency of each estimator computed with respect to SRS is displayed in Figures from A.1 to A.6. Relative bias of $\hat{\mu}_{pHH}$ and $\hat{\mu}_{pHT}$ is displayed in Tables A.1 and A.2 while relative bias of the variance estimators $\widehat{var}(\hat{\mu}_{pHT})$ and $\widehat{var}(\hat{\mu}_{pHH})$ are reported in Tables A.3 and A.4. Finally, the accuracy of the MSE approximations $MSE(\hat{\mu}_{pHH})$ and $MSE(\hat{\mu}_{pHT})$ is examined in Tables A.5 and A.6.

6.1. Efficiency comparisons

With a low correlation at network level the proposed product estimators have a similar behavior of the standard HH and HT estimators when $\lambda_1 = 5$. It is observed a slight loss of efficiency when $\lambda_2 = 10$ especially with the HT product estimator. Overall, this suggests that auxiliary information could be used with ACS without the risk of efficiency loss if correlation is not present at network level. Under ACS, with negative correlation at network level, product estimation becomes useful and more efficient than standard ACS estimators. The efficiency gain of the product estimators $\hat{\mu}_{pHH}$ and $\hat{\mu}_{pHT}$ with respect to $\hat{\mu}_{HH}$ and $\hat{\mu}_{HT}$ increases with the correlation at network level and λ_2 and it is higher for $\lambda_1 = 5$. Furthermore, the efficiency gains are less pronounced with large values of the initial sample size n.

An interesting result is observed for those populations where ACS is less efficient than SRS, *i.e.* $RE_{ACS_{HH}}$ and $RE_{ACS_{HT}}$ are both less than 1 ($\lambda_1 = 10$, $\lambda_2 \leq 10$, $\lambda_2 = 100$ and $n \leq 20$). The presence of negative correlation at network level ($\rho_{w_{xy}} < -0.4$) makes product estimation under ACS more efficient than SRS, *i.e.* $RE_{ACS_{pHH}}$ and $RE_{ACS_{pHT}}$ are both greater than 1.

In general, if the simulated populations have no correlation at unit level then the product estimator $\hat{\mu}_p$ does not show a good performance with respect to both ACS and ACS product estimators. However, the populations with $\lambda_1 = 10$ and $\lambda_2 = 5$ have same negative correlation at unit level and $\hat{\mu}_p$ results to be the most efficient estimator. The comparison of $RE_{ACS_{pHH}}$ and $RE_{ACS_{pHT}}$ supports the general findings about the HH and HT estimators under ACS, *i.e.* $\hat{\mu}_{pHT}$ is in general more efficient than $\hat{\mu}_{pHH}$ especially for high values of n.

6.2. Bias of the product estimators

The bias of $\hat{\mu}_{HT}$ and $\hat{\mu}_{pHH}$ given in (16) and (11) depends on the value of the correlation at network level. Indeed, results of relative bias in the Tables A.1 and A.2 clearly show how the bias is always negative and increases in absolute terms with the correlation level. The bias becomes negligible as the initial sample size becomes large ($n \geq 20$) for any value of network correlation. It is usually found to be unimportant even in samples of small size at any level of network correlation and for $\lambda_2 \leq 15$. It increases with λ_1 and λ_2. With $n = 5, 10$ and at intermediate and high network correlation levels, the product estimators seriously underestimate the mean value when $\lambda_2 = 50, 100$.

6.3. Bias of the variance estimators

Tables A.3 and A.4 provide relative bias of the variance estimators $\widehat{var}(\hat{\mu}_{pHT})$ and $\widehat{var}(\hat{\mu}_{pHH})$ given in (19) and (12), respectively. Results show that the variance estimators are nearly unbiased at low network correlation level. Relative bias is a bit larger when $\lambda_1 = 5$ and tends to increase with λ_2. With an initial sample size of $n = 5$ variance is underestimated at intermediate and high levels of correlation. With $n = 10$, relative bias is almost always less than 10%. For $n > 20$ it is almost always less than 5%.

6.4. Accuracy of the MSE approximation

Results from Tables A.5 and A.6 show that the accuracy of the MSE approximations of $\hat{\mu}_{pHT}$ and $\hat{\mu}_{pHH}$ given in (17) and (13) increases with the initial sample size n. The error is negligible (say a relative bias less than 0.05) for any value of n when $\rho_{w_{xy}}$ is low. With intermediate and high values of the network correlation, relative bias of the MSE approximations are almost always negligible for $n \geq 20$. The approximation error increases with the correlation level and decreases for higher values of λ_2. On average, it is higher in the populations with $\lambda_1 = 5$ for the HH product estimator. The opposite is observed for the HT product estimator where the approximation error is higher when $\lambda_1 = 10$. Grey cells in Tables A.1 and A.2 denote the combinations where the conditions (20) and (21) are not coherent with the results of the simulations in terms of relative efficiency of ACS product with respect to ACS. Some inconsistent result is observed only at low value of network correlation where ACS product and ACS have a similar behavior. In general, we may conclude that the theoretical results given in (20) and (21) are confirmed by the results of the simulations.

7. APPLICATION TO REAL DATA

The same estimators have been used in a case study. The data was obtained from an aerial survey conducted in 2010 at the Amboseli-West Kilimanjaro/ Magadi-Natron region. The region covers parts of Kenya and Tanzania between 1°37’ S and 3°13’ S and between 35°49’ E and 38°00’ E . The survey was conducted by Kenya Wildlife Service and Tanzania Wildlife Research Institute with other partners and covered a region of 24,108km^2. Details on the survey can be found at `http://www.kws.org/info/news/2013/24april2013tanzania.html`. The data was slightly modified by adding extra zeros to create a rectangular region that was divided into, 17 rows and 23 columns, $N = 391$ quadrats. The study variables of interest considered were counts of African Buffalo and Hartebeest. GIS-derived environmental variables were used as auxiliary information. In particular, values of the Minimum temperature (C) and Altitude (Km) for each quadrat over the study area were taken from the "WorldClim" data base freely available for download from http://www.worldclim.org (Hijmans *et al.*, 2005). Values of the variables are provided in Figure 1.

[Figure 1 about here.]

Table 3 reports population parameters of the study variables and auxiliary variables. Buffalo and Minimum temperature show a weak negative correlation at unit level, $\rho_{xy} = -0.087$. At network level the negative correlation increases up to $\rho_{w_{xy}} = -0.1845$. At network level, variability of Minimum temperature ($C_{w_x} = 0.1512$) is much smaller than variability of Buffalo ($C_{w_x} = 5.740$). Consequently, both conditions (20) and (21) are satisfied meaning that we should expect the product estimators to be more efficient than the direct estimators under ACS. The same holds for Hartebeest and Altitude even if correlations at unit and network level are more similar (being equal to $\rho_{xy} = -0.0973$ and $\rho_{w_{xy}} = -0.1187$). This would lower the efficiency gain of the product estimator over the direct estimator under ACS.

