# A COMPARISON OF ESTIMATORS OF MEAN AND ITS FUNCTIONS IN FINITE POPULATION

Anurag Dey and Probal Chaudhuri

*Indian Statistical Institute, Kolkata*

**Supplementary Material**

In this supplement, we discuss conditions C1 through C4 from the main paper and demonstrate situations, where these conditions hold. Then, we state and prove some additional mathematical results. We also give the proofs of Remark 1 and Theorems 2, 3, 6 and 7 of the main text. The biased estimators considered in the main paper are then compared empirically with their bias-corrected versions based on jackknifing in terms of MSE. Finally, we provide the numerical results related to the analysis based on both synthetic and real data.

## S1 Discussion of conditions and related results

In this section, we demonstrate some situations, when conditions C1 through C4 in the main article hold. Before that we prove and state the following lemma. Recall from the paragraph following C2 in the main text that $\gamma=\sum_{i=1}^{n} N_i(N_i-1)/N(N-1)$ with $N_i$ being the size of the $i^{th}$ group formed randomly in RHC sampling design.

**Lemma S 1.** *Suppose that $C0$ holds. Then, $n\gamma \to c$ for some $c \geq 1-\lambda > 0$ as $\nu \to \infty$, where $\lambda$ is as in $C0$.*

*Proof.* Let us first consider the case of $\lambda$=0. Note that

$$\begin{aligned} & n(N/n-1)(N-n)/(N(N-1)) \leq n\gamma \leq \\ & n(N/n+1)(N-n)/(N(N-1)) \end{aligned} \tag{S1.1}$$

by (2.1) in Section 2 of the main text. Moreover, $n(N/n+1)(N-n)/(N(N-1))=(1+n/N)(N-n)/(N-1) \to 1$ and $n(N/n-1)(N-n)/(N(N-1))=(1-n/N)(N-n)/(N-1) \to 1$ as $\nu \to \infty$ because C0 holds and $\lambda$=0. Thus we have $n\gamma \to 1$ as $\nu \to \infty$ in this case.

Next, consider the case, when $\lambda > 0$ and $\lambda^{-1}$ is an integer. Here, we consider the following sub-cases. Let us first consider the sub-case, when $N/n$ is an integer for all sufficiently large $\nu$. Then, by (2.1), we have $n\gamma$=$(N-n)/(N-1)$ for all sufficiently large $\nu$. Now, since C0 holds, we have

$$(N-n)/(N-1) \to 1-\lambda \text{ as } \nu \to \infty. \tag{S1.2}$$

Further, consider the sub-case, when $N/n$ is a non-integer and $N/n - \lambda^{-1} \geq 0$ for all sufficiently large $\nu$. Then by (2.1) in Section 2 of the main text, we have

$$n\gamma = (N/(N-1))(n/N)\lfloor N/n\rfloor\big(2-\big((n/N)\lfloor N/n\rfloor\big)-(n/N)\big) \tag{S1.3}$$

for all sufficiently large $\nu$. Now, since C0 holds, we have $0 \leq N/n - \lambda^{-1} < 1$ for all sufficiently large $\nu$. Then, $\lfloor N/n \rfloor = \lambda^{-1}$ for all sufficiently large $\nu$, and hence

$$(N/(N-1))(n/N)\lfloor N/n \rfloor \left(2 - \left((n/N)\lfloor N/n \rfloor\right) - (n/N)\right) \to 1 - \lambda \quad \text{(S1.4)}$$

as $\nu \to \infty$.

Next, consider the sub-case, when $N/n$ is a non-integer and $N/n - \lambda^{-1} < 0$ for all sufficiently large $\nu$. Then, the result in (S1.3) holds by (2.1), and $-1 \leq N/n - \lambda^{-1} < 0$ for all sufficiently large $\nu$ by C0. Therefore, $\lfloor N/n \rfloor = \lambda^{-1} - 1$ for all sufficiently large $\nu$, and hence the result in (S1.4) holds. Thus, in the case of $\lambda > 0$ and $\lambda^{-1}$ being an integer, $n\gamma$ converges to $1 - \lambda$ as $\nu \to \infty$ through all the sub-sequences, and hence $n\gamma \to 1 - \lambda$ as $\nu \to \infty$. Thus we have $c = 1 - \lambda$ in this case.

Finally, consider the case, when $\lambda > 0$, and $\lambda^{-1}$ is a non-integer. Then, $N/n$ must be a non-integer for all sufficiently large $\nu$, and hence $n\gamma = (N/(N-1))(n/N)\lfloor N/n \rfloor \big(2 - \big((n/N)\lfloor N/n \rfloor\big) - (n/N)\big)$ for all sufficiently large $\nu$ by (2.1) in Section 2 of the main text. Note that in this case, $N/n - \lfloor \lambda^{-1} \rfloor \to \lambda^{-1} - \lfloor \lambda^{-1} \rfloor \in (0,1)$ as $\nu \to \infty$ by C0. Therefore, $\lfloor \lambda^{-1} \rfloor < N/n < \lfloor \lambda^{-1} \rfloor + 1$ for all sufficiently large $\nu$, and hence $\lfloor N/n \rfloor = \lfloor \lambda^{-1} \rfloor$ for all sufficiently large $\nu$. Thus $n\gamma \to \lambda \lfloor \lambda^{-1} \rfloor (2 - \lambda \lfloor \lambda^{-1} \rfloor - \lambda)$ as $\nu \to \infty$ by C0. Now, if $m = \lfloor \lambda^{-1} \rfloor$ and $\lambda^{-1}$ is a non-integer, then

$(m+1)^{-1} < \lambda < m^{-1}$. Therefore, $\lambda\lfloor\lambda^{-1}\rfloor(2-\lambda\lfloor\lambda^{-1}\rfloor-\lambda)-1+\lambda{=}-\big(1-(2m+1)\lambda+m(m+1)\lambda^2\big){=}-(1-m\lambda)(1-(m+1)\lambda)>0$. Thus we have $c{=}\lambda\lfloor\lambda^{-1}\rfloor(2-\lambda\lfloor\lambda^{-1}\rfloor-\lambda)>1-\lambda$ in this case. This completes the proof of the Lemma. □

Next, recall $\{\mathbf{V}_i\}_{i=1}^N$ from the paragraph preceding the condition C3 and $b$ from the condition C5 in the main text. Let us define $\Sigma_1{=}nN^{-2}\sum_{i=1}^N(\mathbf{V}_i-\mathbf{T}\pi_i)^T(\mathbf{V}_i-\mathbf{T}\pi_i)(\pi_i^{-1}-1)$ and $\Sigma_2{=}n\gamma\overline{X}N^{-1}\sum_{i=1}^N(\mathbf{V}_i-X_i\overline{\mathbf{V}}/\overline{X})^T(\mathbf{V}_i-X_i\overline{\mathbf{V}}/\overline{X})/X_i$, where $\mathbf{T}{=}\sum_{i=1}^N\mathbf{V}_i(1-\pi_i)/\sum_{i=1}^N\pi_i(1-\pi_i)$, the $\pi_i$'s are inclusion probabilities and $\overline{\mathbf{V}}{=}\sum_{i=1}^N\mathbf{V}_i/N$. Now, we state the following lemma.

**Lemma S 2.** ($i$) *Suppose that* $C0$ *and* $C5$ *hold, and* $\{(h(Y_i),X_i):1\leq i\leq N\}$ *are generated from a superpopulation distribution* $\mathbb{P}$ *with* $E_{\mathbb{P}}||h(Y_i)||^4<\infty$. *Then,* $C1$, $C2$ *and* $C4$ *hold a.s.* $[\mathbb{P}]$.

($ii$) *Further, if* $C0$ *and* $C5$ *hold, and* $E_{\mathbb{P}}||h(Y_i)||^2<\infty$, *then* $C3$ *holds a.s.* $[\mathbb{P}]$ *under SRSWOR and LMS sampling design. Moreover, if* $C0$ *holds with* $0\leq\lambda<E_{\mathbb{P}}(X_i)/b$, $C5$ *holds, and* $E_{\mathbb{P}}||h(Y_i)||^2<\infty$, *then* $C3$ *holds a.s.* $[\mathbb{P}]$ *under any* $\pi PS$ *sampling design.*

*Proof.* As before, for simplicity, let us write $h(Y_i)$ as $h_i$. Under the conditions C5 and $E_{\mathbb{P}}||h(Y_i)||^4<\infty$, C1 holds *a.s.* $[\mathbb{P}]$ by SLLN. Also, under C5, C2 holds *a.s.* $[\mathbb{P}]$. Next, by SLLN, $\lim_{\nu\to\infty}\Sigma_2{=}cE_{\mathbb{P}}(X_i)E_{\mathbb{P}}[(h_i-(E_{\mathbb{P}}(X_i))^{-1}X_i\,E_{\mathbb{P}}(h_i))^T(h_i-(E_{\mathbb{P}}(X_i))^{-1}X_iE_{\mathbb{P}}(h_i))\,X_i^{-1}]$ *a.s.* $[\mathbb{P}]$ for $\mathbf{V}_i{=}h_i$,

$h_i - \overline{h}X_i/\overline{X}$ and $h_i + \overline{h}X_i/\overline{X}$ because $n\gamma \to c$ as $\nu \to \infty$ by Lemma S1. Similarly, $\lim_{\nu\to\infty} \Sigma_2 = cE_{\mathbb{P}}(X_i)E_{\mathbb{P}}[(h_i - E_{\mathbb{P}}(h_i))^T(h_i - E_{\mathbb{P}}(h_i))/X_i]$ *a.s.* $[\mathbb{P}]$ for $\mathbf{V}_i = h_i - \overline{h}$, and $\lim_{\nu\to\infty} \Sigma_2 = cE_{\mathbb{P}}(X_i)E_{\mathbb{P}}[\,(h_i - E_{\mathbb{P}}(h_i) - C_{xh}(X_i - E_{\mathbb{P}}(X_i)))^T(h_i - E_{\mathbb{P}}(h_i) - C_{xh}(X_i - E_{\mathbb{P}}(X_i)))/X_i]$ *a.s.* $[\mathbb{P}]$ for $\mathbf{V}_i = h_i - \overline{h} - S_{xh}(X_i - \overline{X})/S_x^2$. Here, $C_{xh} = (E_{\mathbb{P}}(h_iX_i) - E_{\mathbb{P}}(h_i)E_{\mathbb{P}}(X_i))/\,(E_{\mathbb{P}}(X_i)^2 - (E_{\mathbb{P}}(X_i))^2)$. Note that the above limits are p.d. matrices because C5 holds. Therefore, C4 holds *a.s.* $[\mathbb{P}]$. This completes the proof of *(i)* in Lemma S2.

Next, note that $\Sigma_1 = (1 - n/N)(\sum_{i=1}^N \mathbf{V}_i^T\mathbf{V}_i/N - \overline{\mathbf{V}}^T\overline{\mathbf{V}})$ under SRSWOR. Then, C3 holds *a.s.* $[\mathbb{P}]$ by directly applying SLLN. Under LMS sampling design, C3 can be shown to hold *a.s.* $[\mathbb{P}]$ in the same way as the proof of the result $\sigma_1^2 = \sigma_2^2$ in the proof of Lemma 2 in the Appendix. Next, we have $\lim_{\nu\to\infty} \Sigma_1 = E_{\mathbb{P}}\big[\big\{h_i + \chi^{-1}(E_{\mathbb{P}}(X_i))^{-1}X_i\big(\lambda E_{\mathbb{P}}(h_iX_i) - E_{\mathbb{P}}(h_i)E_{\mathbb{P}}(X_i)\big)\big\}^T\big\{h_i + \chi^{-1}(E_{\mathbb{P}}(X_i))^{-1}X_i\big(\lambda E_{\mathbb{P}}(h_iX_i) - E_{\mathbb{P}}(h_i)E_{\mathbb{P}}(X_i)\big)\big\}\big\{E_{\mathbb{P}}(X_i)/X_i - \lambda\big\}\big]$ *a.s.* $[\mathbb{P}]$ for $\mathbf{V}_i = h_i$, $h_i - \overline{h}X_i/\overline{X}$ and $h_i + \overline{h}X_i/\overline{X}$ under any $\pi$PS sampling design (i.e., a sampling design with $\pi_i = nX_i/\sum_{i=1}^N X_i$) by SLLN because C0 and C5 hold, and $E_{\mathbb{P}}||h_i||^2 < \infty$. Here, $\chi = E_{\mathbb{P}}(X_i) - \lambda(E_{\mathbb{P}}(X_i)^2/E_{\mathbb{P}}(X_i))$. Moreover, under any $\pi$PS sampling design, we have $\lim_{\nu\to\infty} \Sigma_1 = E_{\mathbb{P}}\big[\big\{h_i - E_{\mathbb{P}}(h_i) + \lambda\chi^{-1}(E_{\mathbb{P}}(X_i))^{-1}X_iC_{xh}\big\}^T\big\{h_i - E_{\mathbb{P}}(h_i) + \lambda\chi^{-1}(E_{\mathbb{P}}(X_i))^{-1}X_iC_{xh}\big\}\times\big\{E_{\mathbb{P}}(X_i)/X_i - \lambda\big\}\big]$ *a.s.* $[\mathbb{P}]$ for $\mathbf{V}_i = h_i - \overline{h}$ and $\lim_{\nu\to\infty} \Sigma_1 = E_{\mathbb{P}}\big[\big\{h_i - E_{\mathbb{P}}(h_i) - C_{xh}(X_i - E_{\mathbb{P}}(X_i))\big\}^T\big\{h_i - E_{\mathbb{P}}(h_i) - C_{xh}(X_i - E_{\mathbb{P}}(X_i))\big\}\big\{E_{\mathbb{P}}(X_i)/X_i - \lambda\big\}\big]$

*a.s.* [$\mathbb{P}$] for $\mathbf{V}_i = h_i - \overline{h} - S_{xh}(X_i - \overline{X})/S_x^2$. Note that the above limits are p.d. matrices because C5 holds and C0 holds with $0 \leq \lambda < E_{\mathbb{P}}(X_i)/b$. Therefore, C3 holds *a.s.* [$\mathbb{P}$] under any $\pi$PS sampling design. This completes the proof of *(ii)* in Lemma S2. □

# S2 Additional mathematical details

In this section, we state and prove some technical results, which will be required to prove the theorems stated in the main text.

**Lemma S 3.** *Suppose that C2 holds. Then, LMS sampling design is a high entropy sampling design. Moreover, under LMS sampling design, there exist constants $L, L' > 0$ such that*

$$L \leq \min_{1 \leq i \leq N}(N\pi_i/n) \leq \max_{1 \leq i \leq N}(N\pi_i/n) \leq L' \tag{S2.1}$$

*for all sufficiently large $\nu$ .*

The condition (S2.1) was considered earlier in Wang and Opsomer (2011), Boistard et al. (2017), etc. However, the above authors did not discuss whether LMS sampling design satisfies (S2.1) or not.

*Proof.* Suppose that $P(s)$ and $R(s)$ denote LMS sampling design and SRSWOR, respectively. Note that SRSWOR is a rejective sampling design. Then, $P(s) = (\overline{x}/\overline{X})/{}^N C_n$ and $R(s) = ({}^N C_n)^{-1}$, where $\overline{x} = \sum_{i \in s} X_i/n$ and $s \in$

$\mathcal{S}$. By Cauchy-Schwarz inequality, we have $D(P||R)$=$E_R((\overline{x}/\overline{X})\log(\overline{x}/\overline{X})) \leq K_1 E_R|\overline{x}/\overline{X}-1| \leq K_1 E_R(\overline{x}/\overline{X}-1)^2$ for some $K_1>0$ since C2 holds, and $\log(x) \leq |x-1|$ for $x>0$. Here $E_R$ denotes the expectation with respect to $R(s)$. Therefore, $nD(P||R) \leq K_1(1-f)(N/(N-1))(S_x^2/\overline{X}^2) \leq 2K_1(\sum_{i=1}^N X_i^2/N\overline{X}^2) \leq 2K_1(\max_{1\leq i\leq N} X_i/\min_{1\leq i\leq N} X_i)^2$=$O(1)$ as $\nu \to \infty$, where $f$=$n/N$. Hence, $D(P||R) \to 0$ as $\nu \to \infty$. Thus LMS sampling design is a high entropy sampling design.

Next, suppose that $\{\pi_i\}_{i=1}^N$ denote inclusion probabilities of $P(s)$. Then, we have $\pi_i$=$(n-1)/(N-1)+(X_i/\sum_{i=1}^N X_i)((N-n)/(N-1))$ and $\pi_i - n/N$=$-(N-n)(N(N-1))^{-1}(X_i/\overline{X}-1)$. Further,

$$\frac{|\pi_i - n/N|}{n/N} = \frac{N-n}{n(N-1)}\left|\frac{X_i}{\overline{X}}-1\right| \leq \frac{N-n}{n(N-1)}\left(\frac{\max_{1\leq i\leq N} X_i}{\min_{1\leq i\leq N} X_i}+1\right).$$

Therefore, $\max_{1\leq i\leq N}|N\pi_i/n-1| \to 0$ as $\nu \to \infty$ by C2. Hence, $K_2 \leq \min_{1\leq i\leq N}(N\pi_i/n) \leq \max_{1\leq i\leq N}(N\pi_i/n) \leq K_3$ for all sufficiently large $\nu$ and some constants $K_2>0$ and $K_3>0$. Thus (S2.1) holds under LMS sampling design. $\square$

Next, suppose that $\{\mathbf{V}_i\}_{i=1}^N$, $\overline{\mathbf{V}}$, $\Sigma_1$ and $\Sigma_2$ are as in the previous Section S1. Let us define $\hat{\overline{V}}_1$= $\sum_{i\in s}(N\pi_i)^{-1}V_i$ and $\hat{\overline{V}}_2$=$\sum_{i\in s} G_i V_i/NX_i$, where $G_i$'s are as in the paragraph containing Table 8 in the main article. Now, we state the following lemma.

