Suplementary Material
Two-Stage Cluster Samples with Ranked Set Sampling Designs

Omer Ozturk
The Ohio State University
Department of Statistics
1958 Neil Avenue, Columbus, OH, 43210, USA

Proof of Corollary 1: We rewrite $\sigma^2_{\bar{D}_2}$ in two different forms

$$\sigma^2_{\bar{D}_2} = -\frac{T^2}{N-1} + \frac{1}{nH}\sum_{h=1}^{H}(T^2_{[h]} - T_{[h,h]}) = \sigma^2_{\bar{D}_1} - G_{\bar{D}_1}$$

and

$$\sigma^2_{\bar{D}_2} = \sigma^2_{\bar{D}_3} - \frac{NT^2}{n(N-1)} + \frac{1}{nH}\sum_{h=1}^{H}\{T^2_{[h]} - T_{[h,h]}\} = \sigma^2_{\bar{D}_3} - G_{\bar{D}_3}.$$

For the proof of $\sigma^2_{\bar{D}_2} \leq \sigma^2_{\bar{D}_1}$, we need to show that $G_{\bar{D}_1} \geq 0$. We first observe from Ozturk (2015) or Patil et al. (1995) that

$$\frac{1}{H^2}\sum_{h=1}^{H}\sum_{h'=1}^{H} T_{[h,h']} = -\frac{T^2}{N-1}. \tag{1}$$

We now write $G_{\bar{D}_1}$ as follows

$$G_{\bar{D}_1} = \frac{T^2}{N-1} + \frac{1}{nH}\sum_{h=1}^{H} T_{[h,h]} = -\frac{1}{H^2}\sum_{h=1}^{H}\sum_{h'=1}^{H} T_{[h,h']} + \frac{1}{nH}\sum_{h=1}^{H} T_{[h,h]}.$$

Since design D_2 is constructed without replacement, all covariances are negative. The quantity $G_{\bar{D}_1}$ then can be bounded below as follows

$$G_{\bar{D}_1} \geq -\frac{1}{H^2}\sum_{h=1}\sum_{h'=1}^{H} T_{[h,h']} + \frac{1}{H^2}\sum_{h=1}^{H} T_{[h,h]} = -\frac{1}{H^2}\sum_{h=1}\sum_{h\neq h'}^{H} T_{[h,h']} \geq 0.$$

For the proof of $\sigma^2_{\bar{D}_2} \leq \sigma^2_{\bar{D}_3}$, we we must show $G_{\bar{D}_3} \geq 0$,

$$\begin{aligned} G_{\bar{D}_3} &= -\frac{NT^2}{n(N-1)} + \frac{1}{nH}\sum_{h=1}^{H}\{T^2_{[h]} - T_{[h,h]}\} = -\{\frac{T^2}{n} - \frac{1}{nH}\sum_{h=1}^{H} T^2_{[h]}\} - \frac{T^2}{n(N-1)} - \frac{1}{nH}\sum_{h=1}^{H} T_{[h,h]} \\ &= -\frac{1}{nH}\sum_{h=1}^{H}\{\bar{y}_{[h]} - \bar{y}_N\}^2 - \frac{T^2}{n(N-1)} - \frac{1}{nH}\sum_{h=1}^{H} T_{[h,h]}. \end{aligned}$$

It is now sufficient to show $-\frac{T^2}{n(N-1)} - \frac{1}{nH}\sum_{h=1}^{H} T_{[h,h]} \leq 0$. Again using the equality in equation (1), we write

$$\begin{aligned} -\frac{T^2}{n(N-1)} - \frac{1}{nH}\sum_{h=1}^{H} T_{[h,h]} &= \frac{1}{nH^2}\sum_{h=1}^{H}\sum_{h'=1}^{H} T_{[h,h']} - \frac{1}{nH}\sum_{h=1}^{H} T_{[h,h]} \\ &\geq \frac{1}{nH}\sum_{h=1}^{H}\sum_{h'=1}^{H} T_{[h,h']} - \frac{1}{nH}\sum_{h=1}^{H} T_{[h,h]} \geq 0. \end{aligned}$$

This completes the proof.

The proof of $\sigma^2_{\bar{D}_1} \leq \sigma^2_{\bar{D}_3}$ is trivial.

Proof of Lemma 2: Using conditional expectation, we write

$$\begin{aligned} E\left(y_{\bar{D}^I_k,\bar{D}^{II}_q}\right) &= E_{D^I_k}\left(E_{D^{II}_q}\left\{y_{\bar{D}^I_k,\bar{D}^{II}_q}|D^I_k\right\}\right), \quad (2) \\ &= E_{D^I_k}\left(\frac{1}{\bar{M}n_1}\sum_{a_i\in D^I_k} M_{a_i} E_{D^{II}_q}\left\{\frac{1}{n_{2,a_i}}\sum_{b_j\in D^{II}_q} y_{a_i,b_j}|a_i\right\}\right), \quad (3) \end{aligned}$$

where $E_{D^I_k}$ indicates that the expected value is computed based on random sample (or design) D^I_k. Conditionally on a_i, Lemma 1 indicates that the sample average of stage two sample is unbiased for the mean of cluster a_i. Using this result, one can write

$$\begin{aligned} E\left(y_{\bar{D}^I_k,\bar{D}^{II}_q}\right) &= E_{D^I_k}\left(\frac{1}{\bar{M}n_1}\sum_{a_i\in D^I_k}\frac{M_{a_i}}{M_{a_i}}\sum_{j=1}^{M_{a_i}} y_{a_i,j}\right) = E_{D^I_k}\left(\frac{1}{\bar{M}n_1}\sum_{a_i\in D^I_k} y_{a_i}\right) \\ &= \frac{1}{\bar{M}N}\sum_{i=1}^{N} y_i = \bar{y}. \end{aligned}$$