[Table 3 about here.]

All the above considerations are supported by the analysis of the real data. We implement the simulation study of Section 6 with study variables Buffalo and Hartebeest and auxiliary variables Minimum temperature and Altitude, respectively. Results of relative efficiencies shown in Table 4 confirm the ones obtained in the simulation study. The proposed product estimators $\hat{\mu}_{pHH}$ and $\hat{\mu}_{pHT}$ are better than the competing estimators. The efficiency gain over direct estimators under ACS decreases as the initial sample size increases. More pronounced efficiency gains are observed over the product estimator under SRS. As expected, efficiency gains are lower for the Hartebeest population than for the Buffalo population. Relative bias of the product estimators $\hat{\mu}_{pHH}$ and $\hat{\mu}_{pHT}$ and the proposed variance estimators $\widehat{var}(\hat{\mu}_{pHH})$

and $\widehat{var}(\hat{\mu}_{pHT})$ are also reported in Table 4. As already seen in the simulation with low negative correlation, product estimators and variance estimators are nearly unbiased with relative bias almost always less than 1%.

[Table 4 about here.]

8. DISCUSSION

For a researcher generally the first question would be what sampling design to use. When studying rare and clustered populations, ACS is often the best option for obtaining the most number of units of interest and is more efficient given the same final sample size. Even if there exist SRS estimators that are more efficient given the same expected final sample size ACS may still be the preferred option when desiring the maximum number of units of interest in the final sample. Thus if the researcher selects ACS the next question is what is the best estimator to use given the sampling design chosen.

The simulations results are based on simulated data assuming varying network level correlations. Network level correlation doesn't guarantee unit level correlation, nor is unit level negative correlation required for the product estimator under ACS. Thus if unit level correlation is very close to zero and this is not the case on a network level then ACS product estimators may be appropriate whereas SRS product estimators are not even appropriate to use. On the other hand if there exists a negative unit level correlation and a positive network correlation then the ACS product estimator is not appropriate and the researcher should not use the product estimator. For example if the auxiliary variable is rare and clustered and often found together with the variable of interest. In this case it could be possible for the two variables to be negatively correlated on a unit level but be positively correlated on a network level. The researcher should be cautious of this scenario even if it is unlikely to occur.

Temperature is a good example of an auxiliary variable in which a researcher would not expect positive/negative unit level correlation but might expect correlation on a network level. For example, there might be various reasons why a species might tend to live in cooler regions. Expecting a species to live in cooler regions though is different than expecting counts of the species to correlate on a unit level. That is within the cooler regions there may not be a pattern on a unit level. The difference is subtle but significant. It is the difference between expecting unit level correlation and network level correlation. This is the case with the African Buffalo where there is a larger negative correlation on a network level than on an unit level. This is also seen to a lesser extent with the Hartebeest and altitude. In both cases ACS product estimators were more efficient than their SRS product estimator counterpart. In addition, the SRS product estimator was actually less efficient at times compared to the sample mean. The real data findings support the findings in the simulation study in Section 6 given further evidence to the preference of ACS product estimation when there exists a network level correlation.

Ideally, the researcher would have an estimate or at least a concept for the network level correlation from a prior study on the same attributes of interest to determine if the ACS product estimator is the best estimator to use. Generally in the simulation as long as the initial sample size is greater than or equal to 20 the bias of the estimators proposed is small. In summary, once you've chosen ACS then as long as you believe there is a negative network level correlation, a sample size of 20 or more, and the auxiliary population mean is known then ACS product estimator would be preferable.

REFERENCES

Barabesi L, Marcheselli M, 2004. Design-based ranked set sampling using auxiliary variables. *Environmental and Ecological Statistics* **11**: 415–430.

Bro-Jorgensen J, 2003. The significance of hotslpots to lekking topi antelopes. *Behavioral Ecology and Sociobiology* **53**: 324–331.

Chao C, 2004. Ratio estimation on adaptive cluster sampling. *Journal of Chinese Statistical Association* **42**(3): 307–327.

Chao C, Dryver A, Chiang T, 2011. Leveraging the raoblackwell theorem to improve ratio estimators in adaptive cluster sampling. *Environmental and Ecological Statistics* **18**(3): 543–568.

Diggle P, 1983. *Statistical analysis of spatial point patterns*. Academic Press, London.

Dryver A, Chao C, 2007. Ratio estimators in adaptive cluster sampling. *Environmetrics* **18**: 607–620.

Dryver A, Netharn U, Smith D, 2012. Partial systematic adaptive cluster sampling. *Environmetrics* **23**: 306–316.

Hijmans R, Cameron E, Parra J, Jones P, Jarvis A, 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**: 1965–1978.

Hopcraft J, Anderson T, Perez-vila S, Mayembe E, 2012. Body size and the division of niche space: food and predation differentially shape the distribution of serengeti grazers. *Ecology* **81**: 201–213.

Lin F, Chao C, 2014. Variances and variance estimators of the improved ratio estimators under adaptive cluster sampling. *Environmental and Ecological Statistics* **21**: 285–311.

Murthy M, 1964. Product method of estimation. *Sankhya, A* **74**: 69–74.

Philippi T, 1995. Adaptive cluster sampling for estimation of abundances within local populations of low-abundance plants. *Ecology* **86**: 1091–1100.

Salehi M, Panahbehagh B, Parvardeh A, Smith D, Lei Y, 2013. Regression-type estimators for adaptive two-stage sequential sampling. *Environmental and Ecological Statistics* **20**: 571–590.

Salehi M, Seber G, 1997. Two-stage adaptive cluster sampling. *Biometrics* **53**: 959–970.

Smith D, Conroy M, Brakhage D, 1995. Efficiency of adaptive cluster sampling for estimating density of wintering waterfowl. *Biometrics* **51**: 777–788.

Thompson S, 1990. Adaptive cluster sampling. *Journal of the American Statistical Association* **85**: 1050–1059.

Thompson S, 1991. Stratified adaptive cluster sampling. *Biometrics* **78**: 389–397.

Thompson S, Seber G, 1996. *Adaptive Sampling*. Wiley, New York.

Winnie J, Cross P, Getz W, 2008. Habitat quality and heterogeneity influence distribution and behaviour in african buffalo. *Ecology* **89**: 1458–1468.