**Lemma S 4.** *Suppose that C0 through C3 hold. Then, under SRSWOR, LMS sampling design and any HE$\pi$PS sampling design, we have* $\sqrt{n}(\hat{\overline{\mathbf{V}}}_1 - \overline{\mathbf{V}}) \xrightarrow{\mathcal{L}} N(0,\Gamma_1)$ *as* $\nu \to \infty$*, where* $\Gamma_1$=$\lim_{\nu\to\infty}\Sigma_1$. *Further, suppose that C0 through C2 and C4 hold. Then, we have* $\sqrt{n}(\hat{\overline{\mathbf{V}}}_2 - \overline{\mathbf{V}}) \xrightarrow{\mathcal{L}} N(0,\Gamma_2)$ *as* $\nu \to \infty$ *under RHC sampling, where* $\Gamma_2$=$\lim_{\nu\to\infty}\Sigma_2$.

*Proof.* Note that SRSWOR is a high entropy sampling design since it is a rejective sampling design. Also, (S2.1) in Lemma S3 holds trivially under SRSWOR. It follows from Lemma S3 that LMS sampling design is a high entropy sampling design, and (S2.1) holds under this sampling design. Further, any HE$\pi$PS sampling design satisfies (S2.1) since C2 holds. Now, fix $\epsilon > 0$ and $\mathbf{m} \in \mathbb{R}^p$. Suppose that $L(\epsilon,\mathbf{m})$=$(n^{-1}N^2\mathbf{m}\Sigma_1\mathbf{m}^T)^{-1}\sum_{i\in G(\epsilon,\mathbf{m})}(\mathbf{m}$ $(\mathbf{V}_i - \mathbf{T}\pi_i)^T)^2(\pi_i^{-1} - 1)$ for $G(\epsilon,\mathbf{m})$=$\{1 \le i \le N : |\mathbf{m}(\mathbf{V}_i - \mathbf{T}\pi_i)^T| > \epsilon\pi_i N(n^{-1}\,\mathbf{m}\Sigma_1\mathbf{m}^T)^{1/2}\}$, $\mathbf{T}$=$\sum_{i=1}^N \mathbf{V}_i(1-\pi_i)/\sum_{i=1}^N \pi_i(1-\pi_i)$ and $\mathbf{Z}_i$=$(n/N\pi_i)\mathbf{V}_i$ $-(n/N)\mathbf{T}$, $i$=$1,\ldots,N$. Then, given any $\eta > 0$, $L(\epsilon,\mathbf{m}) \le (\mathbf{m}\Sigma_1\mathbf{m}^T)^{-(1+\eta/2)}$ $n^{-\eta/2}\epsilon^{-\eta}N^{-1}\sum_{i=1}^N(||\mathbf{m}||||\mathbf{Z}_i||)^{2+\eta}(N\pi_i/n)$ since $|\mathbf{m}\mathbf{Z}_i^T|/(\sqrt{n}\epsilon(\mathbf{m}\Sigma_1\mathbf{m}^T)^{1/2}) >$ 1 for any $i \in G(\epsilon,\mathbf{m})$. It follows from Jensen's inequality that $N^{-1}\sum_{i=1}^N ||\mathbf{Z}_i||^{2+\eta}$ $(N\pi_i/n) \le 2^{1+\eta}(N^{-1}\sum_{i=1}^N ||\mathbf{V}_i(n/N\pi_i)||^{2+\eta}\,(N\pi_i/n) + ||(n/N)\mathbf{T}||^{2+\eta})$ since $\sum_{i=1}^N \pi_i$=$n$. It also follows from C1, C2 and Jensen's inequality that $\sum_{i=1}^N ||\mathbf{V}_i||^{2+\eta}$ $/N$=$O(1)$ as $\nu \to \infty$ for any $0 < \eta \le 2$. Further, $\sum_{i=1}^N \pi_i(1-\pi_i)/n$ is bounded away from 0 as $\nu \to \infty$ under SRSWOR, LMS sampling design and any HE$\pi$PS sampling design because (S2.1) holds under these sampling designs, and C0 holds. Therefore, $N^{-1}\sum_{i=1}^{N}||\mathbf{V}_i(n/N\pi_i)||^{2+\eta}(N\pi_i/n)=O(1)$ and $||(n/N)\mathbf{T}||^{2+\eta}=O(1)$, and hence $N^{-1}\sum_{i=1}^{N}||\mathbf{Z}_i||^{2+\eta}(N\pi_i/n)=O(1)$ as $\nu\to\infty$ under the above sampling designs. Then, $L(\epsilon,\mathbf{m})\to 0$ as $\nu\to\infty$ for any $\epsilon>0$ under all of these sampling designs since C3 holds. Therefore, $\inf\{\epsilon>0: L(\epsilon,\mathbf{m})\leq\epsilon\}\to 0$ as $\nu\to\infty$, and consequently the Hájek-Lindeberg condition holds for $\{\mathbf{m}\mathbf{V}_i^T\}_{i=1}^{N}$ under each of the above sampling designs. Also, $\sum_{i=1}^{N}\pi_i(1-\pi_i)\to\infty$ as $\nu\to\infty$ under these sampling designs. Then, from Theorem 5 in Berger (1998), $\sqrt{n}\mathbf{m}(\hat{\overline{\mathbf{V}}}_1-\overline{\mathbf{V}})^T\xrightarrow{\mathcal{L}}N(0,\mathbf{m}\Gamma_1\mathbf{m}^T)$ as $\nu\to\infty$ under each of the above sampling designs for any $\mathbf{m}\in\mathbb{R}^p$ and $\Gamma_1=\lim_{\nu\to\infty}\Sigma_1$. Hence, $\sqrt{n}(\hat{\overline{\mathbf{V}}}_1-\overline{\mathbf{V}})\xrightarrow{\mathcal{L}}N(0,\Gamma_1)$ as $\nu\to\infty$ under the above-mentioned sampling designs.

Next, define $L(\mathbf{m})=n\gamma(\max_{1\leq i\leq N}X_i)(N^{-1}\sum_{i=1}^{n}N_i^3(N_i-1)\sum_{i=1}^{N}(\mathbf{m}(\mathbf{V}_i\overline{X}/X_i-\overline{\mathbf{V}})^T)^4X_i)^{1/2}(\overline{X}^{3/2}\sum_{i=1}^{n}N_i(N_i-1)\mathbf{m}\Sigma_2\mathbf{m}^T)^{-1}$, where $\gamma=\sum_{i=1}^{n}N_i(N_i-1)/N(N-1)$ as before. Note that as $\nu\to\infty$, $(N^{-1}\sum_{i=1}^{N}(\mathbf{m}(\mathbf{V}_i\overline{X}/X_i-\overline{\mathbf{V}})^T)^4(X_i/\overline{X}))^{1/2}=O(1)$ and $(\max_{1\leq i\leq N}X_i)/\overline{X}=O(1)$ since C1 and C2 hold. Now, recall from Section 2 in the main text that the $N_i$'s are considered as in (2.1). Then, under C0, we have $(\sum_{i=1}^{n}N_i^3(N_i-1))^{1/2}(\sum_{i=1}^{n}N_i(N_i-1))^{-1}=O(1/\sqrt{n})$ and $n\gamma=O(1)$ as $\nu\to\infty$. Therefore, $L(\mathbf{m})\to 0$ as $\nu\to\infty$ since C4 holds. This implies that condition C1 in Ohlsson (1986) holds for $\{\mathbf{m}\mathbf{V}_i^T\}_{i=1}^N$. Therefore, by Theorem 2.1 in Ohlsson (1986), $\sqrt{n}\mathbf{m}(\hat{\overline{\mathbf{V}}}_2 - \overline{\mathbf{V}})^T \xrightarrow{\mathcal{L}} N(0, \mathbf{m}\Gamma_2\mathbf{m}^T)$ as $\nu \to \infty$ under RHC sampling design for any $\mathbf{m} \in \mathbb{R}^p$ and $\Gamma_2 = \lim_{\nu\to\infty} \Sigma_2$. Hence, $\sqrt{n}(\hat{\overline{\mathbf{V}}}_2 - \overline{\mathbf{V}}) \xrightarrow{\mathcal{L}} N(0, \Gamma_2)$ as $\nu \to \infty$ under RHC sampling design. □

Next, suppose that $\overline{\mathbf{W}} = \sum_{i=1}^N \mathbf{W}_i/N$, $\hat{\overline{\mathbf{W}}}_1 = \sum_{i\in s}(N\pi_i)^{-1}\mathbf{W}_i$ and $\hat{\overline{\mathbf{W}}}_2 = \sum_{i\in s} G_i\mathbf{W}_i/NX_i$ for $\mathbf{W}_i = (h_i, X_ih_i,\ X_i^2)$, $i = 1, \ldots, N$. Let us also define $\hat{\overline{X}}_1 = \sum_{i\in s} (N\pi_i)^{-1}X_i$. Now, we state the following lemma.

**Lemma S 5.** *Suppose that $C0$ through $C2$ hold. Then, under SRSWOR, LMS sampling design and any HE$\pi$PS sampling design, we have $\hat{\overline{\mathbf{W}}}_1 - \overline{\mathbf{W}} = o_p(1)$, $\sqrt{n}(\hat{\overline{X}}_1 - \overline{X}) = O_p(1)$ and $\sqrt{n}(\sum_{i\in s}(N\pi_i)^{-1} - 1) = O_p(1)$ as $\nu \to \infty$. Moreover, under RHC sampling design, we have $\hat{\overline{\mathbf{W}}}_2 - \overline{\mathbf{W}} = o_p(1)$ and $\sqrt{n}(\sum_{i\in s} G_i/NX_i - 1) = O_p(1)$ as $\nu \to \infty$.*

*Proof.* We first show that as $\nu \to \infty$, $\hat{\overline{\mathbf{W}}}_1 - \overline{\mathbf{W}} = o_p(1)$, $\sqrt{n}(\hat{\overline{X}}_1 - \overline{X}) = O_p(1)$ and $\sqrt{n}(\sum_{i\in s}( N\pi_i)^{-1} - 1) = O_p(1)$ under a high entropy sampling design $P(s)$ satisfying (S2.1) in Lemma S3. Fix $\mathbf{m} \in \mathbb{R}^{2p+1}$. Suppose that $\tilde{R}(s)$ is a rejective sampling design with inclusion probabilities equal to those of $P(s)$ (cf. Berger (1998)). Under $\tilde{R}(s)$, $var(\mathbf{m}(\sqrt{n}(\hat{\overline{\mathbf{W}}}_1 - \overline{\mathbf{W}})^T)) = \mathbf{m}(nN^{-2} \sum_{i=1}^N(\mathbf{W}_i - \mathbf{T}\pi_i)^T(\mathbf{W}_i - \mathbf{T}\pi_i)(\pi_i^{-1} - 1))\mathbf{m}^T(1 + e)$ (see Theorem 6.1 in Hájek (1964)), where $\mathbf{T} = \sum_{i=1}^N \mathbf{W}_i(1 - \pi_i)/\sum_{i=1}^N \pi_i(1 - \pi_i)$, and $e \to 0$ as

$\nu \to \infty$ whenever $\sum_{i=1}^{N} \pi_i(1-\pi_i) \to \infty$ as $\nu \to \infty$. Note that (S2.1) holds under $\tilde{R}(s)$, and hence $\sum_{i=1}^{N} \pi_i(1-\pi_i) \to \infty$ as $\nu \to \infty$ under $\tilde{R}(s)$ because (S2.1) holds under $P(s)$, and C0 holds. Then, $\mathbf{m}(nN^{-2}\sum_{i=1}^{N}(\mathbf{W}_i - \mathbf{T}\pi_i)^T(\mathbf{W}_i - \mathbf{T}\pi_i)(\pi_i^{-1}-1))\mathbf{m}^T \leq nN^{-2}\sum_{i=1}^{N}(\mathbf{m}\mathbf{W}_i^T)^2/\pi_i = O(1)$ under $\tilde{R}(s)$ since C1 holds. Therefore, $\sqrt{n}(\hat{\overline{\mathbf{W}}}_1 - \overline{\mathbf{W}}) = O_p(1)$ as $\nu \to \infty$ under $\tilde{R}(s)$ since $var(\mathbf{m}(\sqrt{n}(\hat{\overline{\mathbf{W}}}_1 - \overline{\mathbf{W}})^T)) = O(1)$ as $\nu \to \infty$ for any $\mathbf{m} \in \mathbb{R}^{2p+1}$ under $\tilde{R}(s)$. Now, $\sum_{s\in E} P(s) \leq \sum_{s\in E} \tilde{R}(s) + \sum_{s\in\mathcal{S}} |P(s) - \tilde{R}(s)| \leq \sum_{s\in E} \tilde{R}(s) + (2D(P||\tilde{R}))^{1/2} \leq \sum_{s\in E} \tilde{R}(s) + (2D(P||R))^{1/2}$ (see Lemmas 2 and 3 in Berger (1998)), where $E = \{s \in \mathcal{S} : ||\sqrt{n}(\hat{\overline{\mathbf{W}}}_1 - \overline{\mathbf{W}})|| > \delta\}$ for $\delta > 0$ and $R(s)$ is any other rejective sampling design. Let us consider a rejective sampling design $R(s)$ such that $D(P||R) \to 0$ as $\nu \to \infty$. Therefore, given any $\epsilon > 0$, there exists a $\delta > 0$ such that $\sum_{s\in E} P(s) \leq \epsilon$ for all sufficiently large $\nu$. Hence, as $\nu \to \infty$, $\sqrt{n}(\hat{\overline{\mathbf{W}}}_1 - \overline{\mathbf{W}}) = O_p(1)$ and $\hat{\overline{\mathbf{W}}}_1 - \overline{\mathbf{W}} = o_p(1)$ under $P(s)$. Similarly, we can show that as $\nu \to \infty$, $\sqrt{n}(\hat{\overline{X}}_1 - \overline{X}) = O_p(1)$ and $\sqrt{n}(\sum_{i\in s}(N\pi_i)^{-1} - 1) = O_p(1)$ under $P(s)$. Now, recall from the proof of Lemma S4 that SRSWOR and LMS sampling design are high entropy sampling designs, and they satisfy (S2.1). Also, any HE$\pi$PS sampling design satisfies (S2.1). Therefore, as $\nu \to \infty$, $\hat{\overline{\mathbf{W}}}_1 - \overline{\mathbf{W}} = o_p(1)$, $\sqrt{n}(\hat{\overline{X}}_1 - \overline{X}) = O_p(1)$ and $\sqrt{n}(\sum_{i\in s}(N\pi_i)^{-1} - 1) = O_p(1)$ under the above-mentioned sampling designs.

Under RHC sampling design, $var(\mathbf{m}(\sqrt{n}(\hat{\overline{\mathbf{W}}}_2-\overline{\mathbf{W}})^T))=\mathbf{m}(n\gamma\overline{X}N^{-1}\sum_{i=1}^{N}$ $(\mathbf{W}_i-X_i\overline{\mathbf{W}}/\overline{X})^T(\mathbf{W}_i-X_i\overline{\mathbf{W}}/\overline{X})/X_i)\mathbf{m}^T$ (see Ohlsson (1986)). Recall from the proof of Lemma S4 that $n\gamma=O(1)$ as $\nu\rightarrow\infty$. Then, $var(\mathbf{m}(\sqrt{n}(\hat{\overline{\mathbf{W}}}_2-\overline{\mathbf{W}})^T))\leq n\gamma(\overline{X}/N)\sum_{i=1}^{N}(\mathbf{m}\mathbf{W}_i^T)^2/X_i=O(1)$ as $\nu\rightarrow\infty$ since C1 and C2 hold. Hence, as $\nu\rightarrow\infty$, $\sqrt{n}(\hat{\overline{\mathbf{W}}}_2-\overline{\mathbf{W}})=O_p(1)$ and $\hat{\overline{\mathbf{W}}}_2-\overline{\mathbf{W}}=o_p(1)$ under RHC sampling design. Similarly, we can show that as $\nu\rightarrow\infty$, $\sqrt{n}(\sum_{i\in s}G_i/NX_i-1)=O_p(1)$ under RHC sampling design. □

Recall from the $2^{nd}$ paragraph in the Appendix that we denote the HT, the RHC, the Hájek, the ratio, the product, the GREG and the PEML estimators of population means of $h(y)$ by $\hat{\overline{h}}_{HT}$, $\hat{\overline{h}}_{RHC}$, $\hat{\overline{h}}_{H}$, $\hat{\overline{h}}_{RA}$, $\hat{\overline{h}}_{PR}$, $\hat{\overline{h}}_{GREG}$ and $\hat{\overline{h}}_{PEML}$, respectively. Suppose that $\hat{\overline{h}}$ denotes one of $\hat{\overline{h}}_{HT}$, $\hat{\overline{h}}_{H}$, $\hat{\overline{h}}_{RA}$, $\hat{\overline{h}}_{PR}$, and $\hat{\overline{h}}_{GREG}$ with $d(i,s)=(N\pi_i)^{-1}$. Then, a Taylor type expansion of $\hat{\overline{h}}-\overline{h}$ can be obtained as $\hat{\overline{h}}-\overline{h}=\Theta(\hat{\overline{\mathbf{V}}}_1-\overline{\mathbf{V}})+\mathbf{Z}$, where $\hat{\overline{\mathbf{V}}}_1=\sum_{i\in s}(N\pi_i)^{-1}\mathbf{V}_i$, and the $\mathbf{V}_i$'s, $\Theta$ and $\mathbf{Z}$ are as described in Table 1 below. On the other hand, if $\hat{\overline{h}}$ is either $\hat{\overline{h}}_{RHC}$ or $\hat{\overline{h}}_{GREG}$ with $d(i,s)=G_i/NX_i$, a Taylor type expansion of $\hat{\overline{h}}-\overline{h}$ can be obtained as $\hat{\overline{h}}-\overline{h}=\Theta(\hat{\overline{\mathbf{V}}}_2-\overline{\mathbf{V}})+\mathbf{Z}$. Here, $\hat{\overline{\mathbf{V}}}_2=\sum_{i\in s}G_i\mathbf{V}_i/NX_i$, the $G_i$'s are as in the paragraph containing Table 8 in the main text, and the $\mathbf{V}_i$'s, $\Theta$ and $\mathbf{Z}$ are once again described in Table 1. In Table 1, $\hat{\overline{X}}_1=\sum_{i\in s}(N\pi_i)^{-1}X_i$, $\hat{\overline{X}}_2=\hat{\overline{X}}_1/\sum_{i\in s}(N\pi_i)^{-1}$, $\hat{\beta}_1=(\sum_{i\in s}(N\pi_i)^{-1}$ $\sum_{i\in s}(N\pi_i)^{-1}h_iX_i-\hat{\overline{h}}_{HT}\hat{\overline{X}}_1)/(\sum_{i\in s}(N\pi_i)^{-1}\sum_{i\in s}(N\pi_i)^{-1}X_i^2-(\hat{\overline{X}}_1)^2)$ and

Table 1: Expressions of $\mathbf{V}_i$, $\Theta$ and $\mathbf{Z}$ for different $\hat{\overline{h}}$'s

| $\hat{\overline{h}}$ | $\mathbf{V}_i$ | $\Theta$ | $\mathbf{Z}$ |
|---|---|---|---|
| $\hat{\overline{h}}_{HT}$ | $h_i$ | $1$ | $0$ |
| $\hat{\overline{h}}_{H}$ | $h_i-\overline{h}$ | $(\sum_{i\in s}(N\pi_i)^{-1})^{-1}$ | $0$ |
| $\hat{\overline{h}}_{RA}$ | $h_i-\overline{h}X_i/\overline{X}$ | $\overline{X}/\hat{\overline{X}}_1$ | $0$ |
| $\hat{\overline{h}}_{PR}$ | $h_i+\overline{h}X_i/\overline{X}$ | $\hat{\overline{X}}_1/\overline{X}$ | $-(1-\hat{\overline{X}}_1/\overline{X}))^2\overline{h}$ |
| $\hat{\overline{h}}_{GREG}$ with $d(i,s)=(N\pi_i)^{-1}$ | $h_i-\overline{h}-$ $S_{xh}(X_i-\overline{X})/S_x^2$ | $(\sum_{i\in s}(N\pi_i)^{-1})^{-1}$ | $(\hat{\overline{X}}_2-\overline{X})\times$ $(S_{xh}/S_x^2-\hat{\beta}_1)$ |
| $\hat{\overline{h}}_{RHC}$ | $h_i$ | $1$ | $0$ |
| $\hat{\overline{h}}_{GREG}$ with $d(i,s)=G_i/NX_i$ | $h_i-\overline{h}-$ $S_{xh}(X_i-\overline{X})/S_x^2$ | $(\sum_{i\in s}G_i/NX_i)^{-1}$ | $\overline{X}((\sum_{i\in s}G_i/NX_i)^{-1}$ $-1)(S_{xh}/S_x^2-\hat{\beta}_2)$ |

$\hat{\beta}_2=(\sum_{i\in s}(G_i/NX_i)\sum_{i\in s}(G_ih_i/N)-\hat{\overline{h}}_{RHC}\overline{X})/(\sum_{i\in s}(G_i/NX_i)\sum_{i\in s}(G_iX_i/N)$ $-\overline{X}^2)$. Now, we state the following Lemma.