This completes the proof of $y_{\bar{D}^I_k,\bar{D}^{II}_q}$. The proof of $t_{\bar{D}^I_k,\bar{D}^{II}_q}$ follows from the fact that $t_{\bar{D}^I_k,\bar{D}^{II}_q} = \bar{M}Ny_{\bar{D}^I_k,\bar{D}^{II}_q}$.

Proof of Theorem 1: We prove the theorem for $k, q = 1, 2, 3$. The proof for $k = 4$ is similar, but uses slightly different notation as in Horvitz-Thompson estimator. We first use conditional variance to write

$$Var(y_{\bar{D}^I_k,\bar{D}^{II}_q}) = E_{D^I_k}\left\{Var_{D^{II}_q}\left(y_{\bar{D}^I_k,\bar{D}^{II}_q}|D^I_k\right)\right\} + Var_{D^I_k}\left\{E_{D^{II}_q}\left(y_{\bar{D}^I_k,\bar{D}^{II}_q}|D^I_k\right)\right\}$$

$$
\begin{aligned}
&= E_{D_k^I}\left\{\frac{1}{\bar{M}^2 n_1^2}\sum_{a_i\in D_k^I} M_{a_i}^2 Var_{D_q^{II}}\left(y_{a_i,\bar{D}_q^{II}}|D_k^I\right)\right\} \\
&+ Var_{D_k^I}\left\{\frac{1}{\bar{M}n_1}\sum_{a_i\in D^I} M_{a_i} E_{D_q^{II}}\left(y_{a_i,\bar{D}_q^{II}}|D_k^I\right)\right\}. \qquad (4)
\end{aligned}
$$

In the above expression, the conditional variance and mean are given by

$$
Var_{D_q^{II}}\left(y_{a_i,\bar{D}_q^{II}}|D_k^I\right) = \sigma_{a_i,\bar{D}_q^{II}}
$$

and

$$
E_{D_q^{II}}\left(y_{a_i,\bar{D}^{II}}|D_k^I\right) = \frac{1}{M_{a_i}}\sum_{j=1}^{M_{a_i}} y_{a_i,j}.
$$

We now insert these quantities in equation (4) to write

$$
Var(y_{\bar{D}_k^I,\bar{D}_q^{II}}) = \frac{1}{\bar{M}^2 n_1} E_{D_k^I}\left\{\frac{1}{n_1}\sum_{a_i\in D_k^I} M_{a_i}^2\sigma_{a_i,\bar{D}_q^{II}}^2\right\} + \frac{1}{\bar{M}^2} Var_{D_k^I}\left\{\frac{1}{n_1}\sum_{a_i\in D_k^I} y_{a_i}\right\}.
$$

The design D_k^I yields an unbiased estimator for sample mean. Then the expected value of the sample mean in the first curly bracket is equal to $\frac{1}{N}\sum_{i=1}^{N} M_i^2\sigma_{i,\bar{D}_q^{II}}^2$ for $k=1,2,3$. For $k=4$, The variance of the expression in the second curly bracket is the variance of sample mean of stage I sample with respect to design D_k^I. Hence it follows from Lemma 1 that it is equal to $\sigma_{\bar{D}_k^I}^2$. By inserting these values we write

$$
Var(y_{\bar{D}_k^I,\bar{D}_q^{II}}) = \frac{1}{\bar{M}^2 n_1 N}\sum_{i=1}^{N} M_i^2\sigma_{i,\bar{D}_q^{II}}^2 + \frac{1}{\bar{M}^2}\sigma_{\bar{D}_k^I}^2.
$$

This competes the proof. The proof of $Var(t_{\bar{D}_k^I,\bar{D}_q^{II}})$ is similar. Hence, it is omitted.