APPENDIX

To evaluate the empirical efficiency of the proposed estimators we compute for each population:

- relative efficiency of the product estimators under ACS
  $RE_{ACS_{pHH}} = \frac{MSE(\bar{y})}{MSE(\hat{\mu}_{pHH})}$ and $RE_{ACS_{pHT}} = \frac{MSE(\bar{y})}{MSE(\hat{\mu}_{pHT})}$
- relative efficiency of standard ACS estimators
  $RE_{ACS_{HH}} = \frac{MSE(\bar{y})}{MSE(\hat{\mu}_{HH})}$ and $RE_{ACS_{HT}} = \frac{MSE(\bar{y})}{MSE(\hat{\mu}_{HT})}$
- relative efficiency of the product estimator under SRS
  $RE_{SRS_p} = \frac{MSE(\bar{y})}{MSE(\hat{\mu}_p)}$.

Results of relative efficiency are displayed in Figures from A.1 to A.6.

To evaluate the bias of the proposed estimators we compute for each population:

- $RB_{ACS_{pHH}} = \frac{1}{50000} \sum_{b=1}^{50000} \frac{\hat{\mu}_{HHb} - \mu_y}{\mu_y}$
- $RB_{ACS_{pHT}} = \frac{1}{50000} \sum_{b=1}^{50000} \frac{\hat{\mu}_{HTb} - \mu_y}{\mu_y}$.

Results of relative bias are displayed in Tables A.1 and A.2.

Analogously, we compute relative bias of the variance estimators $\widehat{var}(\hat{\mu}_{pHH})$ and $\widehat{var}(\hat{\mu}_{pHT})$ displayed in Web Tables A.3 and A.4.

Finally, to assess the accuracy of the MSE approximations, $MSE(\hat{\mu}_{pHH})$ and $MSE(\hat{\mu}_{pHT})$ are computed and compared to the true MSE. Results of relative error of the MSE first order approximations are presented in Tables A.5 and A.6.

[Figure A.1 about here.]

[Figure A.2 about here.]

[Figure A.3 about here.]

[Figure A.4 about here.]

[Figure A.5 about here.]

[Figure A.6 about here.]

[Table A.1 about here.]

[Table A.2 about here.]

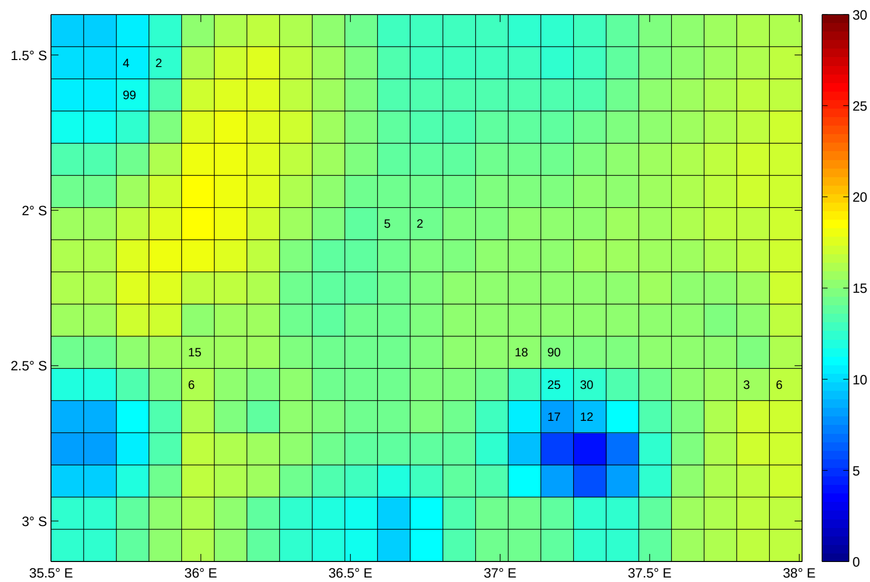
[Table A.3 about here.]

[Table A.4 about here.]

[Table A.5 about here.]

[Table A.6 about here.]

FIGURES



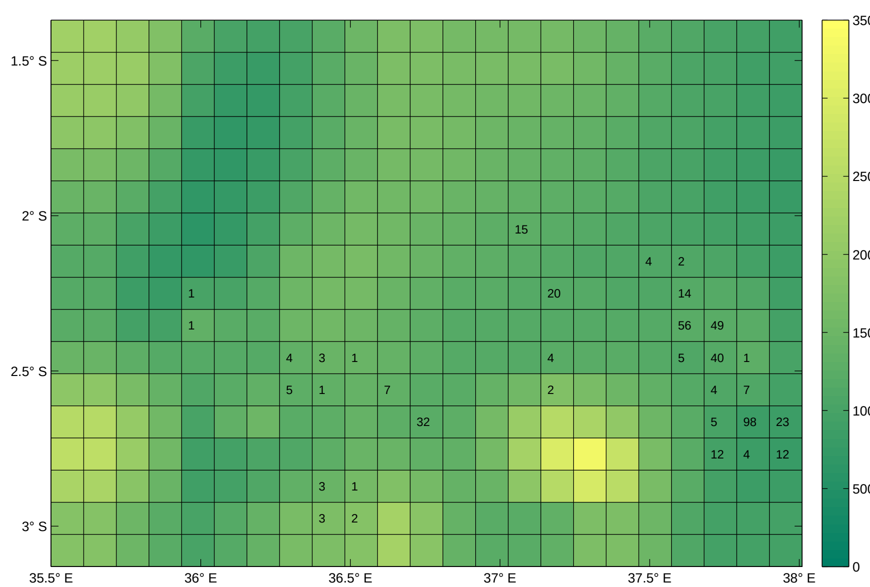


Figure 1. *Top panel*: values of Minimum temperature over the study area and Buffalo abundances. *Bottom panel*: values of Altitude over the study area and Hartebeest abundances.