**Lemma S 6.** (*i*) *Suppose that C0 through C3 hold. Further, suppose that $\hat{\overline{h}}$ is one of $\hat{\overline{h}}_{HT}$, $\hat{\overline{h}}_{H}$, $\hat{\overline{h}}_{RA}$, $\hat{\overline{h}}_{PR}$, and $\hat{\overline{h}}_{GREG}$ with $d(i,s)=(N\pi_i)^{-1}$. Then, under SRSWOR, LMS sampling design and any HE$\pi$PS sampling design,*

$$\sqrt{n}(\hat{\overline{h}}-\overline{h})\xrightarrow{\mathcal{L}}N(0,\Gamma)\ \text{as}\ \nu\to\infty \tag{S2.2}$$

*for some p.d. matrix $\Gamma$.*

(*ii*) *Next, suppose that C0 through C2 and C4 hold, and $\hat{\overline{h}}$ is $\hat{\overline{h}}_{RHC}$ or $\hat{\overline{h}}_{GREG}$ with $d(i,s)=G_i/NX_i$. Then,* (S2.2) *holds under RHC sampling design.*

*Proof.* It can be shown from Lemma S4 that $\sqrt{n}(\hat{\overline{\mathbf{V}}}_1-\overline{\mathbf{V}})\xrightarrow{\mathcal{L}}N(0,\Gamma_1)$ as

$\nu \to \infty$ under SRSWOR, LMS sampling design and any HE$\pi$PS sampling design, where $\Gamma_1$=$\lim_{\nu\to\infty} nN^{-2}\sum_{i=1}^{N}(\mathbf{V}_i - \mathbf{T}\pi_i)^T(\mathbf{V}_i - \mathbf{T}\pi_i)(\pi_i^{-1} - 1)$ with $\mathbf{T}$=$\sum_{i=1}^{N}\mathbf{V}_i(1-\pi_i)/\sum_{i=1}^{N}\pi_i(1-\pi_i)$. Note that $\Gamma_1$ is a p.d. matrix under each of the above sampling designs as C3 holds under these sampling designs. Let us now consider from Table 1 various choices of $\Theta$ and $\mathbf{Z}$ corresponding to $\hat{\overline{h}}_{HT}$, $\hat{\overline{h}}_{H}$, $\hat{\overline{h}}_{RA}$, $\hat{\overline{h}}_{PR}$, and $\hat{\overline{h}}_{GREG}$ with $d(i,s)$=$(N\pi_i)^{-1}$. Then, it can be shown from Lemma S5 that for each of these choices, $\sqrt{n}\mathbf{Z}$=$o_p(1)$ and $\Theta - 1$=$o_p(1)$ as $\nu \to \infty$ under the above-mentioned sampling designs. Therefore, (S2.2) holds under those sampling designs with $\Gamma$=$\Gamma_1$. This completes the proof of $(i)$ in Lemma 6

We can show from Lemma S4 that $\sqrt{n}(\hat{\overline{\mathbf{V}}}_2 - \overline{\mathbf{V}}) \xrightarrow{\mathcal{L}} N(0,\Gamma_2)$ as $\nu \to \infty$ under RHC sampling design, where $\Gamma_2$=$\lim_{\nu\to\infty} n\gamma\ \overline{X}N^{-1}\ \sum_{i=1}^{N}(\mathbf{V}_i - X_i\overline{\mathbf{V}}/\overline{X})^T(\mathbf{V}_i - X_i\overline{\mathbf{V}}/\overline{X})/X_i$ with $\gamma$=$\sum_{i=1}^{n}N_i(N_i-1)/N(N-1)$. Note that $\Gamma_2$ is a p.d. matrix since C4 holds. Let us now consider from Table 1 different choices of $\Theta$ and $\mathbf{Z}$ corresponding to $\hat{\overline{h}}_{RHC}$, and $\hat{\overline{h}}_{GREG}$ with $d(i,s)$=$G_i/NX_i$. Then, it follows from Lemma S5 that for each of these choices, $\sqrt{n}\mathbf{Z}$=$o_p(1)$ and $\Theta - 1$=$o_p(1)$ as $\nu \to \infty$ under RHC sampling design. Therefore, (S2.2) holds under RHC sampling design with $\Gamma$=$\Gamma_2$. This completes the proof of $(ii)$ in Lemma 6 □

Let $\{\mathbf{V}_i\}_{i=1}^{N}$ be as described in Table 1. Recall $\Sigma_1$ and $\Sigma_2$ from the paragraph preceding Lemma S2 in this supplement. Note that the expression of $\Sigma_1$ remains the same for different HE$\pi$PS sampling designs. Also, recall from the paragraph preceding Theorem 3 in the main text that $\phi{=}\overline{X} - (n/N)\sum_{i=1}^{N} X_i^2/N\overline{X}$. Now, we state the following lemma.

**Lemma S 7.** $(i)$ *Suppose that C0 through C3 hold. Further, suppose that* $\sigma_1^2$ *and* $\sigma_2^2$ *denote* $\lim_{\nu\to\infty}\nabla g(\mu_0)\Sigma_1\nabla g(\mu_0)^T$ *under SRSWOR and LMS sampling design, respectively, where* $\mu_0{=}\lim_{\nu\to\infty}\overline{h}$. *Then, we have* $\sigma_1^2{=}\sigma_2^2{=}(1-\lambda)\lim_{\nu\to\infty}\sum_{i=1}^{N}(A_i-\bar{A})^2/N$ *for* $A_i{=}\nabla g(\mu_0)\,\boldsymbol{V}_i^T$, $i{=}1,\ldots,N$.
$(ii)$ *Next, suppose that C4 holds, and* $\sigma_3^2{=}\lim_{\nu\to\infty}\nabla g(\mu_0)\Sigma_2\nabla g(\mu_0)^T$ *in the case of RHC sampling design. Then, we have* $\sigma_3^2{=}\lim_{\nu\to\infty} n\gamma((\overline{X}/N)\sum_{i=1}^{N} A_i^2/X_i{-}\bar{A}^2)$. *On the other hand, if C0 through C3 hold, and* $\sigma_4^2{=}\lim_{\nu\to\infty}\nabla g(\mu_0)\Sigma_1\nabla g(\mu_0)^T$ *under any HE$\pi$PS sampling design, then we have* $\sigma_4^2{=}\lim_{\nu\to\infty}\big\{(1/N)\sum_{i=1}^{N} A_i^2\big((\overline{X}/X_i){-}(n/N)\big){-}\phi^{-1}\overline{X}^{-1}\big((n/N)\sum_{i=1}^{N} A_iX_i/N{-}\overline{AX}\big)^2\big\}$. *Further, if C0 holds with* $\lambda{=}0$ *and C1 through C3 hold, then we have* $\sigma_4^2{=}\sigma_3^2{=}\lim_{\nu\to\infty}((\overline{X}/N)\sum_{i=1}^{N} A_i^2/X_i-\bar{A}^2)$.

*Proof.* Let us first note that the limits in the expressions of $\sigma_1^2$ and $\sigma_2^2$ exist in view of C3. Also, note that $\nabla g(\mu_0)\Sigma_1\nabla g(\mu_0)^T{=}nN^{-2}\sum_{i=1}^{N}(A_i - T_A\pi_i)^2(\pi_i^{-1}-1){=}nN^{-2}\,[\sum_{i=1}^{N} A_i^2(\pi_i^{-1}-1)-(\sum_{i=1}^{N} A_i(1-\pi_i))^2/\sum_{i=1}^{N}\pi_i(1-\pi_i)]$, where $T_A{=}\sum_{i=1}^{N} A_i(1-\pi_i)/\sum_{i=1}^{N}\pi_i(1-\pi_i)$ and $A_i{=}\nabla g(\mu_0)\mathbf{V}_i^T$. Now, substituting $\pi_i{=}n/N$ in the above expression for SRSWOR, we get $\sigma_1^2{=}$

$\lim_{\nu\to\infty} nN^{-2}\,[\sum_{i=1}^N A_i^2(N/n-1)-(\sum_{i=1}^N A_i(1-n/N))^2/n(1-n/N)]=\lim_{\nu\to\infty}$ $(1-n/N)\sum_{i=1}^N(A_i-\bar{A})^2/N$. Since C0 holds, we have $\sigma_1^2=(1-\lambda)\lim_{\nu\to\infty}$ $\sum_{i=1}^N(A_i-\bar{A})^2/N$. Let $\{\pi_i\}_{i=1}^N$ be the inclusion probabilities of LMS sampling design. Then, $\sigma_2^2-\sigma_1^2=\lim_{\nu\to\infty} nN^{-2}[\sum_{i=1}^N A_i^2(\pi_i^{-1}-N/n)-((\sum_{i=1}^N A_i(1-\pi_i))^2/\sum_{i=1}^N\pi_i(1-\pi_i)-(\sum_{i=1}^N A_i(1-n/N))^2/n(1-n/N))]$. Now, it can be shown from the proof of Lemma S3 that $\max_{1\leq i\leq N}|N\pi_i/n-1|\to 0$ as $\nu\to\infty$. Therefore, using C1, we can show that $\lim_{\nu\to\infty} nN^{-2}\sum_{i=1}^N A_i^2(\pi_i^{-1}-N/n)=0$ and $\lim_{\nu\to\infty} nN^{-2}[(\sum_{i=1}^N A_i(1-\pi_i))^2/\sum_{i=1}^N\pi_i(1-\pi_i)-(\sum_{i=1}^N A_i(1-n/N))^2/n(1-n/N)]=0$, and consequently $\sigma_1^2=\sigma_2^2$. This completes the proof of $(i)$ in Lemma 7.

Next, consider the case of RHC sampling design and note that the limit in the expression of $\sigma_3^2$ exists in view of C4. Also, note that $\nabla g(\mu_0)\Sigma_2\nabla g(\mu_0)^T$ $=n\gamma(\overline{X}/N)\sum_{i=1}^N(A_i-\bar{A}X_i/\overline{X})^2/X_i=n\gamma((\overline{X}/N)\sum_{i=1}^N A_i^2/X_i-\bar{A}^2)$, where $\bar{A}=\sum_{i=1}^N A_i/N$ and $\gamma=\sum_{i=1}^n N_i(N_i-1)/N(N-1)$. Thus we have $\sigma_3^2=\lim_{\nu\to\infty}$ $n\gamma((\overline{X}/N)\sum_{i=1}^N A_i^2/X_i-\bar{A}^2)$.

Next, note that the limit in the expression of $\sigma_4^2$ exists in view of C3. Substituting $\pi_i=nX_i/\sum_{i=1}^N X_i$ in $\nabla g(\mu_0)\Sigma_1\nabla g(\mu_0)^T$ for any HE$\pi$PS sampling design, we get $\sigma_4^2=\lim_{\nu\to\infty} nN^{-2}[\sum_{i=1}^N A_i^2(\sum_{i=1}^N X_i/nX_i-1)-(\sum_{i=1}^N A_i(1-nX_i/\sum_{i=1}^N X_i))^2/\sum_{i=1}^N(nX_i/\sum_{i=1}^N X_i)(1-nX_i/\sum_{i=1}^N X_i)]=$ $\lim_{\nu\to\infty}\big\{(1/N)\sum_{i=1}^N A_i^2\big((\overline{X}/X_i)-(n/N)\big)-\phi^{-1}\overline{X}^{-1}\big((n/N)\sum_{i=1}^N A_iX_i/N-$

$\overline{A}\,\overline{X}\big)^2\}$. Further, we can show that $\sigma_4^2{=}\lim_{\nu\to\infty}((\overline{X}/N)\sum_{i=1}^N A_i^2/X_i - \bar{A}^2)$, when C1 and C2 hold, and C0 holds with $\lambda$=0. It also follows from Lemma S1 that $n\gamma \to 1$ as $\nu \to \infty$, when C0 holds with $\lambda$=0. Thus we have $\sigma_3^2{=}\sigma_4^2{=}\lim_{\nu\to\infty}((\overline{X}/N)\sum_{i=1}^N A_i^2/X_i - \bar{A}^2)$. This completes the proof of $(ii)$ in Lemma 7. □

**Lemma S 8.** *Suppose that C0 through C2 hold. Then under SRSWOR, LMS sampling design and any HE$\pi$PS sampling design, we have*

$$(i) \quad u^* = \max_{i\in s}|Z_i| = o_p(\sqrt{n}), \text{ and } \quad (ii) \quad \sum_{i\in s}\pi_i^{-1}Z_i/\sum_{i\in s}\pi_i^{-1}Z_i^2 = O_p(1/\sqrt{n})$$

*as $\nu \to \infty$, where $Z_i{=}X_i - \overline{X}$ for $i{=}1,\ldots,N$*

*Proof.* Let $P(s)$ be any sampling design and $E_P$ be the expectation with respect to $P(s)$. Then, $E_P(u^*/\sqrt{n}) \leq (\max_{1\leq i\leq N} X_i{+}\overline{X})/\sqrt{n} \leq \overline{X}(\max_{1\leq i\leq N} X_i/\min_{1\leq i\leq N} X_i + 1)/\sqrt{n}{=}o(1)$ as $\nu \to \infty$ since C1 and C2 hold. Therefore, $(i)$ holds under $P(s)$ by Markov inequality. Thus $(i)$ holds under SRSWOR, LMS sampling design and any HE$\pi$PS sampling design.

Using similar arguments as in the first paragraph of the proof of Lemma S5, it can be shown that $\sqrt{n}(\sum_{i\in s} Z_i/N\pi_i - \overline{Z}){=}\sqrt{n}\sum_{i\in s} Z_i/N\pi_i{=}O_p(1)$ and $\sum_{i\in s} Z_i^2/N\pi_i - \sum_{i=1}^N Z_i^2/N{=}o_p(1)$ as $\nu \to \infty$ under a high entropy sampling design $P(s)$ satisfying (S2.1) in Lemma S3. Therefore, $1/(\sum_{i\in s} Z_i^2/N\pi_i)$ $=O_p(1)$ as $\nu \to \infty$ under $P(s)$ since $\sum_{i=1}^N Z_i^2/N$ is bounded away from 0 as

$\nu \to \infty$ by C1. Thus under $P(s)$, $\sum_{i\in s}\pi_i^{-1}Z_i/\sum_{i\in s}\pi_i^{-1}Z_i^2 = O_p(1/\sqrt{n})$ as $\nu \to \infty$.

It follows from Lemma S3 that SRSWOR and LMS sampling design are high entropy sampling designs and satisfy (S2.1) Also, any HE$\pi$PS sampling design satisfies (S2.1) since C2 holds. Therefore, the result in $(ii)$ holds under the above-mentioned sampling designs. □

## S3 Proofs of Remark 1 and Theorems 2, 3, 6 and 7

In this section, we give the proofs of Remark 1 and Theorems 2, 3, 6 and 7 of the main text.