Proof of Corollary 3: We first rewrite $\sigma_y^2(\bar{D}_2^I,\bar{D}_2^{II}$ and $\sigma_t^2(\bar{D}_2^I,\bar{D}_2^{II}$ in a slightly different form

$$
\begin{aligned}
\sigma_y^2(\bar{D}_2^I,\bar{D}_2^{II}) &= \sigma_y^2(\bar{D}_k^I,\bar{D}_q^{II}) - \frac{1}{n_1 N\bar{M}^2}\sum_{i=1}^{N} M_i^2 G_{i,\bar{D}_q^{II}} - \frac{1}{\bar{M}^2} G_{\bar{D}_k^I}; k,q=1,3 \\
\sigma_t^2(\bar{D}_2^I,\bar{D}_2^{II}) &= \sigma_t^2(\bar{D}_k^I,\bar{D}_q^{II}) - \frac{N}{n_1}\sum_{i=1}^{N} M_i^2 G_{i,\bar{D}_q^{II}} - N^2 G_{\bar{D}_k^I}; k,q=1,3,
\end{aligned}
$$

where

$$
\begin{aligned}
G_{i,\bar{D}_1^{II}} &= \frac{T_i^2}{M_i-1} + \frac{1}{n_{2,i}H_{2,i}}\sum_{h=1}^{H_{2,i}} T_{[h,h]|i}, \\
G_{i,D_3^{II}} &= \frac{T_i^2}{n_{2,i}(M_i-1)} + \frac{1}{n_{2,i}H_{2,i}}\sum_{h=1}^{H_{2,i}} T_{[h,h]|i} + \frac{1}{n_{2,i}H_{2,i}}\sum_{h=1}^{H_{2,i}}(\bar{y}_{[h]|i}-\bar{y}_i)^2.
\end{aligned}
$$

For the proof of part (i), we need to show that $G_{i,\bar{D}_q^{II}} \geq 0$ and $G_{\bar{D}_k^I} \geq 0$, for $k, q = 1, 3$. The inequality $G_{\bar{D}_k^I} \geq 0$ is proved in Corollary 1. The proof of $G_{i,\bar{D}_q^{II}} \geq 0$ is similar to the poof of $G_{\bar{D}_k^I} \geq 0$ in Corollary 1.

For the proof of (ii), we need to show

$$\begin{aligned}\sigma_y^2(\bar{D}_1^I, \bar{D}_1^{II}) - \sigma_y^2(\bar{D}_3^I, \bar{D}_3^{II}) &= \frac{1}{n_1 N \bar{M}^2} \sum_{i=1}^{N} M_i^2 \left\{ \frac{\sum_{h=1}^{H_{2,i}} T_{[h]|i}^2}{n_{2,i} H_{2,i}} - \frac{(M_i - n_{2,i}) T_i^2}{n_{2,i}(M_i - 1)} \right\} \\ &+ \frac{1}{\bar{M}^2} \left\{ \frac{1}{n_1 H_1} \sum_{h=1}^{H_1} T_{[h]}^2 - \frac{(N - n_1) T^2}{n_1 (N-1)} \right\} \leq 0\end{aligned}$$

The proof follows from the definition of finite population correction factor.

Proof of Lemma 3: Note that using Lemma 1 we write

$$\eta_1 = \frac{var(y_{\bar{D}_2^I})}{var(\bar{y}_{D_3^I})} = \frac{-T^2/(N-1) + \frac{1}{n_1 H_1} \sum_{h=1}^{2} (T_{[h]}^2 - T_{[h,h]})}{\frac{(N-n_1)T^2}{(N-1)n_1}}.$$

With little algebra, one can establish that

$$\frac{1}{H_1} \sum_{h=1}^{2} (T_{[h]}^2 - T_{[h,h]}) = \frac{\{\eta_1 (N - n_1) + n_1\} T^2}{(N-1)}. \tag{5}$$

In a similar fashion, for η_2 we write

$$\eta_2 = \frac{var(W_1)}{var(W_{SRS})} = \frac{-\sum_{i=1}^{N} T_i^2/(M-1) + \frac{1}{n_2 H_2} \sum_{i=1}^{N} \sum_{h=1}^{H_2} (T_{[h]|i}^2 - T_{[h,h]|i})}{\sum_{i=1}^{N} \frac{(M-n_2)T_i^2}{(M-1)n_2}}.$$

Again with some algebra, we obtain

$$\frac{1}{H_2} \sum_{i=1}^{N} \sum_{h=1}^{H_2} (T_{[h]|i}^2 - T_{[h,h]|i}) = \frac{\{\eta_2 (M - n_2) + n_2\} \sum_{i=1}^{N} T_i^2}{M-1}. \tag{6}$$

With these expression, the variance $\sigma_y^2(\bar{D}_2^I, \bar{D}_2^{II})$ can be written as

$$\sigma_y^2(\bar{D}_2^I, \bar{D}_2^{II}) = \frac{1}{n_1} \{ \frac{\eta_1 MSB}{M} - \frac{\eta_2 MSW}{M} \} + \frac{1}{n_1 n_2} \eta_2 MSW - \frac{\eta_1 MSB}{NM},$$

where MSW and MSB are mean square of within- and between-errors, respectively. Using equations

$$\frac{MSW}{MST} = \frac{(1 - ICC)(NM - 1)}{NM}, \quad \frac{MSB}{MST} = \frac{\{(M-1)ICC + 1\}(NM - 1)}{M(N-1)}$$

the ratio $\sigma_y^2(\bar{D}_2^I, \bar{D}_2^{II})/MST$ can be written as

$$\begin{aligned}\sigma_{\bar{y}}^2 = \frac{\sigma_y^2(\bar{D}_2^I, \bar{D}_2^{II})}{MST} &= \frac{\eta_1 (N - n_1)}{n_1 N M} \frac{\{(M-1)ICC + 1\}(NM - 1)}{M(N-1)} \\ &+ \frac{\eta_2 (M - n_2)}{n_1 n_2} \frac{(1 - ICC)(NM - 1)}{NM^2}.\end{aligned}$$