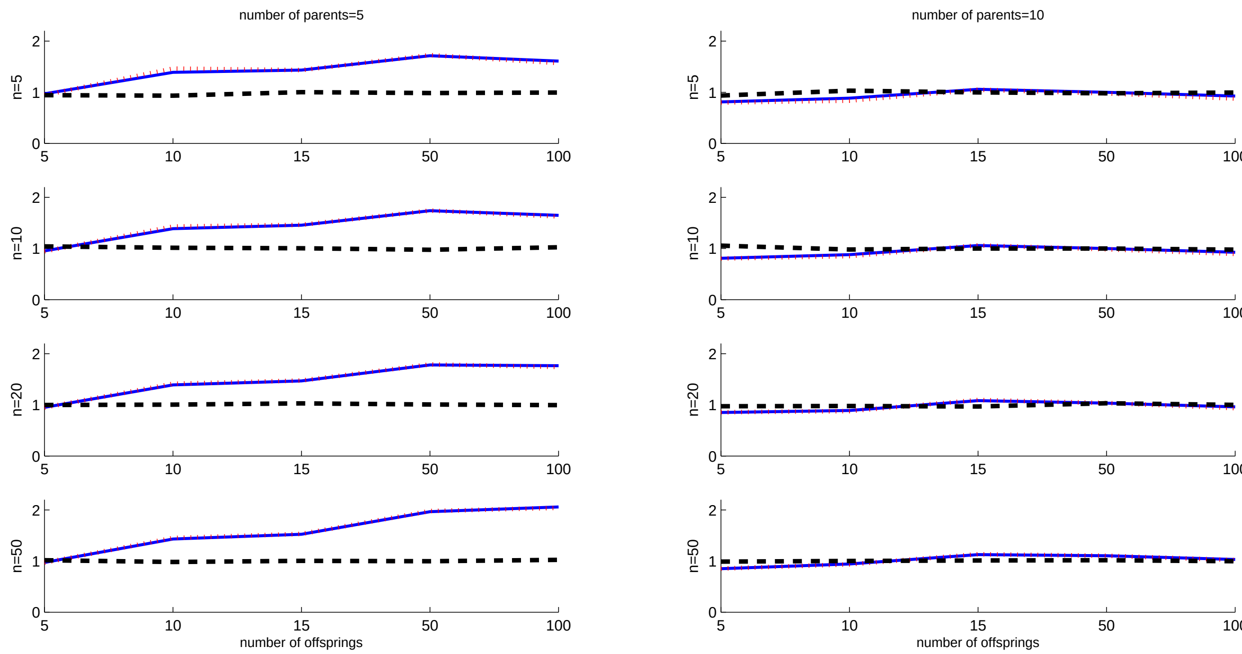


Figure A.1. Relative efficiency of $\hat{\mu}_{pHH}$, $\hat{\mu}_{HH}$ and $\hat{\mu}_p$. Network level correlation: low.

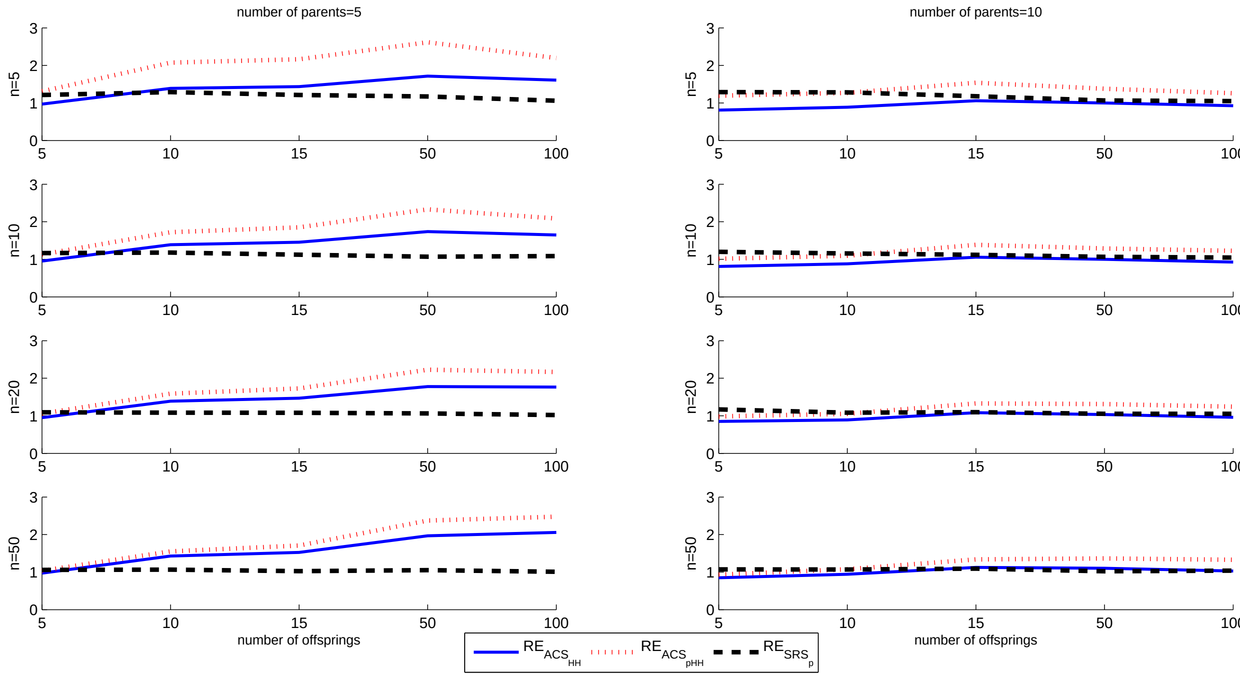


Figure A.2. Relative efficiency of $\hat{\mu}_{pHH}$, $\hat{\mu}_{HH}$ and $\hat{\mu}_p$. Network level correlation: intermediate.

