Two Stage Cluster Sampling and Random Effects
Ed Stanek

1. FINITE POPULATION

Frames

Labels

Expected Response

Response

Notation and terminology

Expected Response: y_{st} $s=1,...,N$ and for each s; $t=1,...,M$

Response: $Y_{stk}=y_{st}+W_{stk}$ $k=1,...,r_{st}$ index of order of measure

Assumptions:

$E\left(W_{stk}\right)=0$ for all $s=1,...,N$; $t=1,...,M$; $k=1,...,r_{st}$.

$\text{var}\left(W_{stk}\right)=\sigma_{st}^2$ for all $s=1,...,N$; $t=1,...,M$; $k=1,...,r_{st}$.

$\text{var}\left(W_{stk},W_{s^*t^*k^*}\right)=0$ for all $s=1,...,N$; $s^*=1,...,N$; $t=1,...,M$; $t^*=1,...,M$ and $st\neq s^*t^*$, and for all $k=1,...,r_{st}$, $k^*=1,...,r_{s^*t^*}$.

Parameters:

$\mu_s=\frac{1}{M}\sum_{t=1}^{M}y_{st}$, $\mu=\frac{1}{N}\sum_{t=1}^{M}\mu_s$, $\beta_s=\mu_s-\mu$, and $\varepsilon_{st}=y_{st}-\mu_s$ so that $y_{st}=\mu+\beta_s+\varepsilon_{st}$

$\sigma^2=\frac{1}{N-1}\sum_{s=1}^{N}\left(\mu_s-\mu\right)^2$, $\sigma_s^2=\frac{1}{M-1}\sum_{t=1}^{M}\left(y_{st}-\mu_s\right)^2$, $\sigma_r^2=\frac{1}{NM}\sum_{s=1}^{N}\sum_{t=1}^{M}\sigma_{st}^2$, $\sigma_e^2=\frac{1}{N}\sum_{s=1}^{N}\sigma_s^2$

and $\sigma^{*2}=\sigma^2-\frac{\sigma_e^2}{M}$.

Additional notation:

Define $\mathbf{U}_i=\left(\left(U_{is}\right)\right)=\left(U_{i1}\quad U_{i2}\quad\cdots\quad U_{iN}\right)'$, $\mathbf{U}_j^{(s)}=\left(\left(U_{jt}^{(s)}\right)\right)=\left(U_{j1}^{(s)}\quad U_{j2}^{(s)}\quad\cdots\quad U_{jM}^{(s)}\right)'$

$\mathbf{U}=\left(\left(\mathbf{U}_i\right)\right)'=\left(\mathbf{U}_1\quad\mathbf{U}_2\quad\cdots\quad\mathbf{U}_N\right)'$, and $\mathbf{U}^{(s)}=\left(\left(\mathbf{U}_j^{(s)}\right)\right)'=\left(\mathbf{U}_1^{(s)}\quad\mathbf{U}_2^{(s)}\quad\cdots\quad\mathbf{U}_M^{(s)}\right)'$

$\mathbf{y}=\left(\left(\mathbf{y}_s'\right)\right)'=\left(\mathbf{y}_1'\quad\mathbf{y}_2'\quad\cdots\quad\mathbf{y}_N'\right)'$, $\mathbf{y}_s=\left(\left(y_{st}\right)\right)=\left(y_{s1}\quad y_{s2}\quad\cdots\quad y_{sM_s}\right)'$

$\boldsymbol{\beta}=\left(\left(\beta_s\right)\right)=\left(\beta_1\quad\beta_2\quad\cdots\quad\beta_N\right)'$

$\mathbf{W}=\left(\left(\mathbf{W}_s'\right)\right)'=\left(\mathbf{W}_1'\quad\mathbf{W}_2'\quad\cdots\quad\mathbf{W}_N'\right)'$, $\mathbf{W}_s=\left(\left(W_{stk}\right)\right)=\left(W_{s1k}\quad W_{s2k}\quad\cdots\quad W_{sMk}\right)'$

2. SIMPLIFICATIONS

Assume $k=1,...,r_{st}=1$ for all $s=1,...,N$; $t=1,...,M$ so that $k=1$. (1 measure of response)