This completes the proof

Proof of Lemma 4: We condition on $a_i \in D_k^I$. In another words, we compute the conditional expectation given that primary sampling unit a_i is selected in stage I based on design D_k^I. It is not difficult to show that

$$E(A_{a_i,\bar{D}_q^{II}}|a_i) = \begin{cases} \frac{1}{H_{2,a_i}^2}\sum_{h=1}^{H_{2,a_i}} T_{[h]|a_i}^2 & q=1 \\ \frac{1}{H_{2,a_i}^2}\sum_{h=1}^{H_{2,a_i}} T_{[h]|a_i}^2 - \frac{1}{H_{2,a_i}^2}\sum_{h=1}^{H_{2,a_i}} T_{[h,h]|a_i} & q=2, \end{cases}$$

where $T_{[h]|a_i}^2 = \sum_{j=1}^{M_{a_i}} P(y_{a_i,r_{1[h]}} = y_{a_i,j})y_{a_i,j}^2 - \bar{y}_{[h]|a_i}^2$. In a similar fashion, the conditional expectation of $B_{a_i,\bar{D}_2^{II}}$ given $a_i \in D^I$ is given by

$$E(B_{a_i,\bar{D}_2^{II}}|a_i) = \frac{H_{2,a_i}-1}{H_2^2}\sum_{h=1}^{H_{2,a_i}} T_{[h]|a_i}^2 + \frac{1}{H_{2,a_i}}\sum_{h=1}^{H_{2,a_i}}\left\{\bar{y}_{[h]|a_i} - \bar{y}_{a_i}\right\}^2 - \frac{1}{H^2}\sum_{h=1}^{H_{2,a_i}}\sum_{h'\neq h}^{H_{2,a_i}} T_{[h,h']|a_i}.$$

Note that

$$\begin{aligned} E(A_{a_i,\bar{D}_1^{II}}|a_i)/d_{2,a_i} &= \frac{1}{d_{2,a_i}H_{2,a_i}}\frac{1}{H_{2,a_i}}\sum_{h=1}^{H_{2,a_i}} T_{[h]|a_i}^2 = \frac{1}{n_{2,a_i}}\left[T_{a_i}^2 - \frac{1}{H_{2,a_i}}\sum_{h=1}^{H_{2,a_i}}\left\{\bar{y}_{[h]|a_i} - \bar{y}_{a_i}\right\}^2\right] \\ &= \sigma_{a_i,\bar{D}_1^{II}}^2. \end{aligned}$$

which completes the proof of $\hat{\sigma}_{a_i,\bar{D}_1^{II}}^2$.

We now consider

$$\begin{aligned} E(A_{a_i,\bar{D}_2^{II}}|a_i + B_{a_i,\bar{D}_2^{II}}|a_i) &= \frac{1}{H_{2,a_i}}\sum_{h=1}^{H_{2,a_i}} T_{[h]|a_i}^2 + \frac{1}{H_{2,a_i}}\sum_{h=1}^{H_{2,a_i}}\left\{\bar{y}_{[h]|a_i} - \bar{y}_{a_i}\right\}^2 - \frac{1}{H_2^2}\sum_{h=1}^{H_{2,a_i}}\sum_{h'=1}^{H_{2,a_i}} T_{[h,h']|a_i} \\ &= \frac{M_{a_i}}{M_{a_i}-1}T_{a_i}^2. \end{aligned} \quad (7)$$

In the last equation above, we used the fact that

$$\frac{1}{H_2^2}\sum_{h=1}^{H_{2,a_i}}\sum_{h'=1}^{H_{2,a_i}} T_{[h,h']|a_i} = \frac{-T_{a_i}^2}{M_{a_i}-1} \text{ and } \frac{1}{H_{2,a_i}}\sum_{h=1}^{H_{2,a_i}} T_{[h]|a_i}^2 + \frac{1}{H_{2,a_i}}\sum_{h=1}^{H_{2,a_i}}\left\{\bar{y}_{[h]|a_i} - \bar{y}_{a_i}\right\}^2 = T_{a_i}^2.$$

We rewrite $E(A_{a_i,\bar{D}_2^{II}}|a_i$ in a slightly different form

$$\begin{aligned} E(A_{a_i,\bar{D}_2^{II}}|a_i) &= \frac{1}{H_{2,a_i}}\left\{\frac{1}{H_{2,a_i}}\sum_{h=1}^{H_{2,a_i}} T_{[h]|a_i}^2\right\} - \frac{1}{H_{2,a_i}^2}\sum_{h=1}^{H_{2,a_i}} T_{[h,h]|a_i} \\ &= \frac{1}{H_{2,a_i}}\left[T_{a_i}^2 - \frac{1}{H_{2,a_i}}\sum_{h=1}^{H_{2,a_i}}\left\{\bar{y}_{[h]|a_i} - \bar{y}_{a_i}\right\}^2\right] - \frac{1}{H_{2,a_i}^2}\sum_{h=1}^{H_{2,a_i}} T_{[h,h]|a_i}. \end{aligned} \quad (8)$$

