Supplementary file with Appendices for the paper-
Regression analysis for exponential family data in a finite population setup using two-stage cluster sample

Brajendra C. Sutradhar

Memorial University, St. John's, Canada

Corresponding email address: bsutradh@mun.ca

Appendix A: Derivation of model assisted between and within cluster covariance matrices under Lemma 11

First we consider the between clusters covariance matrix $\boldsymbol{V}_{1\cdot}(\boldsymbol{\beta}, \sigma^2_\gamma)$ from (88), and re-express it as

$$\begin{aligned}
\boldsymbol{V}_{1\cdot}(\boldsymbol{\beta}, \sigma^2_\gamma) &= \frac{1}{K}\left[\sum_{c=1}^{K} \boldsymbol{Z}_c \boldsymbol{Z}_c^\top - \frac{1}{K}\sum_{c=1}^{K} \boldsymbol{Z}_c \sum_{c=1}^{K} \boldsymbol{Z}_c^\top\right] \\
&= \frac{1}{K}\left[\sum_{c=1}^{K} \boldsymbol{Z}_c \boldsymbol{Z}_c^\top - \frac{1}{K}\sum_{c=1}^{K}\sum_{d=1}^{K} \boldsymbol{Z}_c \boldsymbol{Z}_d^\top\right] \\
&= \frac{1}{K}\left[\sum_{c=1}^{K} \left\{\sum_{i=1}^{N_c} \boldsymbol{z}_{ci}\right\}\left\{\sum_{i=1}^{N_c} \boldsymbol{z}_{ci}^\top\right\} - \frac{1}{K}\sum_{c=1}^{K}\sum_{d=1}^{K} \left\{\sum_{i=1}^{N_c} \boldsymbol{z}_{ci}\right\}\left\{\sum_{i=1}^{N_d} \boldsymbol{z}_{di}^\top\right\}\right] \\
&= \frac{1}{K}\left[\sum_{c=1}^{K} \left\{\sum_{i=1}^{N_c}\sum_{j=1}^{N_c} \boldsymbol{z}_{ci}\boldsymbol{z}_{cj}^\top\right\} - \frac{1}{K}\sum_{c=1}^{K}\sum_{d=1}^{K} \left\{\sum_{i=1}^{N_c}\sum_{j=1}^{N_d} \boldsymbol{z}_{ci}\boldsymbol{z}_{dj}^\top\right\}\right] \\
&= \frac{1}{K}\left[\sum_{c=1}^{K} \left\{\sum_{i,j=1}^{N_c} (y_{ci} - \mu_{ci}(\boldsymbol{\beta}, \sigma^2_\gamma))(y_{cj} - \mu_{cj}(\boldsymbol{\beta}, \sigma^2_\gamma))\boldsymbol{a}_{ci}(\cdot)\boldsymbol{a}_{ci}^\top(\cdot)\right\}\right. \\
&- \frac{1}{K}\sum_{c=1}^{K}\sum_{d=1}^{K} \left\{\sum_{i=1}^{N_c}\sum_{j=1}^{N_d} (y_{ci} - \mu_{ci}(\boldsymbol{\beta}, \sigma^2_\gamma))(y_{dj} - \mu_{dj}(\boldsymbol{\beta}, \sigma^2_\gamma))\right. \\
&\times \left.\left. \boldsymbol{a}_{ci}(\cdot)\boldsymbol{a}_{dj}^\top(\cdot)\right\}\right] \qquad (130)
\end{aligned}$$

Notice that because $\mathcal{F}$ based clusters are independent, the model expectation over (130) yields

$$E_M\left[(Y_{ci}-\mu_{ci}(\boldsymbol{\beta},\sigma_\gamma^2))(Y_{dj}-\mu_{dj}(\boldsymbol{\beta},\sigma_\gamma^2))\right] \;=\; 0, \text{ for } c\neq d, \tag{131}$$

whereas, for $c=d$, we obtain

$$E_M\left[(Y_{ci}-\mu_{ci}(\boldsymbol{\beta},\sigma_\gamma^2))(Y_{dj}-\mu_{dj}(\boldsymbol{\beta},\sigma_\gamma^2))\right] \;=\; \sigma_{c,ij}, \text{ by (23)-(24)}. \tag{132}$$

Hence by using (131)-(132) in (130), we obtain the model assisted between cluster covariance matrix, i.e., $\boldsymbol{V}_{(1\cdot)M}(\boldsymbol{\beta},\sigma_\gamma^2) = E_M[\boldsymbol{V}_{1\cdot}(\boldsymbol{\beta},\sigma_\gamma^2)]$, as shown in (108) under the Lemma 11.