Note: If $k=1,...,r_{st}$ and $r_{st}\geq 1$, we could define $\bar{Y}_{st1}=\frac{1}{r_{st}}\sum_{k=1}^{r_{st}}Y_{stk}$ and $\text{var}\left(\bar{Y}_{st1}\right)=\frac{\sigma_{st}^2}{r_{st}}=\bar{\sigma}_{st}^2$ so that $\bar{Y}_{st1}=y_{st}+\bar{W}_{st1}$ where $\bar{W}_{st1}=\frac{1}{r_{st}}\sum_{k=1}^{r_{st}}W_{stk}$, $E\left(\bar{W}_{st1}\right)=0$ and $\text{var}\left(\bar{W}_{st1}\right)=\bar{\sigma}_{st}^2$. Thus, with suitable definitions, the model we discuss is more general.

3. PERMUTATION

a. Assume the clusters and subjects in clusters are permuted.
b. Assume response error represents measurement error on the subject with $r_{st}=1$ for all $s=1,...,N$ and $t=1,...,M_s$.
c. Assume $M_s=M$ for all $s=1,...,N$.

Define, $\tilde{Y}_{sj}=\sum_{t=1}^{M}U_{jt}^{(s)}y_{st}$, $\tilde{Y}_{sj}=\mathbf{U}_j^{(s)}\mathbf{y}_s$

$$\tilde{\mathbf{Y}}_s=\left(\left(\tilde{Y}_{sj}\right)\right)=\mathbf{U}^{(s)}\mathbf{y}_s$$

$$\left(\bigoplus_{s=1}^{N}\mathbf{U}^{(s)}\right)\mathbf{y}=\left(\left(\tilde{\mathbf{Y}}_s\right)\right)=\begin{pmatrix}\tilde{\mathbf{Y}}_1' & \tilde{\mathbf{Y}}_2' & \cdots & \tilde{\mathbf{Y}}_N'\end{pmatrix}'$$

$$\mathbf{Y}=\left(\mathbf{U}\otimes\mathbf{I}_M\right)\left(\bigoplus_{s=1}^{N}\mathbf{U}^{(s)}\right)\mathbf{y}$$

$$\mathbf{E}=\left(\mathbf{U}\otimes\mathbf{I}_M\right)\left(\bigoplus_{s=1}^{N}\mathbf{U}^{(s)}\right)\boldsymbol{\varepsilon}$$

$$\mathbf{W}^*=\left(\mathbf{U}\otimes\mathbf{I}_M\right)\left(\bigoplus_{s=1}^{N}\mathbf{U}^{(s)}\right)\mathbf{W}$$

$$\mathbf{Y}^*=\left(\left(Y_{ijk}^*\right)\right)=\mathbf{Y}+\mathbf{W}^*$$

$$\mathbf{Y}^*=vec\left(\left(\mathbf{y}+\mathbf{W}\right)'\left(\bigoplus_{s=1}^{N}\mathbf{U}^{(s)\prime}\right)\left(\mathbf{U}'\otimes\mathbf{I}_M\right)\right)=\left(\left(\mathbf{U}\otimes\mathbf{I}_M\right)\otimes\left(\mathbf{y}+\mathbf{W}\right)'\right)vec\left(\bigoplus_{s=1}^{N}\mathbf{U}^{(s)\prime}\right)$$

$$Y_{ij}=\sum_{s=1}^{N}\sum_{t=1}^{M}U_{is}U_{jt}^{(s)}y_{st}$$

$$W_{ijk}^*=\sum_{s=1}^{N}\sum_{t=1}^{M}U_{is}U_{jt}^{(s)}W_{stk}$$

$$\mathbf{Y}_i=\left(\left(Y_{ij}\right)\right)=\begin{pmatrix}Y_{i1} & Y_{i2} & \cdots & Y_{iM}\end{pmatrix}'$$

To summarize,

$$\mathbf{Y}^*=\mathbf{X}\mu+\mathbf{Z}\mathbf{B}+\left(\mathbf{E}+\mathbf{W}^*\right)$$

where $\mathbf{X}=\mathbf{1}_N\otimes\mathbf{1}_M$ and $\mathbf{Z}=\mathbf{I}_N\otimes\mathbf{1}_M$.