Inserting equations (7) and (8) in $\hat{\sigma}_{a_i,\bar{D}_2^{II}}^2$, we write

$$\begin{aligned} E(\hat{\sigma}_{a_i,\bar{D}_2^{II}}^2|a_i) &= \frac{(M_{a_i} - n_{2,a_i} - 1)T_{a_i}^2}{n_{2,a_i}(M_{a_i}-1)} - \frac{1}{n_{2,a_i}H_{2,a_i}}\sum_{h=1}^{H_{2,a_i}}\left\{\bar{y}_{[h]|a_i} - \bar{y}_{a_i}\right\}^2 - \frac{1}{n_{2,a_i}H_{2,a_i}}\sum_{h=1}^{H_{2,a_i}} T_{[h,h]|a_i} \\ &= \sigma_{a_i,\bar{D}_2^{II}}^2. \end{aligned}$$

This completes the proof of $\hat{\sigma}^2_{a_i,\bar{D}_2^{II}}|a_i$.

For the proof $\hat{\sigma}^2_{a_i,D_3^{II}}|a_i$ we observe that

$$
\begin{aligned}
E(C_{a_i,\bar{D}_3^{II}}|a_i) &= \left(1-\frac{n_{2,a_i}}{M_{a_i}}\right)\frac{1}{n_{2,a_i}}E\left\{\frac{1}{n_{2,a_i}-1}\sum_{s_j\in S^{II}}\left(y_{a_i,s_j}-y_{a_i,\bar{S}^{II}}\right)^2|a_i\right\} \\
&= \left(1-\frac{n_{2,a_i}}{M_{a_i}}\right)\frac{M_{a_i}T^2_{a_i}}{(M_{a_i}-1)n_{2,a_i}} = \left(\frac{M_{a_i}-n_{2,a_i}}{M_{a_i}-1}\right)\frac{T^2_{a_i}}{n_{2,a_i}} = \sigma^2_{a_i,\bar{D}_3^{II}}.
\end{aligned}
$$

Proof of Lemma 5: We first look at the expected value $A^*_{\bar{D}_k^I,\bar{D}_q^{II}}$

$$
\begin{aligned}
E(A^*_{\bar{D}_k^I,\bar{D}_q^{II}}) &= E_{\bar{D}_k^I}E_{\bar{D}_q^{II}}(A^*_{\bar{D}_k^I,\bar{D}_q^{II}}|D_k^I) \\
&= E_{D_k^I}\left[\frac{2}{2d_1(d_1-1)H_1^2}\sum_{i=1}^{d_1}\sum_{j\neq i}^{d_1}\sum_{h=1}^{H_1}E_{D_q^{II}}\left\{M^2_{r_{i[h]}}y^2_{r_{i[h]},\bar{D}_q^{II}}|D_k^I\right\}\right] \\
&- E_{D_k^I}\left[\frac{2}{2d_1(d_1-1)H_1^2}\sum_{i=1}^{d_1}\sum_{j\neq i}^{d_1}\sum_{h=1}^{H_1}E_{D_q^{II}}\left\{M_{r_{i[h]}}y_{r_{i[h]},\bar{D}_q^{II}}M_{r_{j[h]}}y_{r_{j[h]},\bar{D}^{II}}|D_k^I\right\}\right].
\end{aligned}
$$

In the above equation, the conditional expectations can be computed as follows

$$
\begin{aligned}
E_{D^{II}}\left\{M^2_{a_i}y_{a_i,\bar{D}^{II}}|a_i\in D_k^I\right\} &= var_{D^{II}}\left\{M_{a_i}y_{a_i,\bar{D}^{II}}|a_i\in D_k^I\right\}+\left(E_{D^{II}}\left\{M_{a_i}y_{a_i,\bar{D}^{II}}|a_i\in D_k^I\right\}\right)^2 \\
&= M^2_{a_i}\sigma^2_{a_i,\bar{D}^{II}}+y^2_{a_i};\quad a_i\in D_k^I;\quad k=1,2
\end{aligned}
$$

and

$$
E_{D_q^{II}}\left\{M_{r_{i[h]}}y_{r_{i[h]},\bar{D}_q^{II}}M_{r_{j[h]}}y_{r_{j[h]},\bar{D}_q^{II}}|D_k^I\right\} = y_{r_{i[h]}}y_{r_{j[h]}};\quad (r_{i[h]},r_{j[h]})\in D_k^I; k=1,2.
$$