Next, we consider the within cluster covariance matrix from (88) and simplify it as

$$\begin{aligned}\boldsymbol{V}_c(\boldsymbol{\beta},\sigma_\gamma^2) &= \frac{1}{N_c}\sum_{i=1}^{N_c}\boldsymbol{z}_{ci}\boldsymbol{z}_{ci}^\top - \frac{2}{N_c(N_c-1)}\sum_{i<j}^{N_c}\boldsymbol{z}_{ci}\boldsymbol{z}_{cj}^\top \\ &= \frac{1}{N_c}\sum_{i=1}^{N_c}(y_{ci}-\mu_{ci}(\boldsymbol{\beta},\sigma_\gamma^2))^2\boldsymbol{a}_{ci}(\cdot)\boldsymbol{a}_{ci}^\top(\cdot) \\ &- \frac{2}{N_c(N_c-1)}\sum_{i<j}^{N_c}[(y_{ci}-\mu_{ci}(\boldsymbol{\beta},\sigma_\gamma^2))(y_{cj}-\mu_{cj}(\boldsymbol{\beta},\sigma_\gamma^2))]\boldsymbol{a}_{ci}(\cdot)\boldsymbol{a}_{cj}^\top(\cdot).\end{aligned} \tag{133}$$

Now the model expectation over (133), by (23)-(24), yields $\boldsymbol{V}_{(c)M}(\boldsymbol{\beta},\sigma_\gamma^2) = E_M[\boldsymbol{V}_c(\boldsymbol{\beta},\sigma_\gamma^2)]$, as shown in (119) under the Lemma 11.

Appendix B: Derivation of var$(S_{2,y})$ under Lemma 13

For $S_{2,y}$ defined in (115), it is indicated in (123) how to compute its variance. Thus we follow (123), and as a first step we compute the TSCS design (D_{s^*})-based variance formula by simplifying each of the two terms given in

$$\text{var}_{D_{s^*}}(S_{2,y}) \;=\; \text{var}_{p_1}E_{p_{2c}}[S_{2,y}] + E_{p_1}\text{var}_{p_{2c}}[S_{2,y}]. \tag{134}$$

Computation of $\mathbf{var}_{p_1} E_{p_{2c}}[S_{2,y}]$:

Because

$$S_{2,y} = \sum_{c=1}^{k}\sum_{i=1}^{n_c} w_{(c,i)\in s^*} z^*_{ci}(\boldsymbol{\beta},\sigma^2_\gamma) + \sum_{c=1}^{k}\sum_{i<j}^{n_c} w_{(c,i)\in s^*}\left(\frac{N_c-1}{n_c-1}\right)\tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma), \quad (135)$$

we take the within cluster expectation following (76) and (77) and write

$$\begin{aligned}\text{var}_{p_1} E_{p_{2c}}[S_{2,y}] &= K^2 \text{var}_{p_1}\left[\frac{1}{k}\sum_{c=1}^{k}\left\{\sum_{i=1}^{N_c} z^*_{ci}(\boldsymbol{\beta},\sigma^2_\gamma) + \sum_{i<j}^{N_c}\tilde{z^*}_{c,ij}(\boldsymbol{\beta},\sigma^2_\gamma)\right\}\right] \\ &= K^2 \text{var}_{p_1}\left[\frac{1}{k}\sum_{c=1}^{k} z^\dagger_c\right], \text{ (say).} \quad (136)\end{aligned}$$

Now write $\bar{z^\dagger} = \frac{1}{K}\sum_{c=1}^{K} {z^\dagger}_c$. Then by similar calculations as in (98), we obtain

$$\text{var}_{p_1}\left[\frac{1}{k}\sum_{c=1}^{k} z^\dagger_c\right] = \left(\frac{K-k}{K}\right)\frac{1}{k}{v^\dagger}_{1\cdot}(\boldsymbol{\beta},\sigma^2_\gamma), \quad (137)$$

where ${v^\dagger}_{1\cdot}(\boldsymbol{\beta},\sigma^2_\gamma) = \frac{1}{K}\sum_{c=1}^{K}\left({z^\dagger}_c - \bar{z^\dagger}\right)^2$, yielding the formula for the variance over the within cluster expectation defined in (136), as

$$\text{var}_{p_1} E_{p_{2c}}(S_{2,y})) = K^2\left(\frac{K-k}{K}\right)\frac{1}{k}{v^\dagger}_{1\cdot}(\boldsymbol{\beta},\sigma^2_\gamma). \quad (138)$$

Computation of $E_{p_1}\mathbf{var}_{p_{2c}}[S_{2,y}]$:

For $S_{2,y}$ given in (135) (see also (115)), we express its second stage design based variance as

$$\begin{aligned} E_{p_1}\text{var}_{p_{2c}}[S_{2,y}] = E_{p_1}\Bigg[(K^2/k^2) \sum_{c=1}^{k} N_c^2 &\left[\text{var}_{p_{2c}}\left\{\frac{1}{n_c}\sum_{i=1}^{n_c} z^*_{ci}(\boldsymbol{\beta},\sigma^2_\gamma)\right\}\right] \\ + &\left[\text{var}_{p_{2c}}\left\{\frac{1}{n_c}\left(\frac{N_c-1}{n_c-1}\right)\sum_{i<j}^{n_c}\tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\right\}\right] \\ + &\, 2\left[\text{cov}_{p_{2c}}\left\{\frac{1}{n_c}\sum_{i=1}^{n_c} z^*_{ci}(\boldsymbol{\beta},\sigma^2_\gamma), \frac{1}{n_c}\left(\frac{N_c-1}{n_c-1}\right)\sum_{i<j}^{n_c}\tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\right\}\right]\Bigg], \quad (139)\end{aligned}$$

and compute the variance and covariances within the square brackets as follows.

(a). Computation of var$_{p2c}\left\{\frac{1}{n_c}\sum_{i=1}^{n_c} z^*_{ci}(\boldsymbol{\beta},\sigma^2_\gamma)\right\}$:

By replacing $z_{ci}(\cdot)$ with $z^*_{ci}(\cdot)$, this variance formula follows from (97) as

$$\begin{aligned}
&\text{var}_{p2c}\left\{\frac{1}{n_c}\sum_{i=1}^{n_c} z^*_{ci}(\boldsymbol{\beta},\sigma^2_\gamma)\right\} = \text{var}_{p2c}\left\{\frac{1}{n_c}\sum_{i=1}^{N_c} \delta_{2,i|c} z^*_{ci}(\boldsymbol{\beta},\sigma^2_\gamma)\right\}\\
=\ & \frac{N_c-n_c}{N_c}\frac{1}{n_c}\left[\frac{1}{N_c}\sum_{i=1}^{N_c}[z^*_{ci}]^2 - \frac{2}{N_c(N_c-1)}\sum_{i<j}^{N_c} {z^*}_{ci}{z^*}_{cj}\right]\\
=\ & \frac{N_c-n_c}{N_c}\frac{1}{n_c}{v^*}_c(\boldsymbol{\beta},\sigma^2_\gamma),\ (say). \qquad (140)
\end{aligned}$$

(b). Computation of var$_{p2c}\left\{\frac{1}{n_c}\left(\frac{N_c-1}{n_c-1}\right)\sum_{i<j}^{n_c} \tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\right\}$

Using the random indicator variables $\delta_{2,i|c}$ and $\delta_{2,j|c}$ as in (69), for pair-wise individuals selection in the sample, we first express the variance formula as

$$\begin{aligned}
&\text{var}_{p2c}\left\{\frac{1}{n_c}\left(\frac{N_c-1}{n_c-1}\right)\sum_{i<j}^{n_c} \tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\right\} = \text{var}_{p2c}\left\{\frac{1}{n_c}\left(\frac{N_c-1}{n_c-1}\right)\sum_{i<j}^{N_c} \delta_{2,i|c}\delta_{2,j|c}\tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\right\}\\
=\ & \frac{(N_c-1)^2}{[n_c(n_c-1)]^2}\left[\sum_{i<j}^{N_c}\text{var}(\delta_{2,i|c}\delta_{2,j|c})\left[\tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\right]^2\right.\\
+\ & \left.\sum_{i<j,k<\ell}^{N_c}\text{cov}[\delta_{2,i|c}\delta_{2,j|c},\delta_{2,k|c}\delta_{2,\ell|c}]\left[\tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\tilde{z^*}_{c,k\ell}(\beta,\sigma^2_\gamma)\right]\right]. \qquad (141)
\end{aligned}$$

We then compute the variances and covariances of the pair-wise indicator variables, as

$$\begin{aligned}
\text{var}(\delta_{2,i|c}\delta_{2,j|c}) &= E[\delta^2_{2,i|c}\delta^2_{2,j|c}] - \Big(E(\delta_{2,i|c}\delta_{2,j|c})\Big)^2 = E[\delta_{2,i|c}\delta_{2,j|c}] - \Big(E(\delta_{2,i|c}\delta_{2,j|c})\Big)^2\\
&= E[\delta_{2,i|c}\delta_{2,j|c}]\Big(1-E[\delta_{2,i|c}\delta_{2,j|c}]\Big)\\
&= \frac{n_c(n_c-1)}{N_c(N_c-1)}\left[1-\frac{n_c(n_c-1)}{N_c(N_c-1)}\right] = \frac{n_c(n_c-1)}{N_c(N_c-1)}g_1(n_c,N_c), \qquad (142)
\end{aligned}$$