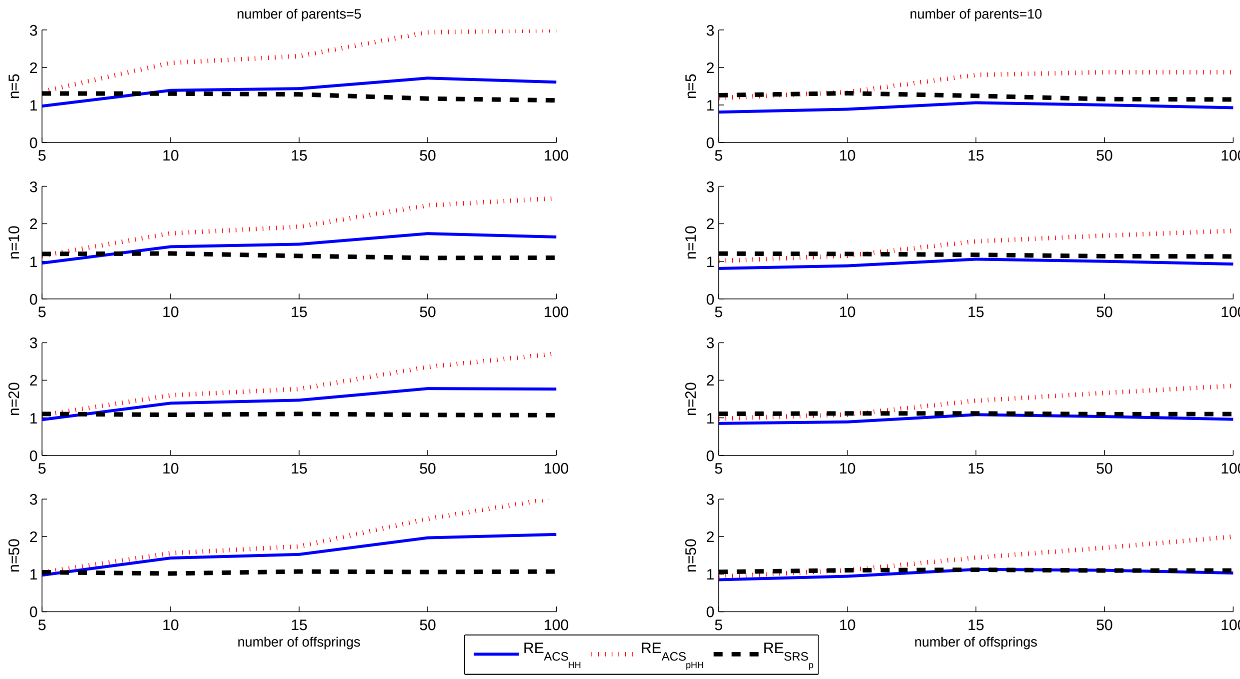


Figure A.3. Relative efficiency of $\hat{\mu}_{pHH}$, $\hat{\mu}_{HH}$ and $\hat{\mu}_p$. Network level correlation: high.

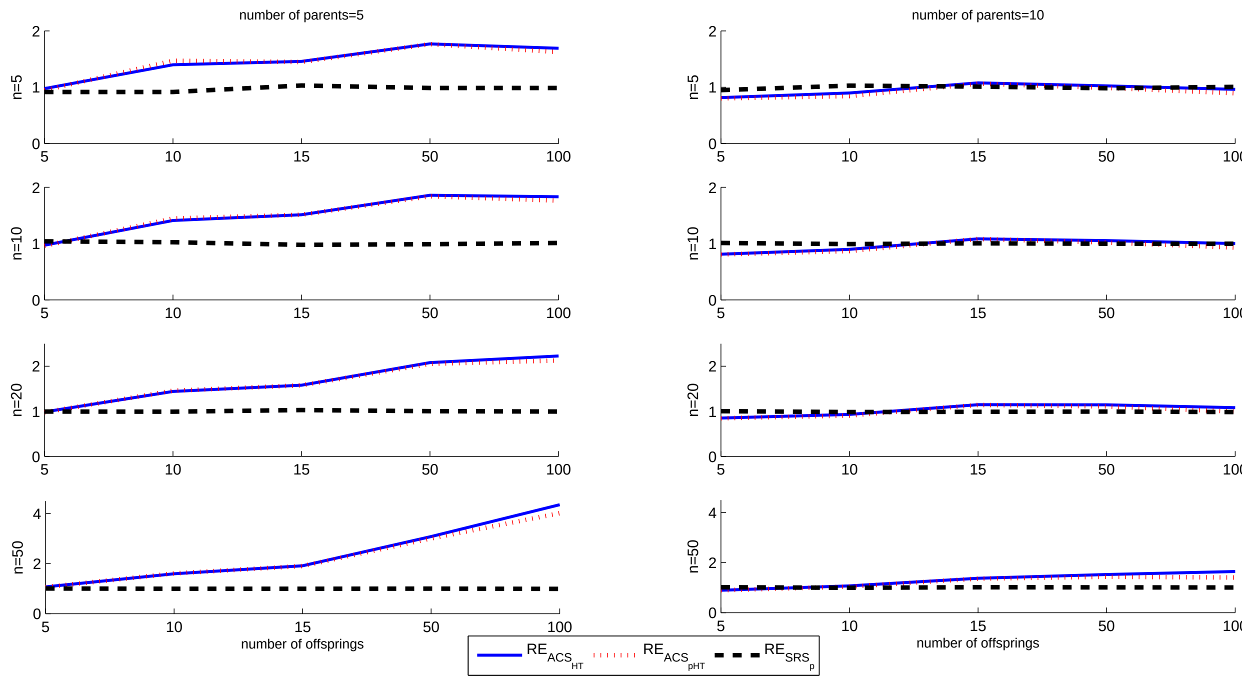


Figure A.4. Relative efficiency of $\hat{\mu}_{pHT}$, $\hat{\mu}_{HT}$ and $\hat{\mu}_p$. Network level correlation: low.