***Proof of Theorem*** 2. Let us first consider a HE$\pi$PS sampling design. Then, it can be shown in the same way as in the $1^{st}$ paragraph of the proof of Theorem 1 that $\sqrt{n}(\hat{\overline{h}}_{PEML} - \hat{\overline{h}}_{GREG}) = o_p(1)$ for $d(i,s) = (N\pi_i)^{-1}$ under this sampling design. It can also be shown in the same way as in the $1^{st}$ paragraph of the proof of Theorem 1 that if $\hat{\overline{h}}$ is one of $\hat{\overline{h}}_{HT}$, $\hat{\overline{h}}_{H}$, and $\hat{\overline{h}}_{GREG}$ and $\hat{\overline{h}}_{PEML}$ with $d(i,s) = (N\pi_i)^{-1}$, then (5.1) in the proof of Theorem 1 holds under the above-mentioned sampling design. Here, we recall from Table 2 in the main text that the HT, the ratio and the product estimators coincide under any HE$\pi$PS sampling design. Further, the asymptotic MSE of $\sqrt{n}(g(\hat{\overline{h}})-g(\overline{h}))$ is $\nabla g(\mu_0)\Gamma_1\ (\nabla g(\mu_0))^T$, where $\mu_0{=}\lim_{\nu\to\infty}\overline{h}$, $\Gamma_1{=}\lim_{\nu\to\infty} nN^{-2}\sum_{i=1}^{N}(\mathbf{V}_i-\mathbf{T}\pi_i)^T(\mathbf{V}_i-\mathbf{T}\pi_i)(\pi_i^{-1}-1)$, and $\mathbf{V}_i$ in $\Gamma_1$ is $h_i$ or $h_i-\overline{h}$ or $h_i-\overline{h}-S_{xh}(X_i-\overline{X})/S_x^2$ if $\hat{\overline{h}}$ is $\hat{\overline{h}}_{HT}$ or $\hat{\overline{h}}_H$, or $\hat{\overline{h}}_{GREG}$ with $d(i,s){=}(N\pi_i)^{-1}$, respectively. Now, since $\sqrt{n}(\hat{\overline{h}}_{PEML}-\hat{\overline{h}}_{GREG}){=}o_p(1)$ for $\nu\to\infty$ under any HE$\pi$PS sampling design, $g(\hat{\overline{h}}_{GREG})$ and $g(\hat{\overline{h}}_{PEML})$ have the same asymptotic distribution under this sampling design. Thus under any HE$\pi$PS sampling design, $g(\hat{\overline{h}}_{GREG})$ and $g(\hat{\overline{h}}_{PEML})$ with $d(i,s){=}(N\pi_i)^{-1}$ form class 5, $g(\hat{\overline{h}}_{HT})$ forms class 6, and $g(\hat{\overline{h}}_H)$ forms class 7 in Table 2 of the main text. This completes the proof of $(i)$ in Theorem 2.

Let us now consider the RHC sampling design. We can show from $(ii)$ in Lemma S6 that $\sqrt{n}(\hat{\overline{h}}-\overline{h})\xrightarrow{\mathcal{L}}N(0,\Gamma)$ as $\nu\to\infty$ for some p.d. matrix $\Gamma$, when $\hat{\overline{h}}$ is either $\hat{\overline{h}}_{RHC}$ or $\hat{\overline{h}}_{GREG}$ with $d(i,s){=}G_i/NX_i$ under RHC sampling design. Further, $\sqrt{n}(\hat{\overline{h}}_{PEML}-\hat{\overline{h}}_{GREG}){=}o_p(1)$ as $\nu\to\infty$ for $d(i,s){=}G_i/NX_i$ under RHC sampling design since C2 holds, and $S_x^2$ is bounded away from 0 as $\nu\to\infty$ (see A2.2 of Appendix 2 in Chen and Sitter (1999)). Therefore, if $\hat{\overline{h}}$ is one of $\hat{\overline{h}}_{RHC}$, and $\hat{\overline{h}}_{GREG}$ and $\hat{\overline{h}}_{PEML}$ with $d(i,s){=}G_i/NX_i$, then we have

$$\sqrt{n}(g(\hat{\overline{h}})-g(\overline{h}))\xrightarrow{\mathcal{L}}N(0,\Delta^2)\text{ as }\nu\to\infty \tag{S3.1}$$

for some $\Delta^2>0$ by the delta method and the condition $\nabla g(\mu_0)\neq 0$ at $\mu_0{=}\lim_{\nu\to\infty}\overline{h}$. Moreover, it follows from the proof of $(ii)$ in Lemma

S6 that $\Delta^2= \nabla g(\mu_0)\Gamma_2(\nabla g(\mu_0))^T$, where $\Gamma_2$=$\lim_{\nu\to\infty} n\gamma\overline{X}N^{-1}\sum_{i=1}^{N}(\mathbf{V}_i - X_i\overline{\mathbf{V}}/\overline{X})^T(\mathbf{V}_i - X_i\overline{\mathbf{V}}/\overline{X})/X_i$. It further follows from Table 1 in this supplement that $\mathbf{V}_i$ in $\Gamma_2$ is $h_i$ if $\hat{\overline{h}}$ is $\hat{\overline{h}}_{RHC}$. Also, $\mathbf{V}_i$ in $\Gamma_2$ is $h_i - \overline{h} - S_{xh}(X_i - \overline{X})/S_x^2$ if $\hat{\overline{h}}$ is $\hat{\overline{h}}_{GREG}$ with $d(i,s)$=$G_i/NX_i$. Now, $g(\hat{\overline{h}}_{GREG})$ and $g(\hat{\overline{h}}_{PEML})$ have the same asymptotic distribution under RHC sampling design since $\sqrt{n}(\hat{\overline{h}}_{PEML} - \hat{\overline{h}}_{GREG})$=$o_p(1)$ for $\nu \to \infty$ under this sampling design as pointed out earlier in this paragraph. Thus $g(\hat{\overline{h}}_{GREG})$ and $g(\hat{\overline{h}}_{PEML})$ with $d(i,s)$=$G_i/NX_i$ under RHC sampling design form class 8, and $g(\hat{\overline{h}}_{RHC})$ forms class 9 in Table 2 of the main article. This completes the proof of $(ii)$ in Theorem 2. □

***Proof of Remark*** 1. It follows from *(ii)* in Lemma S7 that in the case of $\lambda$=0,

$$\sigma_3^2 = \sigma_4^2 = \lim_{\nu\to\infty}((\overline{X}/N)\sum_{i=1}^{N} A_i^2/X_i - \bar{A}^2), \tag{S3.2}$$

where $\sigma_3^2$ and $\sigma_4^2$ are as defined in the statement of Lemma S7, and $A_i$=$\nabla g(\mu_0)\mathbf{V}_i^T$ for different choices of $\mathbf{V}_i$ mentioned in the proof of Theorem 2 above. Thus $g(\hat{\overline{h}}_{GREG})$ with $d(i,s)$=$(N\pi_i)^{-1}$ under any HE$\pi$PS sampling design, and with $d(i,s)$=$G_i/NX_i$ under RHC sampling design have the same asymptotic MSE. Therefore, class 8 is merged with class 5 in Table 2 of the main text. Further, (S3.2) implies that $g(\hat{\overline{h}}_{HT})$ under any HE$\pi$PS sampling design and $g(\hat{\overline{h}}_{RHC})$ have the same asymptotic MSE. Therefore, class 9 is merged with class 6 in Table 2 of the main text. This completes the proof of Remark 1. □

***Proof of Theorem*** 3. Recall the expression of $A_i$'s from the proofs of Theorem 1 and Remark 1. Note that $\lim_{\nu\to\infty}\sum(A_i-\bar{A})^2/N=\lim_{\nu\to\infty}\sum(B_i-\bar{B})^2/N$, $\lim_{\nu\to\infty} n\gamma\big((\overline{X}/N)\sum_{i=1}^{N}A_i^2/X_i-\bar{A}^2\big)=\lim_{\nu\to\infty} n\gamma\big((\overline{X}/N)\sum_{i=1}^{N}B_i^2/X_i-\bar{B}^2\big)$ and $\lim_{\nu\to\infty}\big\{(1/N)\sum_{i=1}^{N}A_i^2\big((\overline{X}/X_i)-(n/N)\big)-\phi^{-1}\overline{X}^{-1}\big((n/N)\times\sum_{i=1}^{N}A_iX_i/N-\bar{A}\,\overline{X}\big)^2\big\}=\lim_{\nu\to\infty}\big\{(1/N)\sum_{i=1}^{N}B_i^2\big((\overline{X}/X_i)-(n/N)\big)-\phi^{-1}\overline{X}^{-1}\times\big((n/N)\sum_{i=1}^{N}B_iX_i/N-\bar{B}\,\overline{X}\big)^2\big\}$ for $B_i=\nabla g(\overline{h})\mathbf{V}_i^T$ and $\mathbf{V}_i$ as in Table 1 in this supplement since $\nabla g(\overline{h})\to\nabla g(\mu_0)$ as $\nu\to\infty$. Here, $\phi=\overline{X}-(n/N)\sum_{i=1}^{N}X_i^2/N\overline{X}$. Then, from Lemma S7 and the expressions of asymptotic MSEs of $\sqrt{n}(g(\hat{\overline{h}})-g(\overline{h}))$ discussed in the proofs of Theorems 1 and 2 , the results in Table 3 of the main text follow. This completes the proof of Theorem 3. □

***Proof of Theorem*** 6. Using similar arguments as in the $1^{st}$ paragraph of the proof of Theorem 4, we can say that under SRSWOR and LMS sampling design, conclusions of Theorems 1 and 3 hold *a.s.* $[\mathbb{P}]$ for $d$=1, $p$=2, $h(y)=(y,y^2)$ and $g(s_1,s_2)=s_2-s_1^2$ in the same way as conclusions of Theorems 1 and 3 hold *a.s.* $[\mathbb{P}]$ for $d$=$p$=1, $h(y)=y$ and $g(s)=s$ in the $1^{st}$ paragraph of the proof of Theorem 4. Note that $W_i=Y_i^2-2Y_i\overline{Y}$ for the above choices of $h$ and $g$. Further, it follows from SLLN and the condition

$E_{\mathbb{P}}(\epsilon_i)^8 < \infty$ that the $\Delta_i^2$'s in Table 3 in the main text can be expressed in terms of superpopulation moments of $(Y_i, X_i)$ $a.s.$ $[\mathbb{P}]$. Note that $\Delta_2^2 - \Delta_1^2 = cov_{\mathbb{P}}^2(\tilde{W}_i, X_i)$ $a.s.$ $[\mathbb{P}]$, where $\tilde{W}_i = Y_i^2 - 2Y_i E_{\mathbb{P}}(Y_i)$. Then, $\Delta_1^2 < \Delta_2^2$ $a.s.$ $[\mathbb{P}]$. This completes the proof of $(i)$ in Theorem 6.

Next consider the case of $0 \le \lambda < E_{\mathbb{P}}(X_i)/b$. Using the same line of arguments as in the $2^{nd}$ paragraph of the proof of Theorem 4, it can be shown that under RHC and any HE$\pi$PS sampling designs, conclusions of Theorems 2 and 3 hold $a.s.$ $[\mathbb{P}]$ for $d$=1, $p$=2, $h(y)=(y, y^2)$ and $g(s_1, s_2) = s_2 - s_1^2$ in the same way as conclusions of Theorems 2 and 3 hold $a.s.$ $[\mathbb{P}]$ for $d$=$p$=1, $h(y)=y$ and $g(s)=s$ in the $2^{nd}$ paragraph of the proof of Theorem 4. Note that $\Delta_7^2 - \Delta_5^2 = \big\{\mu_1^2 cov_{\mathbb{P}}(\tilde{W}_i, X_i)\big(cov_{\mathbb{P}}(\tilde{W}_i, X_i) cov_{\mathbb{P}}(X_i, 1/X_i) - 2cov_{\mathbb{P}}(\tilde{W}_i, 1/X_i)\big)\big\} - \lambda^2 cov_{\mathbb{P}}^2(\tilde{W}_i, X_i)/\chi\mu_1 - \lambda cov_{\mathbb{P}}^2(\tilde{W}_i, X_i) \le \big\{\mu_1^2 \times cov_{\mathbb{P}}(\tilde{W}_i, X_i)\big(cov_{\mathbb{P}}(\tilde{W}_i, X_i) cov_{\mathbb{P}}(X_i, 1/X_i) - 2cov_{\mathbb{P}}(\tilde{W}_i, 1/X_i)\big)\big\}$ $a.s.$ $[\mathbb{P}]$ because $\chi > 0$. Recall from C6 that $\xi = \mu_3 - \mu_2\mu_1$ and $\mu_j = E_{\mathbb{P}}(X_i)^j$ for $j = -1, 1, 2, 3$. Then, from the linear model set up, we have $\big\{\mu_1^2 cov_{\mathbb{P}}(\tilde{W}_i, X_i) \times \big(cov_{\mathbb{P}}(\tilde{W}_i, X_i) cov_{\mathbb{P}}(X_i,\ 1/X_i) - 2cov_{\mathbb{P}}(\tilde{W}_i, 1/X_i)\big)\big\} = (\beta^2\mu_1)^2(\xi - 2\mu_1)((\xi + 2\mu_1)\zeta_1 - 2\zeta_2)$. Here, $\zeta_1 = 1 - \mu_1\mu_{-1}$ and $\zeta_2 = \mu_1 - \mu_2\mu_{-1}$. Note that $(\xi + 2\mu_1)\zeta_1 - 2\zeta_2 = \xi\zeta_1 + 2\mu_{-1}$ and $\zeta_1 < 0$. Therefore, $\big\{\mu_1^2 cov_{\mathbb{P}}(\tilde{W}_i, X_i)\big(cov_{\mathbb{P}}(\tilde{W}_i, X_i) cov_{\mathbb{P}}(X_i, 1/X_i) - 2cov_{\mathbb{P}}(\tilde{W}_i, 1/X_i)\big)\big\} < 0$ if $\xi > 2\max\{\mu_1, \mu_{-1}/(\mu_1\mu_{-1} - 1)\}$. Hence, $\Delta_7^2 - \Delta_5^2 < 0$ $a.s.$ $[\mathbb{P}]$. This completes the proof of $(ii)$ in Theorem 6. $\square$

***Proof of Theorem** 7.* Using the same line of arguments as in the $1^{st}$ paragraph of the proof of Theorem 4, it can be shown that under SRSWOR and LMS sampling design, conclusions of Theorems 1 and 3 hold *a.s.* $[\mathbb{P}]$ for $d$=2, $p$=5, $h(z_1, z_2)$=$(z_1, z_2, z_1^2\ , z_2^2,\ z_1 z_2)$ and $g(s_1, s_2, s_3, s_4, s_5)$=$(s_5 - s_1 s_2)/((s_3 - s_1^2)(s_4 - s_2^2))^{1/2}$ in the case of the correlation coefficient between $z_1$ and $z_2$, and for $d$=2, $p$=4, $h(z_1, z_2)$=$(z_1, z_2, z_2^2, z_1 z_2)$ and $g(s_1, s_2, s_3, s_4)$= $(s_4 - s_1 s_2)/(s_3 - s_2^2)$ in the case of the regression coefficient of $z_1$ on $z_2$ in the same way as conclusions of Theorems 1 and 3 hold *a.s.* $[\mathbb{P}]$ for $d$=$p$=1, $h(y)$=$y$ and $g(s)$=$s$ in the case of the mean of $y$ in the $1^{st}$ paragraph of the proof of Theorem 4. Further, if C0 holds with $0 \leq \lambda < E_{\mathbb{P}}(X_i)/b$, then using similar arguments as in the $2^{nd}$ paragraph of the proof of Theorem 4, it can also be shown that under RHC and any HE$\pi$PS sampling designs, conclusions of Theorems 2 and 3 hold *a.s.* $[\mathbb{P}]$ for $d$=2, $p$=5, $h(z_1, z_2)$=$(z_1, z_2, z_1^2\ , z_2^2,\ z_1 z_2)$ and $g(s_1, s_2, s_3, s_4, s_5)$=$(s_5 - s_1 s_2)/((s_3 - s_1^2)(s_4 - s_2^2))^{1/2}$ in the case of the correlation coefficient between $z_1$ and $z_2$, and for $d$=2, $p$=4, $h(z_1, z_2)$=$(z_1, z_2, z_2^2, z_1 z_2)$ and $g(s_1, s_2, s_3, s_4)$=$(s_4 - s_1 s_2)/(s_3 - s_2^2)$ in the case of the regression coefficient of $z_1$ on $z_2$ in the same way as conclusions of Theorems 2 and 3 hold *a.s.* $[\mathbb{P}]$ for $d$=$p$=1, $h(y)$=$y$ and $g(s)$=$s$ in the case of the mean of $y$ in the $2^{nd}$ paragraph of the proof of Theorem 4. Note that $W_i$=$R_{12}[(\overline{Z}_1/S_1^2 - \overline{Z}_2/S_{12})Z_{1i} + (\overline{Z}_2/S_2^2 - \overline{Z}_1/S_{12})Z_{2i} - Z_{1i}^2/2S_1^2 - Z_{2i}^2/2S_2^2 +$

$Z_{1i}Z_{2i}/S_{12}]$ for the correlation coefficient, and $W_i{=}(1/S_2^2)[-\overline{Z}_2Z_{1i} - (\overline{Z}_1 - 2S_{12}\overline{Z}_2/S_2^2)Z_{2i} - S_{12}Z_{2i}^2/S_2^2 + Z_{1i}Z_{2i}]$ for the regression coefficient. Here, $\overline{Z}_1{=}\sum_{i=1}^N Z_{1i}/N$, $\overline{Z}_2{=}\sum_{i=1}^N Z_{2i}/N$, $S_1^2{=}\sum_{i=1}^N Z_{1i}^2\,/N{-}\overline{Z}_1^2$, $S_2^2{=}\sum_{i=1}^N Z_{2i}^2/N{-}\overline{Z}_2^2$, $S_{12}{=}\sum_{i=1}^N Z_{1i}\;Z_{2i}/N - \overline{Z}_1\overline{Z}_2$ and $R_{12}{=}S_{12}/S_1S_2$. Also, note that since $E_{\mathbb{P}}||\epsilon_i||^8 < \infty$, the $\Delta_i^2$'s in Table 3 in the main text can be expressed in terms of superpopulation moments of $(h(Z_{1i}, Z_{2i}), X_i)$ *a.s.* $[\mathbb{P}]$ for both the parameters by SLLN. Further, for the above parameters, we have $\Delta_2^2 - \Delta_1^2{=}cov_{\mathbb{P}}^2(\tilde{W}_i, X_i) > 0$ and $\Delta_7^2 - \Delta_5^2{=}\big\{\mu_1^2 cov_{\mathbb{P}}(\tilde{W}_i, X_i)\big(cov_{\mathbb{P}}(\tilde{W}_i, X_i)cov_{\mathbb{P}}(X_i, 1/X_i) - 2cov_{\mathbb{P}}(\tilde{W}_i, 1/X_i)\big)\big\} - \lambda^2 cov_{\mathbb{P}}^2(\tilde{W}_i, X_i)/\chi\mu_1 - \lambda cov_{\mathbb{P}}^2(\tilde{W}_i, X_i) \leq \big\{\mu_1^2\times cov_{\mathbb{P}}(\tilde{W}_i, X_i)\big(cov_{\mathbb{P}}(\tilde{W}_i, X_i)cov_{\mathbb{P}}(X_i, 1/X_i){-}2cov_{\mathbb{P}}(\tilde{W}_i, 1/X_i)\big)\big\}$ *a.s.* $[\mathbb{P}]$, where $\tilde{W}_i$ is the same as $W_i$ with all finite population moments in the expression of $W_i$ replaced by their corresponding superpopulation moments. Also, from the linear model set up, we have $\big\{\mu_1^2 cov_{\mathbb{P}}(\tilde{W}_i, X_i)\big(cov_{\mathbb{P}}(\tilde{W}_i, X_i)cov_{\mathbb{P}}(X_i, 1/X_i) - 2cov_{\mathbb{P}}(\tilde{W}_i, 1/X_i)\big)\big\}{=}K(\xi - 2\mu_1)((\xi + 2\mu_1)\zeta_1 - 2\zeta_2)$ for some constant $K > 0$ in the case of the correlation coefficient, and $\big\{\mu_1^2 cov_{\mathbb{P}}(\tilde{W}_i, X_i)\times \big(cov_{\mathbb{P}}(\tilde{W}_i, X_i)cov_{\mathbb{P}}(X_i, 1/X_i){-}2cov_{\mathbb{P}}(\tilde{W}_i, 1/X_i)\big)\big\} = K'(\xi{-}2\mu_1)((\xi{+}2\mu_1)\zeta_1{-}2\zeta_2)$ for some constant $K' > 0$ in the case of the regression coefficient. Thus proofs of both the parts of the theorem follow in the same way as the proof of Theorem 6. □

## S4 Comparison of estimators with their bias-corrected versions

In this section, we empirically compare the biased estimators considered in Table 5 in Section 4 of the main text with their bias-corrected versions based on both synthetic and real data used in Section 4. Following the idea in Stefan and Hidiroglou (2022), we consider the bias-corrected jackknife estimator corresponding to each of the biased estimators considered in Table 5 of the main article. For the mean, we consider the bias-corrected jackknife estimators corresponding to the GREG and the PEML estimators under each of SRSWOR, RS and RHC sampling designs, and the Hájek estimator under RS sampling design. On the other hand, for each of the variance, the correlation coefficient and the regression coefficient, we consider the bias-corrected jackknife estimators corresponding to the estimators that are obtained by plugging in the Hájek and the PEML estimators under each of SRSWOR and RS sampling design, and the PEML estimator under RHC sampling design.