and

$$
\text{cov}[\delta_{2,i|c}\delta_{2,j|c},\delta_{2,k|c}\delta_{2,\ell|c}] = \begin{cases} E[\delta_{2,i|c}\delta_{2,j|c}\delta_{2,k|c}\delta_{2,\ell|c}] - E[\delta_{2,i|c}\delta_{2,j|c}]E[\delta_{2,k|c}\delta_{2,\ell|c}] & \text{for } i \neq j \neq k \neq \ell \\ E[\delta_{2,i|c}\delta_{2,j|c}\delta_{2,k|c}] - E[\delta_{2,i|c}\delta_{2,j|c}]E[\delta_{2,k|c}\delta_{2,\ell|c}] & \text{otherwise} \end{cases}
$$

$$
= \begin{cases} \frac{n_c(n_c-1)(n_c-2)(n_c-3)}{N_c(N_c-1)(N_c-2)(N_c-3)} - \left(\frac{n_c(n_c-1)}{N_c(N_c-1)}\right)^2 & \text{for } i \neq j \neq k \neq \ell \\ \frac{n_c(n_c-1)(n_c-2)}{N_c(N_c-1)(N_c-2)} - \left(\frac{n_c(n_c-1)}{N_c(N_c-1)}\right)^2 & \text{otherwise} \end{cases}
$$

$$
= \begin{cases} \frac{n_c(n_c-1)}{N_c(N_c-1)}\left[\frac{(n_c-2)(n_c-3)}{(N_c-2)(N_c-3)} - \frac{n_c(n_c-1)}{N_c(N_c-1)}\right] & \text{for } i \neq j \neq k \neq \ell \\ \frac{n_c(n_c-1)}{N_c(N_c-1)}\left[\frac{(n_c-2)}{(N_c-2)} - \frac{n_c(n_c-1)}{N_c(N_c-1)}\right] & \text{otherwise} \end{cases}
$$

$$
= \begin{cases} \frac{n_c(n_c-1)}{N_c(N_c-1)}\left[g_2(n_c,N_c)\right] & \text{for } i \neq j \neq k \neq \ell \\ \frac{n_c(n_c-1)}{N_c(N_c-1)}\left[g_3(n_c,N_c)\right] & \text{otherwise.} \end{cases} \tag{143}
$$

Next, putting (142) and (143) in (141), one obtains

$$
\text{var}_{p_{2c}}\left\{\frac{1}{n_c}\left(\frac{N_c-1}{n_c-1}\right)\sum_{i<j}^{n_c}\tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\right\} = \frac{(N_c-1)^2}{[n_c(n_c-1)]^2}\frac{n_c(n_c-1)}{N_c(N_c-1)}
$$
$$
\times \quad \left[g_1(n_c,N_c)\Phi_{1,N_c}(y) + g_2(n_c,N_c)\Phi_{2,N_c}(y) + g_3(n_c,N_c)\Phi_{3,N_c}(y)\right], \tag{144}
$$

where

$$
\Phi_{1,N_c}(y) = \sum_{i<j}^{N_c}\left[\tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\right]^2
$$

$$
\Phi_{2,N_c}(y) = \sum_{i\neq j\neq k\neq \ell}^{N_c}\left[\tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\tilde{z^*}_{c,k\ell}(\beta,\sigma^2_\gamma\right]
$$

$$
\Phi_{3,N_c}(y) = \sum_{i<j,k<\ell,i=k}^{N_c}\left[\tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\tilde{z^*}_{c,k\ell}(\beta,\sigma^2_\gamma)\right] + \sum_{i<j,k<\ell,i=\ell}^{N_c}\left[\tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\tilde{z^*}_{c,k\ell}(\beta,\sigma^2_\gamma)\right]
$$
$$
+ \quad \sum_{i<j,k<\ell,j=k}^{N_c}\left[\tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\tilde{z^*}_{c,k\ell}(\beta,\sigma^2_\gamma)\right] + \sum_{i<j,k<\ell,j=\ell}^{N_c}\left[\tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\tilde{z^*}_{c,k\ell}(\beta,\sigma^2_\gamma)\right]. \tag{145}
$$

(c). Computation of $\mathrm{cov}_{p_{2c}} 2\left\{\frac{1}{n_c}\sum_{i=1}^{n_c} z^*_{ci}(\boldsymbol{\beta},\sigma^2_\gamma), \frac{1}{n_c}\left(\frac{N_c-1}{n_c-1}\right)\sum_{i<j}^{n_c} \tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\right\}$