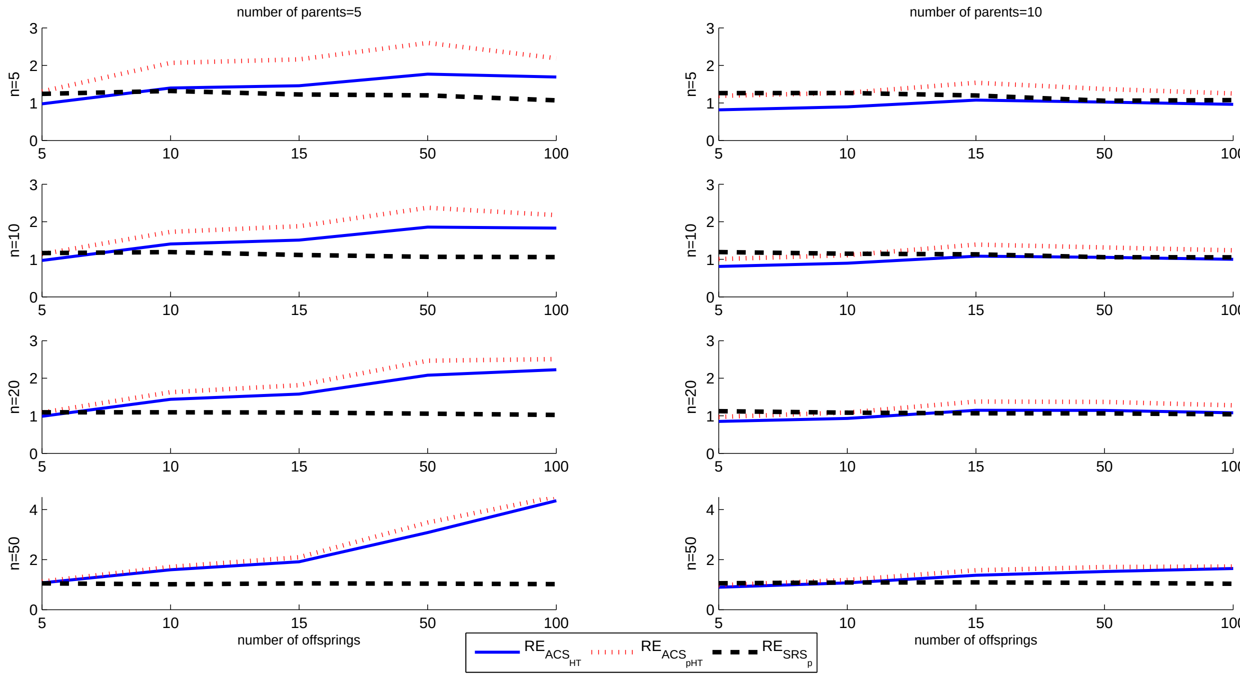


Figure A.5. Relative efficiency of $\hat{\mu}_{pHT}$, $\hat{\mu}_{HT}$ and $\hat{\mu}_p$. Network level correlation: intermediate.

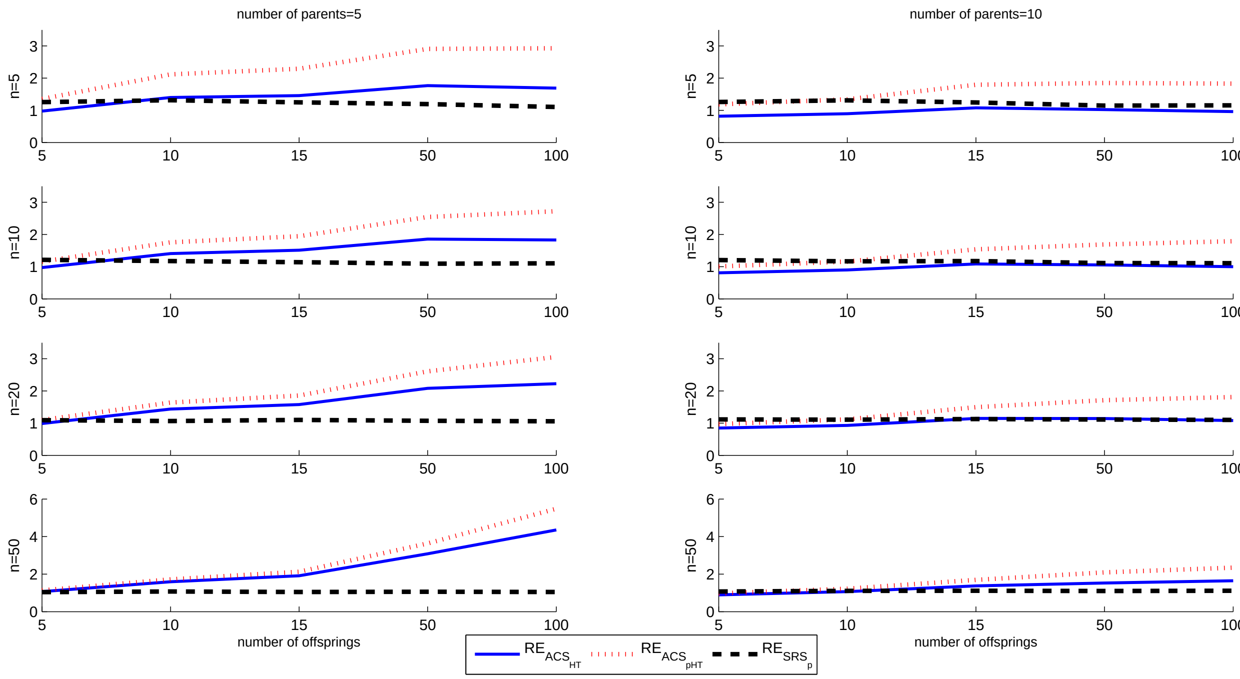


Figure A.6. Relative efficiency of $\hat{\mu}_{pHT}$, $\hat{\mu}_{HT}$ and $\hat{\mu}_{p}$. Network level correlation: high.

TABLES

Unit i	Network i	α_k	y_i	w_{yk}	y_k^*	x_i	w_{xk}	x_k^*
1	1	0.4	1	1	1	7	7	7
2	2	0.4	0	0	0	8	8	8
3	3	0.4	2	2	2	6	6	6
4	4	0.7	10	505	1010	1	3	6
5	4	0.7	1000	505	1010	5	3	6

Table 1. y and x populations and quantities to compute the ACS estimators

Samples	$\bar{y}$	$\hat{\mu}_{HHy}$	$\hat{\mu}_{HTy}$	$\bar{x}$	$\hat{\mu}_{HHx}$	$\hat{\mu}_{HTx}$	$\hat{\mu}_{p}$	$\hat{\mu}_{pHH}$	$\hat{\mu}_{pHT}$
1, 2	0.50	0.50	0.50	7.50	7.50	7.50	0.69	0.69	0.69
1, 3	1.50	1.50	1.50	6.50	6.50	6.50	1.81	1.81	1.81
1, 4; 3, 5	5.50	253.00	289.07	4.00	5.00	5.21	4.07	234.26	279.13
1, 5; 4, 2	500.50	253.00	289.07	6.00	5.00	5.21	556.11	234.26	279.13
2, 3	1.00	1.00	1.00	7.00	7.00	7.00	1.30	1.30	1.30
2, 4; 3, 5	5.00	252.50	288.57	4.50	5.50	5.71	4.17	257.18	305.37
2, 5; 4, 3	500.00	252.50	288.57	6.50	5.50	5.71	601.85	257.18	305.37
3, 4; 5	6.00	253.50	289.57	3.50	4.50	4.71	3.89	211.25	252.80
3, 5; 4	501.00	253.50	289.57	5.50	4.50	4.71	510.28	211.25	252.80
4, 5; 3	505.00	505.00	288.57	3.00	3.00	1.71	280.56	280.56	91.61
Mean	202.60	202.60	202.60	5.40	5.40	5.40	196.47	168.97	177.00
Bias	0.00	0.00	0.00	0.00	0.00	0.00	-6.13	-33.63	-25.60
MSE	59615.30	22861.59	17418.40	2.19	1.60	2.30	62499.67	13579.58	17180.20

Table 2. All possible adaptive cluster samples and the corresponding estimators

$Variables$	μ	C	C_w	ρ_{xy}	$\rho_{w_{xy}}$
Buffalo	0.8542	8.4361	5.8740	-0.087	-0.1845
Minimum Temperature	14.4904	0.1529	0.1512		
Hartebeest	1.1279	6.2215	4.0991	-0.0973	-0.1187
Altitude	1.3561	0.3041	0.3026		