Suppose that $s$ is a sample of size $n$ drawn using one of the sampling designs given in Table 5 of the main text. Further, suppose that $s_{-i}$ is the subset of $s$, which excludes the $i^{th}$ unit for any given $i \in s$. Now, for any $i \in s$, let us denote the estimator $g(\hat{\overline{h}})$ constructed based on $s_{-i}$ by $g(\hat{\overline{h}}_{-i})$. Then, we compute the bias-corrected jackknife estimator of $g(\overline{h})$ corresponding to $g(\hat{\overline{h}})$ as $ng(\hat{\overline{h}}) - (n-1)\sum_{i \in s} g(\hat{\overline{h}}_{-i})/n$ (cf. Stefan and Hidiroglou (2022)). Recall from Section 4 in the main article that we draw $I$=1000 samples each of sizes $n$=75, 100 and 125 from some synthetic as well as real datasets using sampling designs mentioned in Table 5 and compute MSEs of the estimators considered in Table 5 based on these samples. Here, we compute MSEs of the above-mentioned bias-corrected jackknife estimators using the same procedure and compare them with the original biased estimators in terms of their MSEs. We observe from the above analyses that for all the parameters considered in Section 4 of the main text, the bias-corrected jackknife estimators become worse than the original biased estimators in the cases of both the synthetic and the real data (see Tables 2 through 6 and 12 through 21 in Sections S5 and S6 below). Despite reducing the biases of the original biased estimators, bias-correction increases the variances of these estimators significantly. This is the reason why the bias-corrected jackknife estimators have larger MSEs than the original biased estimators in the cases of both the synthetic and the real data.

## S5 Analysis based on synthetic data

The results obtained from the analysis carried out in Section 4.1 of the main paper and Section S4 in this supplement are summarized in these sections. Here, we provide some tables that were mentioned in these sections. Tables 2 through 6 contain relative efficiencies of estimators for the mean, the variance, the correlation coefficient and the regression coefficient in the population. Tables 7 through 11 contain the average and the standard deviation of lengths of asymptotically 95% CIs of the above parameters.

Table 2: Relative efficiencies of estimators for mean of $y$.

| Relative efficiency \ Sample size | $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR \| $\hat{\bar{Y}}_{GREG}$, SRSWOR) | 1.049985 | 1.020252 | 1.035038 |
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR \| $\hat{\bar{Y}}_{H}$, RS) | 4.870516 | 5.370899 | 4.987635 |
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR \| $\hat{\bar{Y}}_{HT}$, RS ) | 2.026734 | 2.061607 | 2.027386 |
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR \| $\hat{\bar{Y}}_{PEML}$, RS) | 1.144439 | 1.124697 | 1.170224 |
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR \| $\hat{\bar{Y}}_{GREG}$, RS) | 1.144455 | 1.124975 | 1.170267 |
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR\| $\hat{\bar{Y}}_{RHC}$, RHC ) | 2.022378 | 1.978623 | 2.143015 |
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR \| $\hat{\bar{Y}}_{PEML}$, RHC) | 1.089837 | 1.030332 | 1.094067 |
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR \| $\hat{\bar{Y}}_{GREG}$, RHC) | 1.089853 | 1.030587 | 1.094108 |
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR \| [1]$\hat{\bar{Y}}_{BCPEML}$, SRSWOR) | 1.050461 | 1.021275 | 1.038282 |
| RE($\hat{\bar{Y}}_{GREG}$, SRSWOR \| [1]$\hat{\bar{Y}}_{BCGREG}$, SRSWOR) | 1.002649 | 1.003156 | 1.005397 |
| RE($\hat{\bar{Y}}_{H}$, RS \| [1]$\hat{\bar{Y}}_{BCH}$, RS) | 1.036379 | 1.006945 | 1.12841 |
| RE($\hat{\bar{Y}}_{PEML}$, RS \| [1]$\hat{\bar{Y}}_{BCPEML}$, RS) | 1.016953 | 1.013402 | 1.011762 |
| RE($\hat{\bar{Y}}_{GREG}$, RS \| [1]$\hat{\bar{Y}}_{BCGREG}$, RS) | 1.016692 | 1.011597 | 1.011493 |
| RE($\hat{\bar{Y}}_{PEML}$, RHC \| [1]$\hat{\bar{Y}}_{BCPEML}$, RHC) | 1.01914 | 1.02292 | 1.024689 |
| RE($\hat{\bar{Y}}_{GREG}$, RHC \| [1]$\hat{\bar{Y}}_{BCGREG}$, RHC) | 1.011583 | 1.052311 | 1.023058 |

[1] BCPEML=Bias-corrected PEML estimator, BCH=Bias-corrected Hájek estimator, and BCGREG=Bias-corrected GREG estimator.

Table 3: Relative efficiencies of estimators for variance of $y$. Recall from Table 4 in Section 2 that for variance of $y$, $h(y)=(y^2, y)$ and $g(s_1, s_2)=s_1 - s_2^2$.

| Relative efficiency \ Sample size | $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| $g(\hat{\hat{h}}_{H})$, SRSWOR) | 1.0926 | 1.0848 | 1.0419 |
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| $g(\hat{\hat{h}}_{H})$, RS) | 1.0367 | 1.0435 | 1.0226 |
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| $g(\hat{\hat{h}}_{PEML})$, RS) | 1.15067 | 1.136 | 1.1635 |
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| $g(\hat{\hat{h}}_{PEML})$, RHC) | 1.141 | 1.1849 | 1.1631 |
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| [2] BC $g(\hat{\hat{h}}_{PEML})$, SRSWOR) | 1.0208 | 1.01 | 1.0669 |
| RE($g(\hat{\hat{h}}_{H})$, SRSWOR \| [2] BC $g(\hat{\hat{h}}_{H})$, SRSWOR) | 38.642 | 50.009 | 65.398 |
| RE($g(\hat{\hat{h}}_{H})$, RS \| [2] BC $g(\hat{\hat{h}}_{H})$, RS) | 1.0029 | 1.0117 | 1.074 |
| RE($g(\hat{\hat{h}}_{PEML})$, RS \| [2] BC $g(\hat{\hat{h}}_{PEML})$, RS) | 1.0112 | 1.023 | 1.0377 |
| RE($g(\hat{\hat{h}}_{PEML})$, RHC \| [2] BC $g(\hat{\hat{h}}_{PEML})$, RHC) | 1.0141 | 1.015 | 1.0126 |

[2] BC=Bias-corrected.

Table 4: Relative efficiencies of estimators for correlation coefficient between $z_1$ and $z_2$. Recall from Table 4 in Section 2 that for correlation coefficient between $z_1$ and $z_2$, $h(z_1, z_2)=(z_1, z_2, z_1^2, z_2^2, z_1 z_2)$ and $g(s_1, s_2, s_3, s_4, s_5)=(s_5 - s_1 s_2)/((s_3 - s_1^2)(s_4 - s_2^2))^{1/2}$.

| Relative efficiency \ Sample size | $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| $g(\hat{\hat{h}}_{H})$, SRSWOR) | 1.0304 | 1.0274 | 1.0385 |
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| $g(\hat{\hat{h}}_{H})$, RS) | 1.0307 | 1.0838 | 1.0515 |
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| $g(\hat{\hat{h}}_{PEML})$, RS) | 1.0573 | 1.1862 | 1.1081 |
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| $g(\hat{\hat{h}}_{PEML})$, RHC) | 1.0847 | 1.1459 | 1.0911 |
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| [2] BC $g(\hat{\hat{h}}_{PEML})$, SRSWOR) | 89.989 | 95.299 | 123.89 |
| RE($g(\hat{\hat{h}}_{H})$, SRSWOR \| [2] BC $g(\hat{\hat{h}}_{H})$, SRSWOR) | 90.407 | 96.79 | 141.989 |
| RE($g(\hat{\hat{h}}_{H})$, RS \| [2] BC $g(\hat{\hat{h}}_{H})$, RS) | 90.037 | 102.914 | 152.993 |
| RE($g(\hat{\hat{h}}_{PEML})$, RS \| [2] BC $g(\hat{\hat{h}}_{PEML})$, RS) | 95.68 | 98.758 | 158.832 |
| RE($g(\hat{\hat{h}}_{PEML})$, RHC \| [2] BC $g(\hat{\hat{h}}_{PEML})$, RHC) | 86.27 | 120.582 | 125.374 |

Table 5: Relative efficiencies of estimators for regression coefficient of $z_1$ on $z_2$. Recall from Table 4 in Section 2 that for regression coefficient of $z_1$ on $z_2$, $h(z_1,z_2)$=$(z_1,z_2,z_2^2,z_1z_2)$ and $g(s_1,s_2,s_3,s_4)$=$(s_4-s_1s_2)/(s_3-s_2^2)$.

| Relative efficiency \ Sample size | $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $g(\hat{\bar{h}}_H)$, SRSWOR) | 1.0389 | 1.0473 | 1.0218 |
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $g(\hat{\bar{h}}_H)$, RS) | 1.0589 | 1.0829 | 1.0827 |
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $g(\hat{\bar{h}}_{PEML})$, RS) | 1.1219 | 1.1334 | 1.2137 |
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $g(\hat{\bar{h}}_{PEML})$, RHC) | 1.2037 | 1.1307 | 1.1399 |
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $^2$ BC $g(\hat{\bar{h}}_{PEML})$, SRSWOR) | 80.64 | 91.707 | 124.476 |
| RE($g(\hat{\bar{h}}_H)$, SRSWOR \| $^2$ BC $g(\hat{\bar{h}}_H)$, SRSWOR) | 79.298 | 89.105 | 123.042 |
| RE($g(\hat{\bar{h}}_{RS})$, RS \| $^2$ BC $g(\hat{\bar{h}}_H)$, RS) | 85.97 | 96.22 | 135.449 |
| RE($g(\hat{\bar{h}}_{PEML})$, RS \| $^2$ BC $g(\hat{\bar{h}}_{PEML})$, RS) | 83.331 | 97.583 | 125.657 |
| RE($g(\hat{\bar{h}}_{PEML})$, RHC \| $^2$ BC $g(\hat{\bar{h}}_{PEML})$, RHC) | 75.343 | 112.619 | 115.594 |

Table 6: Relative efficiencies of estimators for regression coefficient of $z_2$ on $z_1$. Recall from Table 4 in Section 2 that for regression coefficient of $z_2$ on $z_1$, $h(z_1,z_2)$=$(z_2,z_1,z_1^2,z_1z_2)$ and $g(s_1,s_2,s_3,s_4)$=$(s_4-s_1s_2)/(s_3-s_2^2)$.

| Relative efficiency \ Sample size | $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $g(\hat{\bar{h}}_H)$, SRSWOR) | 1.0498 | 1.04 | 1.0301 |
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $g(\hat{\bar{h}}_H)$, RS) | 1.0655 | 1.0652 | 1.0548 |
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $g(\hat{\bar{h}}_{PEML})$, RS) | 1.1073 | 1.1153 | 1.1135 |
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $g(\hat{\bar{h}}_{PEML})$, RHC) | 1.0762 | 1.0905 | 1.1108 |
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $^2$ BC $g(\hat{\bar{h}}_{PEML})$, SRSWOR) | 72.061 | 105.389 | 111.124 |
| RE($g(\hat{\bar{h}}_H)$, SRSWOR \| $^2$ BC $g(\hat{\bar{h}}_H)$, SRSWOR) | 69.114 | 108.837 | 118.675 |
| RE($g(\hat{\bar{h}}_H)$, RS \| $^2$ BC $g(\hat{\bar{h}}_H)$, RS) | 69.16 | 115.113 | 144.811 |
| RE($g(\hat{\bar{h}}_{PEML})$, RS \| $^2$ BC $g(\hat{\bar{h}}_{PEML})$, RS) | 72.448 | 127.387 | 131.558 |
| RE($g(\hat{\bar{h}}_{PEML})$, RHC \| $^2$ BC $g(\hat{\bar{h}}_{PEML})$, RHC) | 90.132 | 104.121 | 148.139 |

Table 7: Average and standard deviation of lengths of asymptotically 95% CIs for mean of $y$.

| Estimator and sampling design based on which CI is constructed \ Sample size | Average length (Standard deviation) $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| $\hat{\overline{Y}}_H$, SRSWOR | 536.821<br>(11.357) | 538.177<br>(9.0784) | 539.218<br>(6.8211) |
| [3] $\hat{\overline{Y}}_{PEML}$, SRSWOR | 44.824<br>(3.7002) | 38.81<br>(2.7727) | 34.648<br>(2.2055) |
| $\hat{\overline{Y}}_{HT}$, RS | 689.123<br>(7.8452) | 597.999<br>(5.7176) | 535.951<br>(4.8422) |
| $\hat{\overline{Y}}_H$, RS | 102.611<br>(10.969) | 87.915<br>(8.453) | 59.98307<br>(6.5828) |
| [3] $\hat{\overline{Y}}_{PEML}$, RS | 345.956<br>(654.77) | 115.944<br>(265.93) | 78.711<br>(1041.2) |
| $\hat{\overline{Y}}_{RHC}$, RHC | 848.033<br>(6.8489) | 624.881<br>(4.9609) | 541.421<br>(4.0927) |
| [3] $\hat{\overline{Y}}_{PEML}$, RHC | 64.573<br>(715.16) | 56.531<br>(275.11) | 50.601<br>(651.31) |

[3] It is to be noted that in the cases of PEML and GREG estimators under any given sampling design, we have the same asymptotic MSE and hence the same asymptotic CI. Therefore, the average and the standard deviation of CIs are not reported for the GREG estimator.