We note that even though $y_{r_{i[h]}}$ and $y_{r_{j[h]}}$ are all independent in design D_1^I, they are not independent in design D_2^I. We then first evaluate $E(A^*_{\bar{D}_2^I,\bar{D}_q^{II}})$

$$
\begin{aligned}
E(A^*_{\bar{D}_2^I,\bar{D}_q^{II}}) &= E_{D_2^I}\left\{\frac{2}{2d_1(d_1-1)H_1^2}\sum_{i=1}^{d_1}\sum_{j\neq i}^{d_1}\sum_{h=1}^{H_1}\left(M^2_{r_{i[h]}}\sigma^2_{r_{i[h]},\bar{D}_q^{II}}+y^2_{r_{i[h]}}\right)\right\} \\
&- E_{D_2^I}\left\{\frac{2}{2d_1(d_1-1)H_1^2}\sum_{i=1}^{d_1}\sum_{j\neq i}^{d_1}\sum_{h=1}^{H_1}y_{r_{i[h]}}y_{r_{j[h]}}\right\} \\
&= \frac{1}{H_1N}\sum_{i=1}^{N}M_i^2\sigma^2_{i,\bar{D}_q^{II}}+\frac{1}{H_1^2}\sum_{h=1}^{H_1}E(y^2_{r_{1[h]}})-\frac{1}{H_1^2}\sum_{h=1}^{H_1}\left\{cov(y_{r_{1[h]}},y_{r_{2[h]}})+\{E(y_{r_{1[h]}})\}^2\right\} \\
&= \frac{1}{H_1N}\sum_{i=1}^{N}M_i^2\sigma^2_{i,\bar{D}_q^{II}}+\frac{1}{H_1}\left(T^2-\frac{1}{H_1}\sum_{h=1}^{H_1}(y_{[h]}-\bar{y}_N)^2\right)-\frac{1}{H_1^2}\sum_{h=1}^{H_1}T_{[h,h]} \qquad (9) \\
&= \frac{1}{H_1N}\sum_{i=1}^{N}M_i^2\sigma^2_{i,\bar{D}_q^{II}}+\frac{1}{H_1^2}\sum_{h=1}^{H_1}T^2_{[h]}-\frac{1}{H_1^2}\sum_{h=1}^{H_1}T_{[h,h]}.
\end{aligned}
$$

For design D_1^I $(k=1)$, since the covariances are zero, the above expression reduces to

$$E(A^*_{\bar{D}_1^I,\bar{D}_q^{II}}) \;\; = \;\; \frac{1}{H_1 N}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}_q^{II}} + \frac{1}{H_1}\left(T^2 - \frac{1}{H_1}\sum_{h=1}^{H_1}(\bar{y}_{[h]} - \bar{y}_N)^2\right).$$

Inserting this expression in $\hat{\sigma}^2_{\bar{D}_1^I,\bar{D}_q^{II}}$ we find

$$\begin{aligned} E(\hat{\sigma}_y^2(\bar{D}_1^I,\bar{D}_q^{II})) &= \frac{1}{n_1\bar{M}^2N}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}_q^{II}} + \frac{1}{\bar{M}^2 n_1}\left(T^2 - \frac{1}{H_1}\sum_{h=1}^{H_1}(\bar{y}_{[h]} - \bar{y}_N)^2\right) \\ &= \sigma_y^2(\bar{D}_1^I,\bar{D}_q^{II}). \end{aligned}$$

We now consider the expected value of $B^*_{D_2^I,D_q^{II}}$

$$\begin{aligned} E(B^*_{\bar{D}_2^I,\bar{D}_q^{II}}) &= E_{D_2^I}\left[\frac{1}{d_1^2H_1^2}\sum_{i=1}^{d_1}\sum_{j=1}^{d_1}\sum_{h=1}^{H_1}\sum_{h'\neq h}^{H_1} E_{D_q^{II}}\left\{M^2_{r_{i[h]}} y^2_{r_{i[h]},\bar{D}_q^{II}} | D_1^I\right\}\right] \\ &- E_{D_2^I}\left[\frac{1}{d_1^2H_1^2}\sum_{i=1}^{d_1}\sum_{j=1}^{d_1}\sum_{h=1}^{H_1}\sum_{h'=\neq h}^{H_1} E_{D_q^{II}}\left\{M_{r_{i[h]}} y_{r_{i[h]},\bar{D}_q^{II}} M_{r_{j[h']}} y_{r_{j[h']},\bar{D}^{II}} | D_2^I\right\}\right] \\ &= \frac{H_1-1}{H_1N}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}_q^{II}} + \frac{H_1-1}{H_1^2}\sum_{h=1}^{H_1} E(y^2_{r_{1[h]}}) \\ &- \frac{1}{H_1^2}\sum_{h=1}^{H_1}\sum_{h'\neq h}^{H-1}\left\{cov(y_{r_{1[h]}}, y_{r_{2[h]}}) + \{E(y_{r_{1[h]}})E(y_{r_{1[h']}})\}\right\} \\ &= \frac{H_1-1}{H_1N}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}_q^{II}} + \frac{H_1-1}{H_1^2}\sum_{h=1}^{H_1} E(y^2_{r_{1[h]}}) \\ &- \frac{1}{H_1^2}\sum_{h=1}^{H_1}\sum_{h'\neq h}^{H-1} T_{[h,h']} - \bar{y}_N^2 + \frac{1}{H_1^2}\sum_{h=1}^{H_1}\bar{y}^2_{[h]} \\ &= \frac{H_1-1}{H_1N}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}_q^{II}} + \frac{H_1-1}{H_1^2}\sum_{h=1}^{H_1}\left\{T^2_{[h]} + \bar{y}^2_{[h]}\right\} \\ &- \frac{1}{H_1^2}\sum_{h=1}^{H_1}\sum_{h'\neq h}^{H-1} T_{[h,h']} - \bar{y}_N^2 + \frac{1}{H_1^2}\sum_{h=1}^{H_1}\bar{y}^2_{[h]} \\ &= \frac{H_1-1}{H_1N}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}_q^{II}} + \frac{H_1-1}{H_1^2}\sum_{h=1}^{H_1} T^2_{[h]} \\ &- \frac{1}{H_1^2}\sum_{h=1}^{H_1}\sum_{h'\neq h}^{H_1} T_{[h,h']} + \frac{1}{H}\sum_{h=1}^{H_1}(\bar{y}_{[h]} - \bar{y}_N)^2. \end{aligned}$$