Table 3. Mean (μ), Coefficient of variation at unit level (C) and at network level (C_w), correlation at unit level (ρ_{xy}) and at network level ($\rho_{w_{xy}}$).

	n	Relative efficiency $\hat{\mu}_p$	$\hat{\mu}_{pHH}$	$\hat{\mu}_{pHT}$	$\hat{\mu}_{HH}$	$\hat{\mu}_{HT}$	Relative bias $\hat{\mu}_{pHH}$	$\hat{\mu}_{pHT}$	$\widehat{var}(\hat{\mu}_{pHH})$	$\widehat{var}(\hat{\mu}_{pHT})$
Buffalo	5	1.05	1.62	1.64	1.48	1.51	−0.036	−0.004	−0.012	−0.012
	10	0.99	1.56	1.62	1.47	1.54	−0.016	−0.015	−0.003	−0.007
	20	0.99	1.54	1.70	1.48	1.65	−0.017	−0.018	−0.017	−0.015
	50	0.99	1.54	2.03	1.49	1.99	−0.001	−0.005	−0.018	−0.008
Hartebeest	5	1.03	1.30	1.29	1.20	1.20	−0.043	−0.038	0.015	−0.001
	10	0.99	1.23	1.25	1.19	1.21	−0.005	−0.001	0.003	−0.008
	20	1.05	1.25	1.33	1.22	1.30	−0.004	−0.005	0.010	0.008
	50	0.98	1.26	1.48	1.24	1.47	−0.011	−0.008	−0.001	−0.014

Table 4. Relative efficiency and relative bias of different estimators for Buffalo data (auxiliary variable is Minimum temperature) and Hartebeest data (auxiliary variable is Altitude)

	$\rho_{w_{xy}}$	λ_2	5	10	15	50	100
$\lambda_1 = 5$		n					
	low	5	-0.001	-0.004	-0.004	-0.018	-0.013
		10	-0.001	-0.002	-0.003	-0.007	-0.007
		20	0.000	-0.001	0.000	-0.002	0.002
		50	-0.001	-0.001	-0.001	0.000	0.001
	intermediate	5	-0.008	-0.019	-0.029	-0.107	-0.153
		10	-0.004	-0.009	-0.015	-0.051	-0.070
		20	-0.002	-0.004	-0.006	-0.021	-0.025
		50	-0.001	-0.002	-0.003	-0.006	-0.010
	high	5	-0.009	-0.019	-0.032	-0.117	-0.250
		10	-0.004	-0.009	-0.016	-0.055	-0.116
		20	-0.002	-0.004	-0.007	-0.023	-0.045
		50	-0.001	-0.002	-0.003	-0.006	-0.010
$\lambda_1 = 10$		n					
	low	5	-0.004	-0.008	-0.001	-0.023	-0.044
		10	-0.001	-0.002	-0.003	-0.006	-0.010
		20	-0.001	-0.001	-0.002	-0.002	0.000
		50	0.000	0.000	-0.001	-0.007	-0.004
	intermediate	5	-0.020	-0.037	-0.056	-0.147	-0.243
		10	-0.008	-0.016	-0.026	-0.068	-0.106
		20	-0.004	-0.008	-0.012	-0.029	-0.041
		50	-0.002	-0.002	-0.004	-0.014	-0.015
	high	5	-0.019	-0.038	-0.067	-0.233	-0.476
		10	-0.008	-0.017	-0.031	-0.109	-0.217
		20	-0.004	-0.008	-0.015	-0.048	-0.092
		50	-0.001	-0.003	-0.005	-0.018	-0.027

Table A.1. Relative bias of $\hat{\mu}_{pHT}$

	$\rho_{w_{xy}}$	λ_2	5	10	15	50	100
		n					
$\lambda_1 = 5$	low	5	-0.001	-0.004	-0.004	-0.018	-0.015
		10	-0.001	-0.002	-0.003	-0.011	-0.018
		20	0.000	-0.001	-0.001	-0.010	-0.010
		50	0.000	-0.001	-0.001	-0.001	-0.001
	intermediate	5	-0.008	-0.019	-0.030	-0.111	-0.162
		10	-0.004	-0.009	-0.016	-0.059	-0.088
		20	-0.002	-0.004	-0.007	-0.033	-0.044
		50	-0.001	-0.002	-0.004	-0.010	-0.015
	high	5	-0.009	-0.019	-0.032	-0.121	-0.264
		10	-0.004	-0.009	-0.017	-0.062	-0.139
		20	-0.002	-0.004	-0.008	-0.034	-0.069
		50	-0.001	-0.002	-0.004	-0.010	-0.024
	n						
$\lambda_1 = 10$	low	5	-0.004	-0.008	-0.010	-0.024	-0.046
		10	-0.001	-0.003	-0.003	-0.011	-0.021
		20	-0.002	-0.001	-0.003	-0.012	-0.013
		50	0.000	-0.001	-0.001	-0.009	-0.004
	intermediate	5	-0.020	-0.037	-0.057	-0.152	-0.253
		10	-0.009	-0.017	-0.027	-0.077	-0.128
		20	-0.005	-0.009	-0.014	-0.041	-0.060
		50	-0.002	-0.003	-0.005	-0.020	-0.022
	high	5	-0.019	-0.039	-0.068	-0.240	-0.496
		10	-0.009	-0.018	-0.032	-0.120	-0.245
		20	-0.005	-0.009	-0.017	-0.061	-0.119
		50	-0.001	-0.003	-0.006	-0.026	-0.044

Table A.2. Relative bias of $\hat{\mu}_{pHH}$

	$\rho_{w_{xy}}$	λ_2	5	10	15	50	100
		n					
$\lambda_1 = 5$	low	5	0.00	0.00	0.012	0.020	0.018
		10	-0.010	0.00	0.005	0.001	0.00
		20	-0.010	0.00	0.00	-0.001	0.008
		50	-0.010	0.020	0.007	0.001	-0.002
	intermediate	5	-0.07	-0.012	-0.106	-0.102	-0.041
		10	-0.030	-0.040	-0.049	-0.060	-0.044
		20	-0.020	-0.020	-0.029	-0.043	-0.028
		50	-0.020	0.00	-0.014	-0.022	-0.030
	high	5	-0.080	-0.012	-0.147	-0.184	-0.165
		10	-0.040	-0.040	-0.061	-0.094	-0.136
		20	-0.020	-0.020	-0.035	-0.058	-0.080
		50	-0.020	0.00	-0.017	-0.039	-0.062
		n					
$\lambda_1 = 10$	low	5	-0.006	0.013	0.000	0.010	0.020
		10	-0.007	0.011	0.001	-0.003	0.010
		20	0.006	-0.002	0.005	0.003	-0.006
		50	0.025	0.024	0.029	-0.008	-0.003
	intermediate	5	-0.100	-0.074	-0.072	-0.043	-0.021
		10	-0.042	-0.037	-0.036	-0.032	-0.021
		20	-0.008	-0.022	-0.020	-0.019	-0.021
		50	0.013	0.006	0.008	-0.0029	-0.016
	high	5	-0.093	-0.115	-0.158	-0.133	0.015
		10	-0.039	-0.045	-0.079	-0.114	-0.096
		20	-0.008	-0.027	-0.038	-0.064	-0.078
		50	0.001	0.006	0.003	-0.053	-0.038