Table 8: Average and standard deviation of lengths of asymptotically 95% CIs for variance of $y$. Recall from Table 4 in Section 2 that for variance of $y$, $h(y_1)$=$(y^2, y)$ and $g(s_1, s_2)$=$s_1 - s_2^2$.

| Estimator and sampling design based on which CI is constructed \ Sample size | Average length (Standard deviation) $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| $g(\hat{\bar{h}}_H)$, SRSWOR | 1010775<br>(34245.5) | 878689.4<br>(26373.9) | 786228<br>(20414.5) |
| $g(\hat{\bar{h}}_{PEML})$, SRSWOR | 29432.4<br>(6076.97) | 25929<br>(4441.2) | 23422<br>(3526.8) |
| $g(\hat{\bar{h}}_H)$, RS | 444594.4<br>(44701.7) | 434160.7<br>(31965.2) | 239065<br>(26739.6) |
| $g(\hat{\bar{h}}_{PEML})$, RS | 1152403<br>(9083944) | 1290084<br>(869339.1) | 235909.1<br>(1183961) |
| $g(\hat{\bar{h}}_{PEML})$, RHC | 1031407<br>(7311193) | 895639<br>(1530759) | 801178.9<br>(417582.9) |

Table 9: Average and standard deviation of lengths of asymptotically 95% CIs for correlation coefficient between $z_1$ and $z_2$. Recall from Table 4 in Section 2 that for correlation coefficient between $z_1$ and $z_2$, $h(z_1, z_2)$=$(z_1, z_2, z_1^2, z_2^2, z_1 z_2)$ and $g(s_1, s_2, s_3, s_4, s_5)$=$(s_5 - s_1 s_2)/((s_3 - s_1^2)(s_4 - s_2^2))^{1/2}$.

| | Average length (Standard deviation) | | |
|---|---|---|---|
| Sample size / Estimator and sampling design based on which CI is constructed | $n$=75 | $n$=100 | $n$=125 |
| $g(\hat{\bar{h}}_H)$, SRSWOR | 8.2191<br>(2.429) | 8.0909<br>(1.889) | 8.0897<br>(1.449) |
| $g(\hat{\bar{h}}_{PEML})$, SRSWOR | 0.2542<br>(0.0467) | 0.2575<br>(0.0365) | 0.2583<br>(0.0294) |
| $g(\hat{\bar{h}}_H)$, RS | 4.6847<br>(2.555) | 3.3135<br>(1.884) | 1.3942<br>(1.421) |
| $g(\hat{\bar{h}}_{PEML})$, RS | 5.0473<br>(162.9) | 4.3229<br>(17.19) | 3.1306<br>(21.04) |
| $g(\hat{\bar{h}}_{PEML})$, RHC | 8.3174<br>(15.82) | 8.3898<br>(41.88) | 8.3514<br>(19.62) |

Table 10: Average and standard deviation of lengths of asymptotically 95% CIs for regression coefficient of $z_1$ on $z_2$. Recall from Table 4 in Section 2 that for regression coefficient of $z_1$ on $z_2$, $h(z_1, z_2)=(z_1, z_2, z_2^2, z_1 z_2)$ and $g(s_1, s_2, s_3, s_4)=(s_4-s_1 s_2)/(s_3-s_2^2)$.

| Estimator and sampling design based on which CI is constructed \ Sample size | Average length (Standard deviation) $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| $g(\hat{\bar{h}}_H)$, SRSWOR | 5.9565 (2.013) | 5.068 (1.514) | 4.4818 (1.135) |
| $g(\hat{\bar{h}}_{PEML})$, SRSWOR | 0.2596 (0.0429) | 0.2251 (0.0324) | 0.2032 (0.025) |
| $g(\hat{\bar{h}}_H)$, RS | 3.0488 (2.178) | 1.469 (1.517) | 1.1532 (1.171) |
| $g(\hat{\bar{h}}_{PEML})$, RS | 3.6477 (19.09) | 1.8558 (4.697) | 1.4023 (4.672) |
| $g(\hat{\bar{h}}_{PEML})$, RHC | 6.111 (25.16) | 5.1324 (38.36) | 4.6658 (11.17) |

Table 11: Average and standard deviation of lengths of asymptotically 95% CIs for regression coefficient of $z_2$ on $z_1$. Recall from Table 4 in Section 2 that for regression coefficient of $z_2$ on $z_1$, $h(z_1, z_2)=(z_2, z_1, z_1^2, z_1 z_2)$ and $g(s_1, s_2, s_3, s_4)=(s_4-s_1 s_2)/(s_3-s_2^2)$.

| | Average length (Standard deviation) | | |
|---|---|---|---|
| Estimator and sampling design based on which CI is constructed \ Sample size | $n$=75 | $n$=100 | $n$=125 |
| $g(\hat{\bar{h}}_H)$, SRSWOR | 11.2173<br>(3.238) | 9.6463<br>(2.418) | 8.5885<br>(1.877) |
| $g(\hat{\bar{h}}_{PEML})$, SRSWOR | 0.4198<br>(0.0661) | 0.3652<br>(0.0531) | 0.3307<br>(0.0405) |
| $g(\hat{\bar{h}}_H)$, RS | 6.7247<br>(3.546) | 3.3547<br>(2.539) | 1.7421<br>(1.921) |
| $g(\hat{\bar{h}}_{PEML})$, RS | 11.3373<br>(151.9) | 9.988<br>(31.83) | 8.7889<br>(7.405) |
| $g(\hat{\bar{h}}_{PEML})$, RHC | 19.9049<br>(28.77) | 3.5595<br>(321.7) | 1.8327<br>(8.164) |

## S6 Analysis based on real data

The results obtained from the analyses carried out in Section 4.2 of the main paper and Section S4 in this supplement are summarized in these sections. Here, we provide some scatter plots and tables that were mentioned in these sections. Figures 1 through 4 present scatter plots and least square regression lines between different study and size variables drawn based on all the population values. Tables 12 through 21 contain relative efficiencies of estimators for the mean, the variance, the correlation coefficient and the regression coefficient in the population. Tables 22 through 31 contain the average and the standard deviation of lengths of asymptotically 95% CIs of the above parameters.

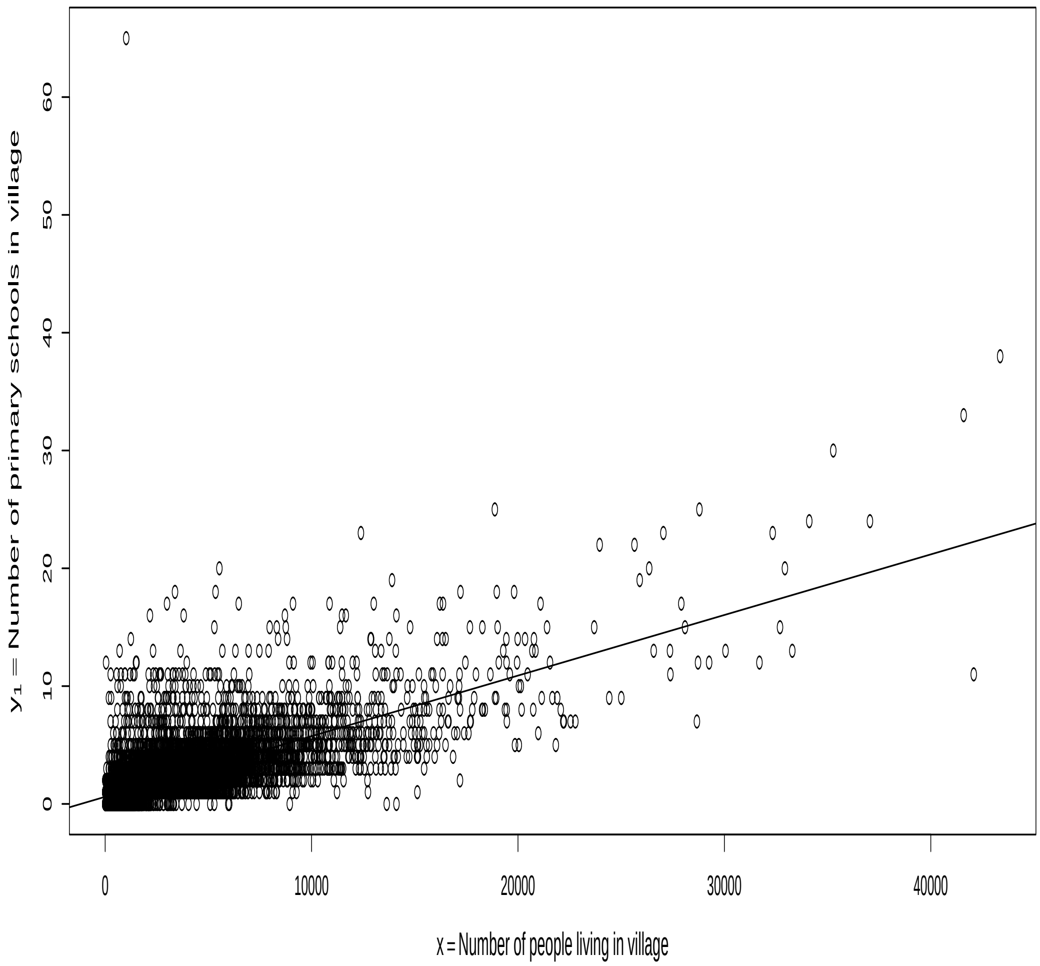


Figure 1: Scatter plot and least square regression line for variables $y_1$ and $x$

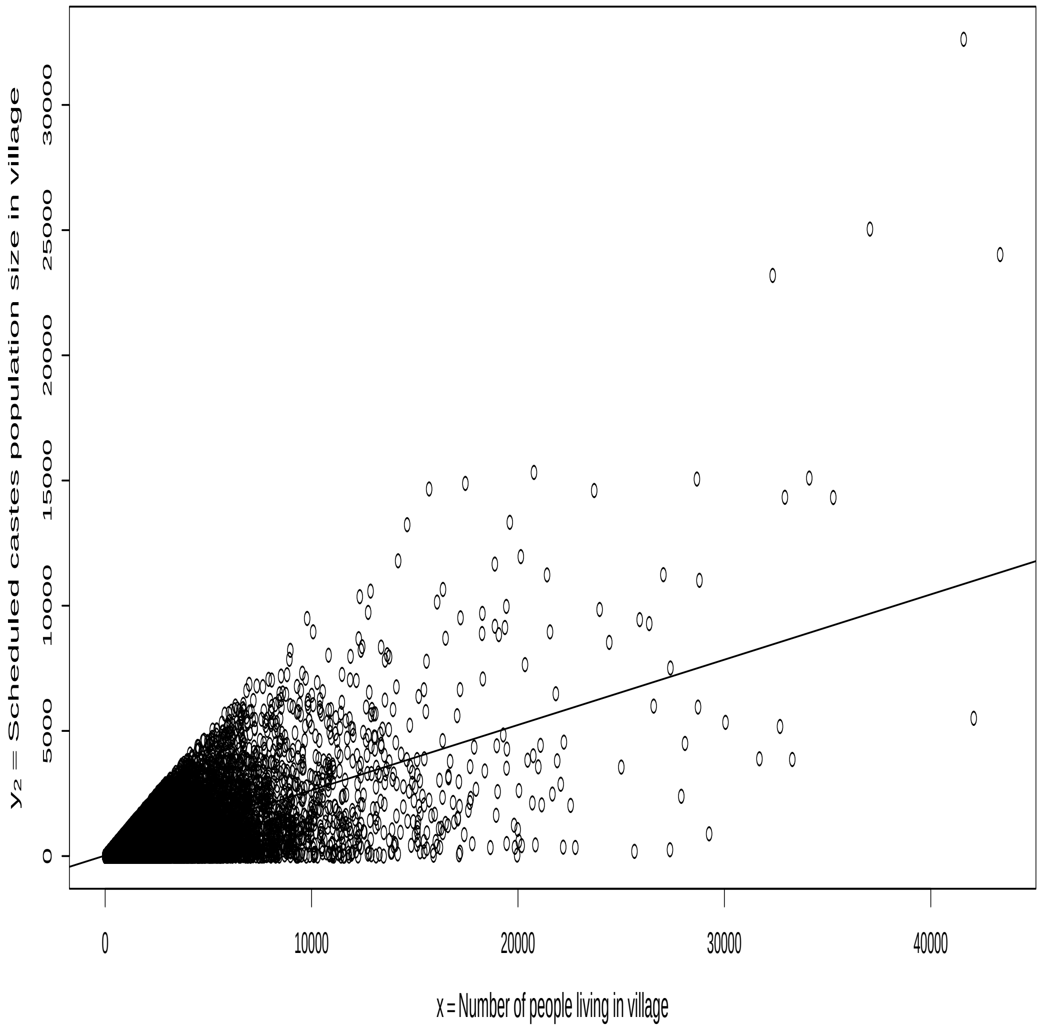


Figure 2: Scatter plot and least square regression line for variables $y_2$ and $x$

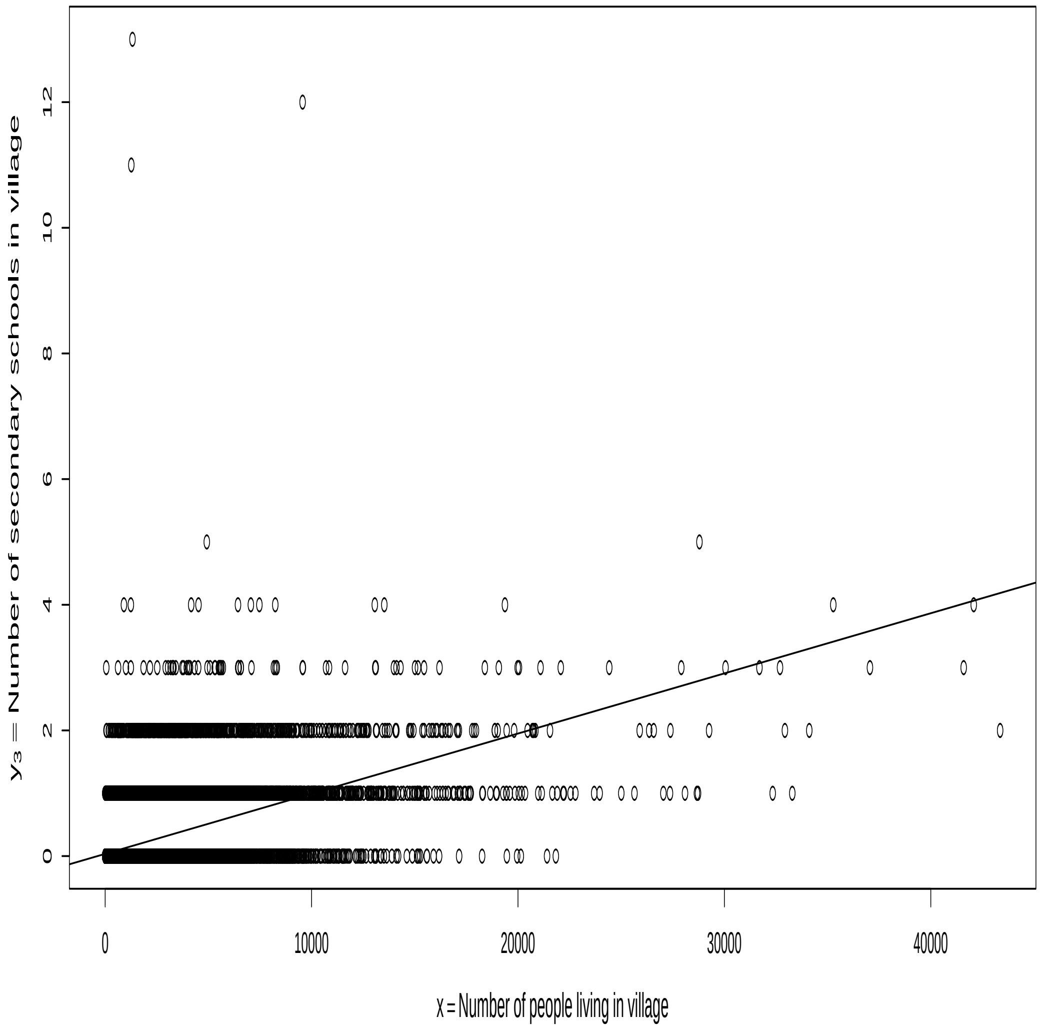


Figure 3: Scatter plot and least square regression line for variables $y_3$ and $x$

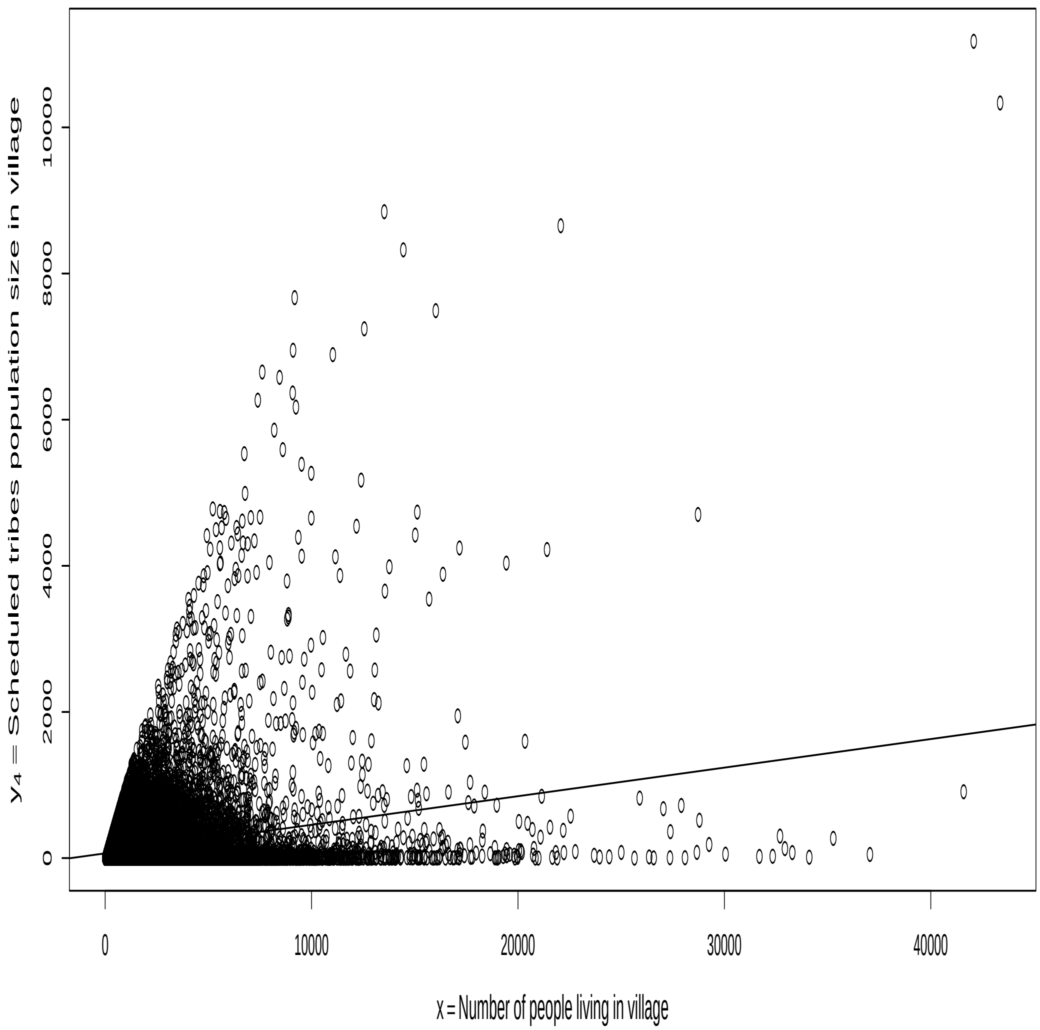


Figure 4: Scatter plot and least square regression line for variables $y_4$ and $x$

Table 12: Relative efficiencies of estimators for mean of $y_1$.