In the fashion similar to that of (b), we first express this covariance using indicator variables, as

$$
\begin{aligned}
& 2\left[\mathrm{cov}_{p_{2c}}\left\{\frac{1}{n_c}\sum_{i=1}^{n_c} z^*_{ci}(\boldsymbol{\beta},\sigma^2_\gamma), \frac{1}{n_c}\left(\frac{N_c-1}{n_c-1}\right)\sum_{i<j}^{n_c} \tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\right\}\right] \\
= \; & 2\frac{1}{n_c^2}\left(\frac{N_c-1}{n_c-1}\right)\mathrm{cov}\left[\sum_{i=1}^{N_c}\delta_{2,i|c} z^*_{ci}(\boldsymbol{\beta},\sigma^2_\gamma), \sum_{i<j}^{N_c}\delta_{2,i|c}\delta_{2,j|c}\tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\right] \\
= \; & 2\frac{1}{n_c^2}\left(\frac{N_c-1}{n_c-1}\right)\left[\sum_{i=1}^{N_c}\sum_{j<k}^{N_c}\mathrm{cov}(\delta_{2,i|c},\delta_{2,j|c}\delta_{2,k|c}) z^*_{ci}(\boldsymbol{\beta},\sigma^2_\gamma)\tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\right] \quad (146)
\end{aligned}
$$

We then compute

$$
\begin{aligned}
\mathrm{cov}[\delta_{2,i|c},\delta_{2,j|c}\delta_{2,k|c}] &= E[\delta_{2,i|c}\delta_{2,j|c}\delta_{2,k|c}] - E[\delta_{2,i|c}]E[\delta_{2,j|c}\delta_{2,k|c}] \\
&= \frac{n_c(n_c-1)(n_c-2)}{N_c(N_c-1)(N_c-2)} - \frac{n_c}{N_c}\frac{n_c(n_c-1)}{N_c(N_c-1)} \\
&= \frac{n_c(n_c-1)}{N_c(N_c-1)} g_4(n_c,N_c), \quad (147)
\end{aligned}
$$

and for $i=j, j<k$ and $i=k, j<k$ we use

$$
\begin{aligned}
\mathrm{cov}[\delta_{2,i|c},\delta_{2,i|c}\delta_{2,k|c}] &= \mathrm{cov}[\delta_{2,i|c},\delta_{2,i|c}\delta_{2,j|c}] \equiv E[\delta_{2,i|c}\delta_{2,j|c}] - E[\delta_{2,i|c}]E[\delta_{2,i|c}\delta_{2,j|c}] \\
&= \frac{n_c(n_c-1)}{N_c(N_c-1)}\left(1-\frac{n_c}{N_c}\right) = \frac{n_c(n_c-1)}{N_c(N_c-1)} g_5(n_c,N_c). \quad (148)
\end{aligned}
$$

By putting (147)-(148) into (146), one obtains

$$
\begin{aligned}
& 2\left[\mathrm{cov}_{p_{2c}}\left\{\frac{1}{n_c}\sum_{i=1}^{n_c} z^*_{ci}(\boldsymbol{\beta},\sigma^2_\gamma), \frac{1}{n_c}\left(\frac{N_c-1}{n_c-1}\right)\sum_{i<j}^{n_c} \tilde{z^*}_{c,ij}(\beta,\sigma^2_\gamma)\right\}\right] \\
= \; & 2\frac{1}{n_c^2}\left(\frac{N_c-1}{n_c-1}\right)\frac{n_c(n_c-1)}{N_c(N_c-1)} \\
\times \; & \left[g_4(n_c,N_c)\Phi_{4,N_c}(y) + g_5(n_c,N_c)\Phi_{5,N_c}(y)\right], \quad (149)
\end{aligned}
$$

where

$$
\Phi_{4,N_c}(y) = \sum_{i\neq j, i\neq k, j<k}^{N_c}\left[z^*_{ci}(\boldsymbol{\beta},\sigma^2_\gamma)\tilde{z^*}_{c,jk}(\beta,\sigma^2_\gamma)\right] \quad (150)
$$

$$\Phi_{5,N_c}(y) = \sum_{i=j,j<k}^{N_c} \left[z^*_{ci}(\boldsymbol{\beta}, \sigma^2_\gamma) \tilde{z^*}_{c,jk}(\beta, \sigma^2_\gamma) \right] + \sum_{i=k,j<k}^{N_c} \left[z^*_{ci}(\boldsymbol{\beta}, \sigma^2_\gamma) \tilde{z^*}_{c,jk}(\beta, \sigma^2_\gamma) \right].$$