We now add $E(A^*_{\bar{D}_2^I,\bar{D}_q^{II}})$ and $E(B^*_{\bar{D}_2^I,\bar{D}_q^{II}})$ and write

$$E\left(A^*_{\bar{D}_2^I,\bar{D}_q^{II}}) + B^*_{\bar{D}_2^I,\bar{D}_q^{II}}\right) \;\; = \;\; \frac{1}{N}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}_q^{II}} - \frac{1}{H_1^2}\sum_{h=1}^{H_1}\sum_{h'=1}^{H_1} T_{[h,h']}$$

$$
\begin{aligned}
&+ \quad \frac{1}{H_1}\sum_{h=1}^{H_1} T_{[h]}^2 + \frac{1}{H}\sum_{h=1}^{H_1}(y_{[h]} - \bar{y}_N)^2 \\
&= \quad \frac{1}{N}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}_q^{II}} + \frac{T^2}{N-1} + T^2 = \frac{1}{N}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}_q^{II}} + \frac{NT^2}{N-1}.
\end{aligned}
$$

We now insert the above equation and equation (9) in the following expression

$$
\begin{aligned}
E\left(\frac{A^*_{\bar{D}_2^I,\bar{D}_q^{II}}}{d_1\bar{M}^2} - \frac{A^*_{\bar{D}_2^I,\bar{D}_q^{II}} + B^*_{\bar{D}_2^I,\bar{D}_q^{II}}}{N\bar{M}^2}\right) &= \frac{1}{n_1 N\bar{M}^2}\sum_{i=1}^{N} M_i^2\sigma_{i,\bar{D}_q^{II}} - \frac{1}{N^2\bar{M}^2}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}_q^{II}} \\
+ \quad \frac{1}{\bar{M}^2}&\left\{\frac{(N-1-n_1)T^2}{n_1(N-1)} - \frac{1}{n_1H_1}\sum_{h=1}^{H-1}(y_{[h]} - \bar{y}_N)^2 - \frac{1}{n_1 H}\sum_{h1}^{H_1} T_{[h,h]}\right\} \\
= \quad \sigma_y^2(\bar{D}_2^I,\bar{D}_q^{II}) &- \frac{1}{N^2\bar{M}^2}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}_q^{II}}.
\end{aligned}
$$

We observe from Lemma 3 that

$$
E_{D_2^I}\left(\frac{1}{d_1H_1}\sum_{h_1}^{d_1}\sum_{i=1}^{H_1} M^2_{r_{i[h]}}\hat{\sigma}^2_{r_{i[h]},\bar{D}^{II}}\right) = \frac{1}{N}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}^{II}}.
$$

The unbiased estimator of $\sigma^2_{\bar{D}_2^I,\bar{D}^{II_q}}$ is then given by

$$
\hat{\sigma}_y^2(\bar{D}_2^I,\bar{D}_q^{II}) = \frac{A^*_{\bar{D}_2^I,\bar{D}_q^{II}}}{d_1\bar{M}^2} - \frac{A^*_{\bar{R}_2^I,\bar{D}_q^{II}} + B^*_{\bar{D}_2^I,\bar{D}_q^{II}}}{N\bar{M}^2} + \frac{1}{\bar{M}^2 N d_1 H_1}\sum_{i=1}^{d_1}\sum_{h=1}^{H_1} M^2_{r_{i[h]}}\hat{\sigma}^2_{r_{i[h]},\bar{D}_q^{II}}.
$$