| Relative efficiency \ Sample size | $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR \| $\hat{\bar{Y}}_{GREG}$, SRSWOR) | 1.008215 | 1.005233 | 1.020408 |
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR \| $\hat{\bar{Y}}_{H}$, RS) | 3.503939 | 3.880443 | 4.175886 |
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR \| $\hat{\bar{Y}}_{HT}$, RS) | 1.796937 | 2.182675 | 1.8311 |
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR \| $\hat{\bar{Y}}_{PEML}$, RS) | 1.20961 | 1.228022 | 1.50233 |
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR \| $\hat{\bar{Y}}_{GREG}$, RS) | 1.21831 | 1.237737 | 1.553863 |
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR \| $\hat{\bar{Y}}_{RHC}$, RHC) | 3.274031 | 2.059141 | 2.030995 |
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR \| $\hat{\bar{Y}}_{PEML}$, RHC) | 1.088166 | 1.388563 | 1.51547 |
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR \| $\hat{\bar{Y}}_{GREG}$, RHC) | 1.097934 | 1.398241 | 1.567545 |
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR \| ${}^{1}\hat{\bar{Y}}_{BCPEML}$, SRSWOR) | 1.070226 | 1.019958 | 1.007533 |
| RE($\hat{\bar{Y}}_{GREG}$, SRSWOR \| ${}^{1}\hat{\bar{Y}}_{BCGREG}$, SRSWOR) | 1.146007 | 1.116225 | 1.117507 |
| RE($\hat{\bar{Y}}_{H}$, RS \| ${}^{1}\hat{\bar{Y}}_{BCH}$, RS) | 1.240493 | 1.012969 | 1.155246 |
| RE($\hat{\bar{Y}}_{PEML}$, RS \| ${}^{1}\hat{\bar{Y}}_{BCPEML}$, RS) | 1.374578 | 1.046986 | 1.055930 |
| RE($\hat{\bar{Y}}_{GREG}$, RS \| ${}^{1}\hat{\bar{Y}}_{BCGREG}$, RS) | 1.466647 | 1.138300 | 1.205053 |
| RE($\hat{\bar{Y}}_{PEML}$, RHC \| ${}^{1}\hat{\bar{Y}}_{BCPEML}$, RHC) | 1.566827 | 1.083589 | 1.132790 |
| RE($\hat{\bar{Y}}_{GREG}$, RHC \| ${}^{1}\hat{\bar{Y}}_{BCGREG}$, RHC) | 1.460886 | 1.037045 | 1.028358 |

Table 13: Relative efficiencies of estimators for variance of $y_1$. Recall from Table 4 in Section 2 that for variance of $y_1$, $h(y_1)$=$(y_1^2, y_1)$ and $g(s_1, s_2)$=$s_1 - s_2^2$.

| Relative efficiency \ Sample size | $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $g(\hat{\bar{h}}_H)$, SRSWOR) | 1.3294 | 1.2413 | 1.1476 |
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $g(\hat{\bar{h}}_H)$, RS) | 2.5303 | 1.6656 | 1.5374 |
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $g(\hat{\bar{h}}_{PEML})$, RS) | 3.1642 | 2.4051 | 2.5831 |
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $g(\hat{\bar{h}}_{PEML})$, RHC) | 2.5499 | 4.7704 | 3.0985 |
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $^2$ BC $g(\hat{\bar{h}}_{PEML})$, SRSWOR) | 1.1812 | 1.2736 | 1.8669 |
| RE($g(\hat{\bar{h}}_H)$, SRSWOR \| $^2$ BC $g(\hat{\bar{h}}_H)$, SRSWOR) | 4.3526 | 4.8948 | 6.0349 |
| RE($g(\hat{\bar{h}}_H)$, RS \| $^2$ BC $g(\hat{\bar{h}}_H)$, RS) | 1.115 | 1.1239 | 1.2269 |
| RE($g(\hat{\bar{h}}_{PEML})$, RS \| $^2$ BC $g(\hat{\bar{h}}_{PEML})$, RS) | 1.4373 | 1.1739 | 1.6481 |
| RE($g(\hat{\bar{h}}_{PEML})$, RHC \| $^2$ BC $g(\hat{\bar{h}}_{PEML})$, RHC) | 1.8502 | 1.0186 | 1.0384 |

Table 14: Relative efficiencies of estimators for mean of $y_2$.

| Relative efficiency \ Sample size | $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| RE($\hat{\bar{Y}}_{HT}$, RS \| $\hat{\bar{Y}}_{H}$, RS) | 4.367712 | 4.008655 | 4.463214 |
| RE($\hat{\bar{Y}}_{HT}$, RS \| $\hat{\bar{Y}}_{PEML}$, RS) | 1.148074 | 1.082488 | 1.088804 |
| RE($\hat{\bar{Y}}_{HT}$, RS \| $\hat{\bar{Y}}_{GREG}$, RS) | 1.216958 | 1.115967 | 1.154132 |
| RE($\hat{\bar{Y}}_{HT}$, RS \| $\hat{\bar{Y}}_{RHC}$, RHC) | 1.073138 | 1.03213 | 1.07484 |
| RE($\hat{\bar{Y}}_{HT}$, RS \| $\hat{\bar{Y}}_{PEML}$, RHC) | 1.230884 | 1.0937 | 1.207308 |
| RE($\hat{\bar{Y}}_{HT}$, RS \| $\hat{\bar{Y}}_{GREG}$, RHC) | 1.304737 | 1.127526 | 1.279746 |
| RE($\hat{\bar{Y}}_{HT}$, RS \| $\hat{\bar{Y}}_{PEML}$, SRSWOR) | 2.440441 | 2.305339 | 2.350916 |
| RE($\hat{\bar{Y}}_{HT}$, RS \| $\hat{\bar{Y}}_{GREG}$, SRSWOR) | 2.58687 | 2.376638 | 2.49197 |
| RE($\hat{\bar{Y}}_{H}$, RS \| ${}^{1}\hat{\bar{Y}}_{BCH}$, RS) | 1.252123 | 1.325047 | 1.241809 |
| RE($\hat{\bar{Y}}_{PEML}$, RS \| ${}^{1}\hat{\bar{Y}}_{BCPEML}$, RS) | 1.988105 | 2.146357 | 2.260343 |
| RE($\hat{\bar{Y}}_{GREG}$, RS \| ${}^{1}\hat{\bar{Y}}_{BCGREG}$, RS) | 2.055588 | 2.018015 | 2.287817 |
| RE($\hat{\bar{Y}}_{PEML}$, RHC \| ${}^{1}\hat{\bar{Y}}_{BCPEML}$, RHC) | 1.831377 | 2.083210 | 2.006134 |
| RE($\hat{\bar{Y}}_{GREG}$, RHC \| ${}^{1}\hat{\bar{Y}}_{BCGREG}$, RHC) | 1.925938 | 1.983984 | 2.091003 |
| RE($\hat{\bar{Y}}_{PEML}$, SRSWOR \| ${}^{1}\hat{\bar{Y}}_{BCPEML}$, SRSWOR) | 1.001786 | 1.004973 | 1.060588 |
| RE($\hat{\bar{Y}}_{GREG}$, SRSWOR \| ${}^{1}\hat{\bar{Y}}_{BCGREG}$, SRSWOR) | 1.021103 | 1.008525 | 1.003390 |

Table 15: Relative efficiencies of estimators for variance of $y_2$. Recall from Table 4 in Section 2 that for variance of $y_2$, $h(y_2)=(y_2^2, y_2)$ and $g(s_1, s_2)=s_1 - s_2^2$.

| Relative efficiency \ Sample size | $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| RE($g(\hat{\bar{h}}_H)$, RS \| $g(\hat{\bar{h}}_{PEML})$, RS) | 11.893 | 6.967 | 34.691 |
| RE($g(\hat{\bar{h}}_H)$, RS \| $g(\hat{\bar{h}}_{PEML})$, RHC) | 5.0093 | 19.456 | 21.919 |
| RE($g(\hat{\bar{h}}_H)$, RS \| $g(\hat{\bar{h}}_H)$, SRSWOR) | 9.8232 | 10.27 | 16.763 |
| RE($g(\hat{\bar{h}}_H)$, RS \| $g(\hat{\bar{h}}_{PEML})$, SRSWOR) | 2.4768 | 4.8093 | 6.2264 |
| RE($g(\hat{\bar{h}}_H)$, RS \| $^2$ BC $g(\hat{\bar{h}}_H)$, RS) | 13.301 | 6.3589 | 33.579 |
| RE($g(\hat{\bar{h}}_{PEML})$, RS \| $^2$ BC $g(\hat{\bar{h}}_{PEML})$, RS) | 4.448 | 7.4621 | 7.989 |
| RE($g(\hat{\bar{h}}_{PEML})$, RHC \| $^2$ BC $g(\hat{\bar{h}}_{PEML})$, RHC) | 21.855 | 3.0076 | 11.368 |
| RE($g(\hat{\bar{h}}_H)$, SRSWOR \| $^2$ BC $g(\hat{\bar{h}}_H)$, SRSWOR) | 8.7641 | 5.6119 | 13.7 |
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $^2$ BC $g(\hat{\bar{h}}_{PEML})$, SRSWOR) | 6.2655 | 2.0015 | 6.959 |

Table 16: Relative efficiencies of estimators for correlation coefficient between $y_1$ and $y_3$. Recall from Table 4 in Section 2 that for correlation coefficient between $y_1$ and $y_3$, $h(y_1, y_3)=(y_1, y_3, y_1^2, y_3^2, y_1 y_3)$ and $g(s_1, s_2, s_3, s_4, s_5)=(s_5 - s_1 s_2)/((s_3 - s_1^2)(s_4 - s_2^2))^{1/2}$.

| Relative efficiency \ Sample size | $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $g(\hat{\bar{h}}_H)$, SRSWOR) | 1.0967 | 1.0369 | 1.0374 |
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $g(\hat{\bar{h}}_H)$, RS) | 1.317 | 1.4831 | 1.2561 |
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $g(\hat{\bar{h}}_{PEML})$, RS) | 1.9803 | 1.9874 | 1.8441 |
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $g(\hat{\bar{h}}_{PEML})$, RHC) | 2.0562 | 1.9651 | 1.8541 |
| RE($g(\hat{\bar{h}}_{PEML})$, SRSWOR \| $^2$ BC $g(\hat{\bar{h}}_{PEML})$, SRSWOR) | 23.149 | 51.887 | 45.976 |
| RE($g(\hat{\bar{h}}_H)$, SRSWOR \| $^2$ BC $g(\hat{\bar{h}}_H)$, SRSWOR) | 90.769 | 163.74 | 154.97 |
| RE($g(\hat{\bar{h}}_H)$, RS \| $^2$ BC $g(\hat{\bar{h}}_H)$, RS) | 72.604 | 79.355 | 163.03 |
| RE($g(\hat{\bar{h}}_{PEML})$, RS \| $^2$ BC $g(\hat{\bar{h}}_{PEML})$, RS) | 24.483 | 35.874 | 43.164 |
| RE($g(\hat{\bar{h}}_{PEML})$, RHC \| $^2$ BC $g(\hat{\bar{h}}_{PEML})$, RHC) | 29.189 | 65.949 | 43.13 |

Table 17: Relative efficiencies of estimators for regression coefficient of $y_1$ on $y_3$. Recall from Table 4 in Section 2 that for regression coefficient of $y_1$ on $y_3$, $h(y_1, y_3)=(y_1, y_3, y_3^2, y_1 y_3)$ and $g(s_1, s_2, s_3, s_4)=(s_4 - s_1 s_2)/(s_3 - s_2^2)$.

| Relative efficiency \ Sample size | $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| $g(\hat{\hat{h}}_H)$, SRSWOR) | 1.0298 | 1.0504 | 1.0423 |
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| $g(\hat{\hat{h}}_H)$, RS) | 1.8046 | 1.2304 | 1.3482 |
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| $g(\hat{\hat{h}}_{PEML})$, RS) | 2.2709 | 1.5949 | 1.854 |
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| $g(\hat{\hat{h}}_{PEML})$, RHC) | 1.8719 | 1.5069 | 1.5626 |
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| $^2$ BC $g(\hat{\hat{h}}_{PEML})$, SRSWOR) | 31.789 | 50.26 | 50.107 |
| RE($g(\hat{\hat{h}}_H)$, SRSWOR \| $^2$ BC $g(\hat{\hat{h}}_H)$, SRSWOR) | 236.49 | 119.88 | 222.23 |
| RE($g(\hat{\hat{h}}_H)$, RS \| $^2$ BC $g(\hat{\hat{h}}_H)$, RS) | 63.933 | 77.049 | 184.45 |
| RE($g(\hat{\hat{h}}_{PEML})$, RS \| $^2$ BC $g(\hat{\hat{h}}_{PEML})$, RS) | 31.503 | 44.945 | 263.5 |
| RE($g(\hat{\hat{h}}_{PEML})$, RHC \| $^2$ BC $g(\hat{\hat{h}}_{PEML})$, RHC) | 65.145 | 76.533 | 90.413 |

Table 18: Relative efficiencies of estimators for regression coefficient of $y_3$ on $y_1$. Recall from Table 4 in Section 2 that for regression coefficient of $y_3$ on $y_1$, $h(y_1, y_3)=(y_3, y_1, y_1^2, y_1 y_3)$ and $g(s_1, s_2, s_3, s_4)=(s_4 - s_1 s_2)/(s_3 - s_2^2)$.

| Relative efficiency \ Sample size | $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| $g(\hat{\hat{h}}_H)$, SRSWOR) | 1.0997 | 1.2329 | 1.1529 |
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| $g(\hat{\hat{h}}_H)$, RS) | 1.3948 | 1.3329 | 1.368 |
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| $g(\hat{\hat{h}}_{PEML})$, RS) | 3.6069 | 1.5532 | 1.8035 |
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| $g(\hat{\hat{h}}_{PEML})$, RHC) | 2.5567 | 1.4867 | 1.5335 |
| RE($g(\hat{\hat{h}}_{PEML})$, SRSWOR \| $^2$ BC $g(\hat{\hat{h}}_{PEML})$, SRSWOR) | 26.09 | 29.557 | 32.345 |
| RE($g(\hat{\hat{h}}_H)$, SRSWOR \| $^2$ BC $g(\hat{\hat{h}}_H)$, SRSWOR) | 98.43 | 104.19 | 165.95 |
| RE($g(\hat{\hat{h}}_H)$, RS \| $^2$ BC $g(\hat{\hat{h}}_H)$, RS) | 100.3 | 110.15 | 196.34 |
| RE($g(\hat{\hat{h}}_{PEML})$, RS \| $^2$ BC $g(\hat{\hat{h}}_{PEML})$, RS) | 11.416 | 71.664 | 23.433 |
| RE($g(\hat{\hat{h}}_{PEML})$, RHC \| $^2$ BC $g(\hat{\hat{h}}_{PEML})$, RHC) | 13.268 | 28.198 | 50.571 |

Table 19: Relative efficiencies of estimators for correlation coefficient between $y_2$ and $y_4$. Recall from Table 4 in Section 2 that for correlation coefficient between $y_2$ and $y_4$, $h(y_2, y_4)$=$(y_2, y_4, y_2^2, y_4^2, y_2y_4)$ and $g(s_1, s_2, s_3, s_4, s_5)$=$(s_5 - s_1s_2)/((s_3 - s_1^2)(s_4 - s_2^2))^{1/2}$.

| Relative efficiency \ Sample size | $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| RE($g(\hat{\overline{h}}_H)$, RS \| $g(\hat{\overline{h}}_{PEML})$, RS) | 1.448 | 1.696 | 2.027 |
| RE($g(\hat{\overline{h}}_H)$, RS \| $g(\hat{\overline{h}}_{PEML})$, RHC) | 1.491 | 2.135 | 2.27 |
| RE($g(\hat{\overline{h}}_H)$, RS \| $g(\hat{\overline{h}}_H)$, SRSWOR) | 2.39 | 2.521 | 2.849 |
| RE($g(\hat{\overline{h}}_H)$, RS \| $g(\hat{\overline{h}}_{PEML})$, SRSWOR) | 2.185 | 2.396 | 2.594 |
| RE($g(\hat{\overline{h}}_H)$, RS \| $^2$ BC $g(\hat{\overline{h}}_{PEML})$, RS) | 79.092 | 58.241 | 120.229 |
| RE($g(\hat{\overline{h}}_{PEML})$, RS \| $^2$ BC $g(\hat{\overline{h}}_{PEML})$, RS) | 82.309 | 61.995 | 316.929 |
| RE($g(\hat{\overline{h}}_{PEML})$, RHC \| $^2$ BC $g(\hat{\overline{h}}_{PEML})$, RHC) | 175.22 | 74.847 | 220.74 |
| RE($g(\hat{\overline{h}}_H)$, SRSWOR \| $^2$ BC $g(\hat{\overline{h}}_H)$, SRSWOR) | 87.942 | 36.363 | 97.432 |
| RE($g(\hat{\overline{h}}_{PEML})$, SRSWOR \| $^2$ BC $g(\hat{\overline{h}}_{PEML})$, SRSWOR) | 120.02 | 51.959 | 121.42 |

Table 20: Relative efficiencies of estimators for regression coefficient of $y_2$ on $y_4$. Recall from Table 4 in Section 2 that for regression coefficient of $y_2$ on $y_4$, $h(y_2, y_4)$=$(y_2, y_4, y_4^2, y_2y_4)$ and $g(s_1, s_2, s_3, s_4)$=$(s_4 - s_1s_2)/(s_3 - s_2^2)$.

| Relative efficiency \ Sample size | $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| RE($g(\hat{\overline{h}}_H)$, RS \| $g(\hat{\overline{h}}_{PEML})$, RS) | 1.8158 | 2.3771 | 3.2021 |
| RE($g(\hat{\overline{h}}_H)$, RS \| $g(\hat{\overline{h}}_{PEML})$, RHC) | 2.5985 | 2.6002 | 3.4744 |
| RE($g(\hat{\overline{h}}_H)$, RS \| $g(\hat{\overline{h}}_H)$, SRSWOR) | 3.3278 | 4.5041 | 6.312 |
| RE($g(\hat{\overline{h}}_H)$, RS \| $g(\hat{\overline{h}}_{PEML})$, SRSWOR) | 2.9788 | 3.9417 | 6.0391 |
| RE($g(\hat{\overline{h}}_H)$, RS \| $^2$ BC $g(\hat{\overline{h}}_H)$, RS) | 125.17 | 256.45 | 260.15 |
| RE($g(\hat{\overline{h}}_{PEML})$, RS \| $^2$ BC $g(\hat{\overline{h}}_{PEML})$, RS) | 145.1 | 333.5 | 135.65 |
| RE($g(\hat{\overline{h}}_{PEML})$, RHC \| $^2$ BC $g(\hat{\overline{h}}_{PEML})$, RHC) | 86.93 | 238.32 | 292.89 |
| RE($g(\hat{\overline{h}}_{PEML})$, SRSWOR \| $^2$ BC $g(\hat{\overline{h}}_{PEML})$, SRSWOR) | 93.707 | 101.93 | 121.44 |
| RE($g(\hat{\overline{h}}_H)$, SRSWOR \| $^2$ BC $g(\hat{\overline{h}}_H)$, SRSWOR) | 115.85 | 146.16 | 104.66 |