Table A.3. Relative bias of the variance estimator $\widehat{var}(\hat{\mu}_{pHT})$

	$\rho_{w_{xy}}$	λ_2	5	10	15	50	100
		n					
$\lambda_1 = 5$	low	5	0.001	0.001	0.013	0.021	0.011
		10	-0.005	0.003	0.003	-0.002	-0.004
		20	0.009	-0.002	0.002	-0.014	0.011
		50	-0.015	0.017	-0.028	0.001	-0.019
	intermediate	5	-0.062	-0.110	-0.099	-0.086	-0.031
		10	-0.029	-0.035	-0.041	-0.039	-0.028
		20	-0.015	-0.013	-0.017	-0.033	0.006
		50	-0.018	0.007	-0.025	-0.006	-0.023
	high	5	-0.075	-0.119	-0.139	-0.169	-0.127
		10	-0.033	-0.035	-0.053	-0.080	-0.108
		20	-0.018	-0.014	-0.023	-0.049	-0.032
		50	-0.018	0.006	-0.027	-0.016	-0.042
	n						
$\lambda_1 = 10$	low	5	-0.008	0.010	-0.005	0.006	-0.002
		10	-0.001	-0.003	0.003	0.008	-0.021
		20	0.023	0.016	0.019	0.014	-0.016
		50	0.023	0.036	0.022	0.002	-0.003
	intermediate	5	-0.020	-0.037	-0.057	-0.152	-0.253
		10	-0.098	-0.071	-0.070	-0.042	-0.0.041
		20	0.008	-0.006	-0.007	0.002	0.003
		50	0.012	0.024	0.010	-0.006	-0.016
	high	5	-0.092	-0.112	-0.154	-0.120	0.022
		10	-0.031	-0.048	-0.067	-0.111	-0.016
		20	0.005	-0.009	-0.016	-0.041	-0.053
		50	0.011	0.024	0.006	-0.025	-0.034

Table A.4. Relative bias of the variance estimator $\widehat{var}(\hat{\mu}_{pHH})$

	$\rho_{w_{xy}}$	λ_2	5	10	15	50	100
		n					
$\lambda_1 = 5$	low	5	0.020	-0.044	0.017	0.003	0.019
		10	0.019	-0.018	0.003	0.014	0.02
		20	0.019	-0.004	0.013	0.012	0.010
		50	0.009	-0.026	-0.007	0.005	0.010
	intermediate	5	-0.230	-0.285	-0.255	-0.201	-0.095
		10	-0.106	-0.137	-0.120	-0.078	-0.015
		20	-0.043	-0.061	-0.040	-0.013	0.015
		50	-0.007	-0.035	-0.012	0.011	0.035
	high	5	-0.253	-0.299	-0.289	-0.256	-0.176
		10	-0.118	-0.146	-0.138	-0.104	-0.040
		20	-0.047	-0.064	-0.049	-0.024	0.018
		50	-0.008	-0.036	-0.012	0.021	0.058
		n					
$\lambda_1 = 10$	low	5	0.015	-0.048	0.012	-0.005	0.011
		10	0.016	-0.023	0.003	0.005	0.004
		20	0.017	-0.009	0.007	0.008	-0.010
		50	0.011	-0.025	0.015	0.002	0.019
	intermediate	5	-0.236	-0.292	-0.268	-0.229	-0.134
		10	-0.112	-0.146	-0.135	-0.119	-0.062
		20	-0.050	-0.073	-0.061	-0.049	-0.047
		50	-0.011	-0.041	-0.006	-0.020	0.007
	high	5	-0.259	-0.307	-0.301	-0.284	-0.225
		10	-0.124	-0.156	-0.154	-0.141	-0.110
		20	-0.055	-0.075	-0.070	-0.061	-0.070
		50	-0.014	-0.041	-0.009	-0.020	0.001

Table A.5. Relative error of the MSE first order approximation of $\hat{\mu}_{pHT}$

	$\rho_{w_{xy}}$	λ_2	5	10	15	50	100
		n					
$\lambda_1 = 5$	low	5	0.015	-0.048	0.012	-0.005	0.011
		10	0.016	-0.023	0.003	0.005	0.004
		20	0.017	-0.009	0.007	0.008	-0.010
		50	0.011	-0.025	0.015	0.002	0.019
	intermediate	5	-0.236	-0.292	-0.268	-0.229	-0.134
		10	-0.112	-0.146	-0.135	-0.119	-0.062
		20	-0.050	-0.073	-0.061	-0.049	-0.047
		50	-0.011	-0.041	-0.006	-0.020	0.007
	high	5	-0.259	-0.30	-0.289	-0.256	-0.176
		10	-0.118	-0.146	-0.138	-0.104	-0.040
		20	-0.047	-0.064	-0.049	-0.024	0.018
		50	-0.008	-0.036	-0.012	0.021	0.058
		n					
$\lambda_1 = 10$	low	5	0.036	0.049	0.013	0.031	0.044
		10	0.033	0.033	0.011	0.026	0.029
		20	-0.001	0.020	0.000	0.013	0.034
		50	-0.013	-0.016	-0.024	0.013	0.015
	intermediate	5	-0.252	-0.215	-0.182	-0.087	-0.033
		10	-0.114	-0.091	-0.083	-0.025	-0.005
		20	-0.067	-0.040	-0.035	-0.002	0.026
		50	-0.030	-0.025	-0.023	0.026	0.026
	high	5	-0.247	-0.241	-0.253	-0.158	-0.015
		10	-0.112	-0.111	-0.117	-0.058	0.014
		20	-0.066	-0.047	-0.050	-0.010	0.049
		50	-0.027	-0.029	-0.028	0.038	0.044

Table A.6. Relative error of the MSE first order approximation of $\hat{\mu}_{pHH}$