Finally, by applying the results from (a), (b), and (c), into (139), and also by taking the first-stage expectation (E_{p_1}), we obtain the formula for $E_{p_1}\mathrm{var}_{p_{2c}}[S_{2,y}]$, as

$$\begin{aligned} E_{p_1}\mathrm{var}_{p_{2c}}[S_{2,y}] = \frac{K}{k}\sum_{c=1}^{K} N_c^2 & \left[\left\{ \frac{N_c - n_c}{N_c} \frac{1}{n_c} {v^*}_c(\boldsymbol{\beta}, \sigma^2_\gamma) \right\} \right. \\ + & \left\{ \frac{1}{n_c N_c} \frac{N_c - 1}{n_c - 1} \left[g_1(n_c, N_c)\Phi_{1,N_c}(\boldsymbol{\beta}, \sigma^2_\gamma) + g_2(n_c, N_c)\Phi_{2,N_c}(\boldsymbol{\beta}, \sigma^2_\gamma) + g_3(n_c, N_c)\Phi_{3,N_c}(\boldsymbol{\beta}, \sigma^2_\gamma) \right] \right\} \\ + & \left. \left\{ \frac{2}{n_c N_c} \left[g_4(n_c, N_c)\Phi_{4,N_c}(\boldsymbol{\beta}, \sigma^2_\gamma) + g_5(n_c, N_c)\Phi_{5,N_c}(\boldsymbol{\beta}, \sigma^2_\gamma) \right] \right\} \right], \end{aligned} \tag{151}$$

which, after taking model assisted expectation (E_M), yields the result in (125) under the Lemma 13. We also need the model assisted expectation of ${v^\dagger}_{1\cdot}(\boldsymbol{\beta}, \sigma^2_\gamma)$, which is used in (124). All these model expected functions are computed as follows:

(i) Computation of ${v^\dagger}_{(1\cdot)M}(\boldsymbol{\beta}, \sigma^2_\gamma)$

Recall the formula for ${v^\dagger}_{1\cdot}(\boldsymbol{\beta}, \sigma^2_\gamma)$ from (137) and write its model assisted formula as

$$\begin{aligned} {v^\dagger}_{(1\cdot)M}(\boldsymbol{\beta}, \sigma^2_\gamma) &= E_M\left[{v^\dagger}_{1\cdot}(\boldsymbol{\beta}, \sigma^2_\gamma) \right] \\ &= \frac{1}{K} E_M \sum_{c=1}^{K} \left({z^\dagger}_c - \bar{z^\dagger} \right)^2 \\ &= \frac{1}{K} E_M \left[\sum_{c=1}^{K} ({z^\dagger}_c)^2 - \frac{1}{K}\left(\sum_{c=1}^{K} z^\dagger_c \right)^2 \right] = \frac{1}{K} E_M \left[\sum_{c=1}^{K} ({z^\dagger}_c)^2 - \frac{1}{K}\left\{ \sum_{c=1}^{K}\sum_{d=1}^{K} z^\dagger_c z^\dagger_d \right\} \right] \\ &= \frac{1}{K}\left[\sum_{c=1}^{K} E_M({z^\dagger}_c)^2 - \frac{1}{K}\left\{ \sum_{c=1}^{K} E_M({z^\dagger}_c)^2 + \sum_{c \neq d}^{K} E_M(z^\dagger_c z^\dagger_d) \right\} \right]. \end{aligned} \tag{152}$$

Now because by (136) and (115),

$$z^\dagger_c = \sum_{i=1}^{N_c} z^*_{ci}(\boldsymbol{\beta}, \sigma^2_\gamma) + \sum_{i<j}^{N_c} \tilde{z^*}_{c,ij}(\boldsymbol{\beta}, \sigma^2_\gamma) \tag{153}$$

$$= \sum_{i=1}^{N_c} \xi_{ci}(\beta, \sigma_\gamma^2)(y_{ci}^2 - \lambda_{c,ii}(\boldsymbol{\beta}, \sigma_\gamma^2)) + \sum_{i<j}^{N_c} \tilde{\xi}_{c,ij}(\beta, \sigma_\gamma^2)(y_{ci}y_{cj} - \lambda_{c,ij}(\boldsymbol{\beta}, \sigma_\gamma^2)),$$

it follows by using the model (M) properties (23)-(24), that

$$E_M(z_c^\dagger) = 0 \Rightarrow E_M(z_c^\dagger z_d^\dagger) = 0, \tag{154}$$

because the clusters (c and d) in the $\mathcal{F}$ are pair-wise independent. Furthermore