For the proof of $\hat{\sigma}^2_{\bar{D}_3^I,\bar{D}_q^{II}}$, we first look at the following expectation

$$
\begin{aligned}
E(s_y^{*2}) &= E_{D_3^I}\left[\frac{1}{n_1-1}\sum_{s_i\in D_3^I}^{n_1} E_{D_q^{II}}\left\{M_{s_i}y_{s_i,\bar{D}_q^{II}} - \frac{1}{n_1}\sum_{s_j\in D_3^I}^{n_1} M_{s_j}y_{s_j,\bar{D}_q^{II}}\right\}^2 |D_3^I\right] \\
&= \frac{n_1}{n_1-1}E_{D_3^I}\left(\frac{1}{n_1}\sum_{s_i\in D_3^I}(M_{s_i}^2 y^2_{s_i,\bar{D}_q^{II}}) + \frac{1}{n_1}\sum_{s_i\in D_3^I} y^2_{s_i}\right) - \frac{n_1}{n_1-1}E(\bar{M}y_{\bar{D}_3^I,\bar{D}_q^{II}})^2 \\
&= \frac{n_1}{n_1-1}\left(\frac{1}{N}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}_q^{II}} + \frac{1}{N}\sum_{i=1}^{N} y_i^2 - \bar{M}^2\sigma^2_{\bar{D}_3^I,\bar{D}_q^{II}} - \bar{y}_N^2\right) \\
&= \frac{n_1}{n_1-1}\left(\frac{1}{N}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}_q^{II}} + T^2 - \bar{M}^2\sigma^2_{\bar{D}_3^I,\bar{D}_q^{II}}\right) \\
&= \frac{n_1}{n_1-1}\left(\frac{1}{N}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}^{II}} + \frac{N-1}{N}S^{*2} - \bar{M}^2\sigma^2_{\bar{D}_3^I,\bar{D}_q^{II}}\right),
\end{aligned}
$$

where $S^{*2} = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{y}_N)^2$. We now consider

$$
E\left((1-\frac{n_1}{N})\frac{s_y^{*2}}{n_1}\right) = \frac{1}{n_1-1}(1-\frac{n_1}{N})\left[\frac{1}{N}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}^{II}} + S^{*2} - \frac{1}{N}S^{*2} - \bar{M}^2\sigma^2_{\bar{D}_3^I,\bar{D}_q^{II}}\right]
$$

$$
\begin{aligned}
&= \frac{n_1}{n_1-1}\left(\frac{1}{Nn_1}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}_q^{II}} + (1-\frac{n_1}{N})\frac{S^{*2}}{n_1}\right) - (1-\frac{n_1}{N})\frac{S^{*2}}{N(n_1-1)} \\
&\quad - \frac{n_1}{N^2(n_1-1)}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}_q^{II}} - \frac{\bar{M}^2}{n_1-1}(1-\frac{n_1}{N})\sigma^2_{\bar{D}_3^I,\bar{D}_q^{II}}. \qquad (10)
\end{aligned}
$$

Note that

$$
\left(\frac{1}{Nn_1}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}^{II}} + (1-\frac{n_1}{N})\frac{S^{*2}}{n_1}\right) = \bar{M}^2\sigma^2_{\bar{D}_3^I,\bar{D}_q^{II}}
$$

and

$$
(1-\frac{n_1}{N})\frac{S^{*2}}{n_1} = \bar{M}^2\sigma^2_{\bar{D}_3^I,\bar{D}_q^{II}} - \frac{1}{Nn_1}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}_q^{II}}.
$$

Inserting these expressions in equation (10) and grouping the similar terms we obtain

$$
\begin{aligned}
E\left((1-\frac{n_1}{N})\frac{s_y^{*2}}{n_1}\right) &= \frac{\bar{M}^2(Nn_1-N+n_1)}{(n_1-1)N}\sigma^2_{\bar{D}_3^I,\bar{D}_q^{II}} - \frac{n_1}{N^2(n_1-1)}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}^{II}} \\
&\quad - \frac{n_1}{N(n_1-1)}\left(\bar{M}^2\sigma^2_{\bar{D}_3^I,\bar{D}_q^{II}} - \frac{1}{Nn_1}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}_q^{II}}\right) \\
&= \bar{M}^2\sigma^2_{\bar{D}_3^I,\bar{D}_q^{II}} - \frac{1}{N^2}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}_q^{II}}.
\end{aligned}
$$

We note that

$$
E\left(\frac{1}{Nn_1}\sum_{i=1}^{n_1} M_{s_i}^2\hat{\sigma}^2_{s_i,\bar{D}^{II}}\right) = \frac{1}{N^2}\sum_{i=1}^{N} M_i^2\sigma^2_{i,\bar{D}^{II}}.
$$

We then conclude that

$$
\hat{\sigma}^2_{\bar{D}_3^I,\bar{D}_q^{II}} = (1-\frac{n_1}{N})\frac{s_y^{*2}}{\bar{M}^2 n_1} + \frac{1}{\bar{M}^2 Nn_1}\sum_{i=1}^{n_1} M_{s_i}^2\hat{\sigma}^2_{s_i,\bar{D}_q^{II}}
$$

is an unbiased estimator for $\sigma^2_{\bar{D}_3^I,\bar{D}_q^{II}}$.

The proof of $\sigma^2_{\bar{D}_4^I,\bar{D}_q^{II}}$ is similar and omitted here.

1 References

Ozturk, O. (2016). Statistical Inference Based on Judgment Post-Stratified Samples in Finite Population. *Survey Methodology*, 42, 239–262.

Patil, G. P. and Sinha, A. K. and Taillie, C. (1995). Finite population correction for ranked set sampling. *Annals of the Institute of Statistical Mathematics*, 47, 621–636.