Table 21: Relative efficiencies of estimators for regression coefficient of $y_4$ on $y_2$. Recall from Table 4 in Section 2 that for regression coefficient of $y_4$ on $y_2$, $h(y_2, y_4)$=$(y_4, y_2, y_2^2, y_2 y_4)$ and $g(s_1, s_2, s_3, s_4)$=$(s_4 - s_1 s_2)/(s_3 - s_2^2)$.

| Relative efficiency \ Sample size | $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| RE($g(\hat{\bar{h}}_H)$, RS \| $g(\hat{\bar{h}}_{PEML})$, RS) | 1.3146 | 1.6055 | 1.937 |
| RE($g(\hat{\bar{h}}_H)$, RS \| $g(\hat{\bar{h}}_{PEML})$, RHC) | 1.652 | 2.7715 | 2.0362 |
| RE($g(\hat{\bar{h}}_H)$, RS \| $g(\hat{\bar{h}}_H)$, SRSWOR) | 3.8248 | 2.4388 | 3.4371 |
| RE($g(\hat{\bar{h}}_H)$, RS \| $g(\hat{\bar{h}}_{PEML})$, SRSWOR) | 3.1843 | 2.3399 | 3.038 |
| RE($g(\hat{\bar{h}}_H)$, RS \| $^2$ BC $g(\hat{\bar{h}}_H)$, RS) | 47.3317 | 73.749 | 52.592 |
| RE($g(\hat{\bar{h}}_H)$, RS \| $^2$ BC $g(\hat{\bar{h}}_{PEML})$, RS) | 105.87 | 126.42 | 323.82 |
| RE($g(\hat{\bar{h}}_H)$, RS \| $^2$ BC $g(\hat{\bar{h}}_{PEML})$, RHC) | 93.403 | 79.453 | 91.347 |
| RE($g(\hat{\bar{h}}_H)$, RS \| $^2$ BC $g(\hat{\bar{h}}_{PEML})$, SRSWOR) | 530.94 | 173.19 | 191.26 |
| RE($g(\hat{\bar{h}}_H)$, RS \| $^2$ BC $g(\hat{\bar{h}}_H)$, SRSWOR) | 394.29 | 156.27 | 164.7 |

Table 22: Average and standard deviation of lengths of asymptotically 95% CIs for mean of $y_1$.

| Estimator and sampling design based on which CI is constructed \ Sample size | Average length (Standard deviation) | | |
|---|---|---|---|
| | $n$=75 | $n$=100 | $n$=125 |
| $\hat{\overline{Y}}_H$, SRSWOR | 0.7233<br>(0.2304) | 0.7303<br>(0.1885) | 0.7333<br>(0.1431) |
| $^3\,\hat{\overline{Y}}_{PEML}$, SRSWOR | 0.3703<br>(0.1608) | 0.3734<br>(0.1534) | 0.3847<br>(0.1074) |
| $\hat{\overline{Y}}_{HT}$, RS | 0.7738<br>(0.2724) | 0.7735<br>(1.071) | 0.8271<br>(0.2001) |
| $\hat{\overline{Y}}_H$, RS | 0.4345<br>(0.8312) | 0.455<br>(8.807) | 0.5414<br>(0.5479) |
| $^3\,\hat{\overline{Y}}_{PEML}$, RS | 0.6784<br>(0.3945) | 0.7207<br>(12.176) | 0.7896<br>(0.2694) |
| $\hat{\overline{Y}}_{RHC}$, RHC | 0.7415<br>(0.4007) | 0.7716<br>(0.6359) | 0.8014<br>(0.2931) |
| $^3\,\hat{\overline{Y}}_{PEML}$, RHC | 0.4911<br>(0.9865) | 0.5078<br>(0.4992) | 0.5289<br>(0.3594) |

Table 23: Average and standard deviation of lengths of asymptotically 95% CIs for variance of $y_1$. Recall from Table 4 in Section 2 that for variance of $y_1$, $h(y_1)=(y_1^2, y_1)$ and $g(s_1, s_2)=s_1 - s_2^2$.

| Estimator and sampling design based on which CI is constructed \ Sample size | Average length (Standard deviation) $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| $g(\hat{\bar{h}}_H)$, SRSWOR | 5.2879 (8.762) | 4.2111 (9.309) | 4.4304 (6.856) |
| $g(\hat{\bar{h}}_{PEML})$, SRSWOR | 2.7519 (7.181) | 2.9935 (8.622) | 3.0013 (5.952) |
| $g(\hat{\bar{h}}_H)$, RS | 3.5121 (1.345) | 3.1177 (11.37) | 3.1095 (10.88) |
| $g(\hat{\bar{h}}_{PEML})$, RS | 3.7475 (4.041) | 3.939 (16.14) | 3.792 (11.08) |
| $g(\hat{\bar{h}}_{PEML})$, RHC | 3.6365 (14.99) | 3.4972 (8.278) | 3.4158 (10.95) |

Table 24: Average and standard deviation of lengths of asymptotically 95% CIs for mean of $y_2$.

| Estimator and sampling design based on which CI is constructed \ Sample size | Average length (Standard deviation) $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| $\hat{\overline{Y}}_H$, SRSWOR | 312.1 (150.08) | 322.48 (121.86) | 326.36 (93.707) |
| ${}^3\,\hat{\overline{Y}}_{PEML}$, SRSWOR | 243.23 (65.059) | 216.42 (55.256) | 198.11 (44.972) |
| $\hat{\overline{Y}}_{HT}$, RS | 184.98 (24.336) | 160.79 (17.942) | 144.43 (13.89) |
| $\hat{\overline{Y}}_H$, RS | 189.49 (314.18) | 163.19 (209.6) | 145.82 (164.32) |
| ${}^3\,\hat{\overline{Y}}_{PEML}$, RS | 343.6 (60.804) | 300.14 (20.411) | 272.63 (21.998) |
| $\hat{\overline{Y}}_{RHC}$, RHC | 277.91 (16.039) | 240.09 (12.042) | 214.78 (9.2784) |
| ${}^3\,\hat{\overline{Y}}_{PEML}$, RHC | 279.97 (52.788) | 242.43 (58.394) | 217.09 (21.356) |

Table 25: Average and standard deviation of lengths of asymptotically 95% CIs for variance of $y_2$. Recall from Table 4 in Section 2 that for variance of $y_2$, $h(y_2)=(y_2^2, y_2)$ and $g(s_1, s_2)=s_1 - s_2^2$.

| Estimator and sampling design based on which CI is constructed \ Sample size | Average length (Standard deviation) $n$=75 | $n$=100 | $n$=125 |
|---|---|---|---|
| $g(\hat{\bar{h}}_H)$, SRSWOR | 1498664 (3236118) | 1588740 (2694726) | 2418155 (3205532) |
| $g(\hat{\bar{h}}_{PEML})$, SRSWOR | 1035032 (1472036) | 1077345 (1376947) | 1002397 (1573834) |
| $g(\hat{\bar{h}}_H)$, RS | 887813.9 (464853) | 764055.6 (377760) | 684218.5 (298552) |
| $g(\hat{\bar{h}}_{PEML})$, RS | 1385778 (1584677) | 1168689 (1339377) | 1055339 (1177054) |
| $g(\hat{\bar{h}}_{PEML})$, RHC | 1319413 (1473379) | 1134532 (1384754) | 1072290 (1472584) |

Table 26: Average and standard deviation of lengths of asymptotically 95% CIs for correlation coefficient between $y_1$ and $y_3$. Recall from Table 4 in Section 2 that for correlation coefficient between $y_1$ and $y_3$, $h(y_1, y_3)=(y_1, y_3, y_1^2, y_3^2, y_1 y_3)$ and $g(s_1, s_2, s_3, s_4, s_5)=(s_5 - s_1 s_2)/((s_3 - s_1^2)(s_4 - s_2^2))^{1/2}$.

| | Average length (Standard deviation) | | |
|---|---|---|---|
| Estimator and sampling design based on which CI is constructed \ Sample size | $n$=75 | $n$=100 | $n$=125 |
| $g(\hat{\bar{h}}_H)$, SRSWOR | 0.3682 (0.1138) | 0.3753 (0.1039) | 0.3893 (0.0936) |
| $g(\hat{\bar{h}}_{PEML})$, SRSWOR | 0.2747 (0.1095) | 0.2881 (0.1008) | 0.2884 (0.0879) |
| $g(\hat{\bar{h}}_H)$, RS | 0.3351 (0.1652) | 0.3453 (0.0938) | 0.3587 (0.1034) |
| $g(\hat{\bar{h}}_{PEML})$, RS | 592.48 (0.2859) | 260.44 (0.3441) | 469.36 (2.738) |
| $g(\hat{\bar{h}}_{PEML})$, RHC | 3838.4 (1.2271) | 2740.5 (0.1467) | 2238.3 (0.1104) |

Table 27: Average and standard deviation of lengths of asymptotically 95% CIs for regression coefficient of $y_1$ on $y_3$. Recall from Table 4 in Section 2 that for regression coefficient of $y_1$ on $y_3$, $h(y_1, y_3)=(y_1, y_3, y_3^2, y_1 y_3)$ and $g(s_1, s_2, s_3, s_4)=(s_4 - s_1 s_2)/(s_3 - s_2^2)$.

| | Average length (Standard deviation) | | |
|---|---|---|---|
| Estimator and sampling design based on which CI is constructed \ Sample size | $n$=75 | $n$=100 | $n$=125 |
| $g(\hat{\bar{h}}_H)$, SRSWOR | 1.6443<br>(1.223) | 1.781<br>(1.127) | 1.8077<br>(0.8849) |
| $g(\hat{\bar{h}}_{PEML})$, SRSWOR | 1.3984<br>(0.8867) | 1.4239<br>(0.7898) | 1.491<br>(0.6645) |
| $g(\hat{\bar{h}}_H)$, RS | 1.4072<br>(0.6463) | 1.5299<br>(0.4833) | 1.5449<br>(0.4883) |
| $g(\hat{\bar{h}}_{PEML})$, RS | 3240.4<br>(4.3202) | 4938.4<br>(1.659) | 1705.3<br>(2.017) |
| $g(\hat{\bar{h}}_{PEML})$, RHC | 50701.7<br>(2.659) | 17291.2<br>(3.93) | 22245.7<br>(1.51) |

Table 28: Average and standard deviation of lengths of asymptotically 95% CIs for regression coefficient of $y_3$ on $y_1$. Recall from Table 4 in Section 2 that for regression coefficient of $y_3$ on $y_1$, $h(y_1, y_3)$=$(y_3, y_1, y_1^2, y_1 y_3)$ and $g(s_1, s_2, s_3, s_4)$=$(s_4 - s_1 s_2)/(s_3 - s_2^2)$.

| Estimator and sampling design based on which CI is constructed \ Sample size | Average length (Standard deviation) | | |
|---|---|---|---|
| | $n$=75 | $n$=100 | $n$=125 |
| $g(\hat{\bar{h}}_H)$, SRSWOR | 0.1387<br>(0.091) | 0.1449<br>(0.072) | 0.1508<br>(0.0616) |
| $g(\hat{\bar{h}}_{PEML})$, SRSWOR | 0.1015<br>(0.0868) | 0.0994<br>(0.0692) | 0.1002<br>(0.0593) |
| $g(\hat{\bar{h}}_H)$, RS | 0.1305<br>(0.0919) | 0.1379<br>(0.0438) | 0.1447<br>(0.0357) |
| $g(\hat{\bar{h}}_{PEML})$, RS | 113.4<br>(0.1712) | 263.23<br>(0.0725) | 78.782<br>(0.0545) |
| $g(\hat{\bar{h}}_{PEML})$, RHC | 798.95<br>(0.6227) | 490.91<br>(0.0862) | 286.92<br>(0.1107) |

Table 29: Average and standard deviation of lengths of asymptotically 95% CIs for correlation coefficient between $y_2$ and $y_4$. Recall from Table 4 in Section 2 that for correlation coefficient between $y_2$ and $y_4$, $h(y_2, y_4)=(y_2, y_4, y_2^2, y_4^2, y_2y_4)$ and $g(s_1, s_2, s_3, s_4, s_5)=(s_5-s_1s_2)/((s_3-s_1^2)(s_4-s_2^2))^{1/2}$.

| | Average length (Standard deviation) | | |
|---|---|---|---|
| Estimator and sampling design based on which CI is constructed \ Sample size | $n$=75 | $n$=100 | $n$=125 |
| $g(\hat{\bar{h}}_H)$, SRSWOR | 0.3428<br>(0.191) | 0.359<br>(0.1783) | 0.3821<br>(0.1844) |
| $g(\hat{\bar{h}}_{PEML})$, SRSWOR | 0.3088<br>(0.1886) | 0.3279<br>(0.171) | 0.3537<br>(0.1773) |
| $g(\hat{\bar{h}}_H)$, RS | 0.2924<br>(0.1561) | 0.2926<br>(0.1491) | 0.298<br>(0.1568) |
| $g(\hat{\bar{h}}_{PEML})$, RS | 833.87<br>(0.5226) | 300.13<br>(0.4406) | 242.51<br>(0.8658) |
| $g(\hat{\bar{h}}_{PEML})$, RHC | 7593.1<br>(0.4385) | 3526.1<br>(0.4869) | 2390.9<br>(0.2661) |

Table 30: Average and standard deviation of lengths of asymptotically 95% CIs for regression coefficient of $y_2$ on $y_4$. Recall from Table 4 in Section 2 that for regression coefficient of $y_2$ on $y_4$, $h(y_2, y_4)$=$(y_2, y_4, y_4^2, y_2y_4)$ and $g(s_1, s_2, s_3, s_4)$=$(s_4 - s_1s_2)/(s_3 - s_2^2)$.

| | Average length (Standard deviation) | | |
|---|---|---|---|
| Estimator and sampling design based on which CI is constructed \ Sample size | $n$=75 | $n$=100 | $n$=125 |
| $g(\hat{\bar{h}}_H)$, SRSWOR | 1.1188 (1.251) | 1.1117 (1.061) | 1.1566 (1.171) |
| $g(\hat{\bar{h}}_{PEML})$, SRSWOR | 0.9865 (0.9935) | 1.0005 (0.8784) | 1.0534 (0.8758) |
| $g(\hat{\bar{h}}_H)$, RS | 0.8575 (0.6472) | 0.847 (0.5219) | 0.8427 (0.4524) |
| $g(\hat{\bar{h}}_{PEML})$, RS | 1583.8 (1.733) | 1647.2 (1.822) | 1533.9 (1.302) |
| $g(\hat{\bar{h}}_{PEML})$, RHC | 24127.4 (2.05) | 10798.8 (1.468) | 5076.1 (2.385) |

Table 31: Average and standard deviation of lengths of asymptotically 95% CIs for regression coefficient of $y_4$ on $y_2$. Recall from Table 4 in Section 2 that for regression coefficient of $y_4$ on $y_2$, $h(y_2, y_4)$=$(y_4, y_2, y_2^2, y_2 y_4)$ and $g(s_1, s_2, s_3, s_4)$=$(s_4 - s_1 s_2)/(s_3 - s_2^2)$.

| Estimator and sampling design based on which CI is constructed \ Sample size | Average length (Standard deviation) | | |
|---|---|---|---|
| | $n$=75 | $n$=100 | $n$=125 |
| $g(\hat{\bar{h}}_H)$, SRSWOR | 0.1607 (0.2236) | 0.1727 (0.2175) | 0.1682 (0.1744) |
| $g(\hat{\bar{h}}_{PEML})$, SRSWOR | 0.1456 (0.2018) | 0.1586 (0.1868) | 0.1577 (0.1616) |
| $g(\hat{\bar{h}}_H)$, RS | 0.1219 (0.0798) | 0.1232 (0.0663) | 0.1273 (0.0615) |
| $g(\hat{\bar{h}}_{PEML})$, RS | 236.81 (0.3529) | 108.3 (0.1879) | 85.466 (0.3227) |
| $g(\hat{\bar{h}}_{PEML})$, RHC | 1568.1 (0.4045) | 2215.1 (0.197) | 659.3 (0.1416) |

## Bibliography

Berger, Y. G. (1998). Rate of convergence to normal distribution for the Horvitz-Thompson estimator, *J. Statist. Plann. Inference* **67**, 209–226.

Boistard, H., Lopuhaä, H. P. and Ruiz-Gazen, A. (2017). Functional central limit theorems for single-stage sampling designs. *Ann. Stat.* **45**, 1728–1758.

Chen, J. and Sitter, R. R. (1999). A pseudo empirical likelihood approach to the effective use of auxiliary information in complex surveys. *Statist. Sinica* **9**, 385–406.

Hájek, J. (1964). Asymptotic theory of rejective sampling with varying probabilities from a finite population. *Ann. Math. Stat.* **35**, 1491–1523.

Ohlsson, E. (1986). Asymptotic normality of the Rao-Hartley-Cochran estimator: an application of the martingale CLT. *Scand. J. Stat.* **13**, 17–28.

Rao, J. N., Hartley, H. and Cochran, W. (1962). On a simple procedure of unequal probability sampling without replacement. *J. R. Stat. Soc. Ser. B Methodol.* **24**, 482–491.

Stefan, M. and Hidiroglou, M. A. (2022). Jackknife Bias-Corrected Generalized Regression Estimator in Survey Sampling. *Journal of Survey Statistics and Methodology*.

Wang, J. C. and Opsomer, J. D. (2011). On asymptotic normality and variance estimation for nondifferentiable survey estimators. *Biometrika* **98**, 91–106.

Anurag Dey

*Indian Statistical Institute*

E-mail: deyanuragsaltlake64@gmail.com

Probal Chaudhuri

*Indian Statistical Institute*

E-mail:probalchaudhuri@gmail.com