$$\begin{aligned} E_M(z_c^\dagger)^2 &= \sum_{i=1}^{N_c}\sum_{j=1}^{N_c} \xi_{ci}(\beta, \sigma_\gamma^2)\xi_{cj}(\beta, \sigma_\gamma^2) E_M \left\{(y_{ci}^2 - \lambda_{c,ii}(\boldsymbol{\beta}, \sigma_\gamma^2))(y_{cj}^2 - \lambda_{c,jj}(\boldsymbol{\beta}, \sigma_\gamma^2))\right\} \\ &+ \sum_{i<j}^{N_c}\sum_{k<\ell}^{N_c} \tilde{\xi}_{c,ij}(\beta, \sigma_\gamma^2)\tilde{\xi}_{c,k\ell}(\beta, \sigma_\gamma^2) E_M \left\{(y_{ci}y_{cj} - \lambda_{c,ij}(\boldsymbol{\beta}, \sigma_\gamma^2))(y_{ck}y_{c\ell} - \lambda_{c,k\ell}(\boldsymbol{\beta}, \sigma_\gamma^2))\right\} \\ &+ \sum_{i=1}^{N_c}\sum_{j<k}^{N_c} \xi_{ci}(\beta, \sigma_\gamma^2)\tilde{\xi}_{c,jk}(\beta, \sigma_\gamma^2) E_M \left\{(y_{ci}^2 - \lambda_{c,ii}(\boldsymbol{\beta}, \sigma_\gamma^2))\right. \\ &\times \left.(y_{cj}y_{ck} - \lambda_{c,jk}(\boldsymbol{\beta}, \sigma_\gamma^2))\right\}, \end{aligned} \tag{155}$$

which by using the fourth order moments from (47)-(48), reduces to

$$\begin{aligned} E_M(z_c^\dagger)^2 &= \sum_{i=1}^{N_c}\sum_{j=1}^{N_c} \xi_{ci}(\beta, \sigma_\gamma^2)\xi_{cj}(\beta, \sigma_\gamma^2)\psi_{c,ij}(\boldsymbol{\beta}, \sigma_\gamma^2) \\ &+ \sum_{i<j}^{N_c}\sum_{k<\ell}^{N_c} \tilde{\xi}_{c,ij}(\beta, \sigma_\gamma^2)\tilde{\xi}_{c,k\ell}(\beta, \sigma_\gamma^2)\omega_{c,ij,k\ell}(\beta, \sigma_\gamma^2) \\ &+ \sum_{i=1}^{N_c}\sum_{j<k}^{N_c} \xi_{ci}(\beta, \sigma_\gamma^2)\tilde{\xi}_{c,jk}(\beta, \sigma_\gamma^2)\omega_{c,ii,jk}(\beta, \sigma_\gamma^2) \\ &= \Upsilon_{c,N_c}(\beta, \sigma_\gamma^2), \text{ (say)}. \end{aligned} \tag{156}$$

Then by using (154) and (156) into (152), one obtains the model assisted between variance as

$$v^\dagger{}_{(1\cdot)M}(\boldsymbol{\beta}, \sigma_\gamma^2) = \frac{1}{K}\left(1 - \frac{1}{K}\right)\sum_{c=1}^{K} \Upsilon_{c,N_c}(\beta, \sigma_\gamma^2). \tag{157}$$

(ii) Computation of $v^*{}_{(c)M}(\boldsymbol{\beta}, \sigma_\gamma^2)$ from (140)

It follows from (140) that

$$v^*{}_c(\boldsymbol{\beta}, \sigma_\gamma^2) \;=\; \frac{1}{N_c}\sum_{i=1}^{N_c}[z^*_{ci}]^2 - \frac{2}{N_c(N_c-1)}\sum_{i<j}^{N_c} z^*{}_{ci}z^*{}_{cj}, \tag{158}$$

where $z^*_{ci}(\cdot) = \xi_{ci}(\beta, \sigma_\gamma^2)(y_{ci}^2 - \lambda_{c,ii}(\boldsymbol{\beta}, \sigma_\gamma^2))$ as in (135)(see also (115)). Now by using the model based fourth order moments, namely var$[Y_{ci}^2]$ and cov$[Y_{ci}^2, Y_{cj}^2]$, we take the model expectation over (158) and obtain

$$\begin{aligned} & v^*{}_{(c)M}(\boldsymbol{\beta}, \sigma_\gamma^2) = E_M\left[v^*{}_c(\boldsymbol{\beta}, \sigma_\gamma^2)\right] \qquad (159) \\ = \;& \frac{1}{N_c}\sum_{i=1}^{N_c} E_M[z^*_{ci}]^2 - \frac{2}{N_c(N_c-1)}\sum_{i<j}^{N_c} E_M[z^*{}_{ci}z^*{}_{cj}] \\ = \;& \frac{1}{N_c}\sum_{i=1}^{N_c} \xi_{ci}^2(\beta, \sigma_\gamma^2)\text{var}[Y_{ci}^2] - \frac{2}{N_c(N_c-1)}\sum_{i<j}^{N_c} \xi_{ci}(\beta, \sigma_\gamma^2)\xi_{cj}(\beta, \sigma_\gamma^2)\text{cov}[Y_{ci}^2, Y_{cj}^2] \\ = \;& \frac{1}{N_c}\sum_{i=1}^{N_c} \xi_{ci}^2(\beta, \sigma_\gamma^2)\psi_{c,ii}(\beta, \sigma_\gamma^2) - \frac{2}{N_c(N_c-1)}\sum_{i<j}^{N_c} \xi_{ci}(\beta, \sigma_\gamma^2)\xi_{cj}(\beta, \sigma_\gamma^2)\psi_{c,ij}(\beta, \sigma_\gamma^2), \end{aligned}$$

by (47).

(iii) Computation of $\{\Phi_{(u,N_c)M}(\boldsymbol{\beta}, \sigma_\gamma^2); u = 1, \ldots, 5\}$ from (145) and (150)

Because, $\tilde{z^*}_{c,ij}(\boldsymbol{\beta}, \sigma_\gamma^2) = \tilde{\xi}_{c,ij}(\beta, \sigma_\gamma^2)(y_{ci}y_{cj} - \lambda_{c,ij}(\boldsymbol{\beta}, \sigma_\gamma^2))$ by (115) (see also (153)), the model based expectations of $\{\Phi_{u,N_c}(\boldsymbol{\beta}, \sigma_\gamma^2); u = 1, \ldots, 5\}$, using the fourth order moments from (48), have the formulas as follows:

$$\begin{aligned} & \Phi_{(1,N_c)M}(\boldsymbol{\beta}, \sigma_\gamma^2) = \sum_{i<j}^{N_c} \tilde{\xi}_{c,ij}^2(\beta, \sigma_\gamma^2)\omega_{c,ij,ij}(\beta, \sigma_\gamma^2) \qquad (160) \\ & \Phi_{(2,N_c)M}(\beta, \sigma_\gamma^2) = \sum_{i\neq j\neq k\neq \ell}^{N_c} \tilde{\xi}_{c,ij}(\beta, \sigma_\gamma^2)\tilde{\xi}_{c,k\ell}(\beta, \sigma_\gamma^2)\omega_{c,ij,k\ell}(\beta, \sigma_\gamma^2) \\ & \Phi_{(3,N_c)M}(\beta, \sigma_\gamma^2) \\ = \;& \sum_{i<j,k<\ell,i=k}^{N_c} \tilde{\xi}_{c,ij}(\beta, \sigma_\gamma^2)\tilde{\xi}_{c,i\ell}(\beta, \sigma_\gamma^2)\omega_{c,ij,i\ell}(\beta, \sigma_\gamma^2) + \sum_{i<j,k<\ell,i=\ell}^{N_c} \tilde{\xi}_{c,ij}(\beta, \sigma_\gamma^2)\tilde{\xi}_{c,ki}(\beta, \sigma_\gamma^2)\omega_{c,ij,ki}(\beta, \sigma_\gamma^2) \end{aligned}$$

$$+ \quad \sum_{i<j,k<\ell,j=k}^{N_c} \tilde{\xi}_{c,ij}(\beta,\sigma_\gamma^2)\tilde{\xi}_{c,j\ell}(\beta,\sigma_\gamma^2)\omega_{c,ij,j\ell}(\beta,\sigma_\gamma^2) + \sum_{i<j,k<\ell,j=\ell}^{N_c} \tilde{\xi}_{c,ij}(\beta,\sigma_\gamma^2)\tilde{\xi}_{c,kj}(\beta,\sigma_\gamma^2)\omega_{c,ij,kj}(\beta,\sigma_\gamma^2)$$

$$\Phi_{(4,N_c)M}(y) = \sum_{i\neq j,i\neq k,j<k}^{N_c} \xi_{ci}(\beta,\sigma_\gamma^2)\tilde{\xi}_{c,jk}(\beta,\sigma_\gamma^2)\omega_{c,ii,jk}(\beta,\sigma_\gamma^2)$$

$$\Phi_{(5,N_c)M}(\beta,\sigma_\gamma^2) = \sum_{i=j,j<k}^{N_c} \xi_{ci}(\beta,\sigma_\gamma^2)\tilde{\xi}_{c,jk}(\beta,\sigma_\gamma^2)\omega_{c,ii,jk}(\beta,\sigma_\gamma^2)$$

$$+ \quad \sum_{i=k,j<k}^{N_c} \xi_{ci}(\beta,\sigma_\gamma^2)\tilde{\xi}_{c,ji}(\beta,\sigma_\gamma^2)\omega_{c,ii,ji}(\beta,\sigma_\gamma^2).$$