4. EXPECTED VALUE AND VARIANCE

ξ_1 Subscript representing expectation with respect to permutations of the clusters.
ξ_2 Subscript representing expectation with respect to permutations of units in a cluster.
ξ_3 Subscript representing expectation with respect to replication.

Expected Value:

$$E_{\xi_1\xi_2\xi_3}\left(\mathbf{Y}^*\right) = \mathbf{X}\mu$$

Variance:

$$\text{var}_{\xi_1\xi_2\xi_3}\left(\mathbf{Y}^*\right) = \left(\sigma_r^2 + \sigma_e^2\right)\mathbf{I}_{NM} + \sigma^{*2}\left(\mathbf{I}_N \otimes \mathbf{J}_M\right) - \frac{\sigma^2}{N}\mathbf{J}_{NM}$$

5. REARRANGING AND COLLAPSING THE RANDOM VARIABLES

The random variables in $\mathbf{Y}^*$ correspond those that are commonly used. We rearrange the random variables into those that are in the sample and remainder. It is also possible to collapse the random variables to sample and remainder PSU totals. We note this fact, but do not introduce additional notation to illustrate it here.

To arrange the random variables into a sample and a remainder, let $\mathbf{L} = \begin{pmatrix} \mathbf{L}_I \\ \mathbf{L}_{II} \end{pmatrix}$

$\begin{pmatrix} \mathbf{Y}_I^* \\ \hline \mathbf{Y}_{II}^* \end{pmatrix} = \mathbf{L}\mathbf{Y}^*$ where $\mathbf{L} = \left(\mathbf{L}_I' \mid \mathbf{L}_{II}'\right)'$, $\mathbf{L}_I = \left(\mathbf{I}_n \mid \underset{n\times(N-n)}{\mathbf{0}}\right) \otimes \left(\mathbf{I}_m \mid \underset{m\times(M-m)}{\mathbf{0}}\right)$ and

$$\mathbf{L}_{II} = \begin{pmatrix} \left(\mathbf{I}_n \mid \underset{n\times(N-n)}{\mathbf{0}}\right) \otimes \left(\underset{(M-m)\times m}{\mathbf{0}} \mid \mathbf{I}_{M-m}\right) \\ \left(\underset{(N-n)\times n}{\mathbf{0}} \mid \mathbf{I}_{N-n}\right) \otimes \mathbf{I}_M \end{pmatrix}.$$

In general, we represent the result as $\begin{pmatrix} \mathbf{Y}_I^* \\ \hline \mathbf{Y}_{II}^* \end{pmatrix} = \mathbf{L}\mathbf{Y}^*$. Using properties of $\mathbf{Y}^*$, we evaluate

$E_{\xi_1\xi_2\xi_3}\begin{pmatrix} \mathbf{Y}_I^* \\ \mathbf{Y}_{II}^* \end{pmatrix} = \begin{pmatrix} \mathbf{X}_I \\ \mathbf{X}_{II} \end{pmatrix}\boldsymbol{\alpha}$ and $\text{var}_{\xi_1\xi_2\xi_3}\begin{pmatrix} \mathbf{Y}_I^* \\ \hline \mathbf{Y}_{II}^* \end{pmatrix} = \text{var}_{\xi_1\xi_2}\begin{pmatrix} \mathbf{Y}_I \\ \hline \mathbf{Y}_{II} \end{pmatrix} + \sigma_r^2\mathbf{I}_{NM}$. This results in the following:

$$\begin{pmatrix}\mathbf{X}_I \\ \mathbf{X}_{II}\end{pmatrix} = \begin{pmatrix}\mathbf{1}_{nm} \\ \mathbf{1}_{NM-nm}\end{pmatrix}, \text{ and } \text{var}_{\xi_1\xi_2}\begin{pmatrix}\mathbf{Y}_I \\ \mathbf{Y}_{II}\end{pmatrix} = \begin{pmatrix}\mathbf{V}_I & \mathbf{V}_{I,II} \\ \mathbf{V}_{II,I} & \mathbf{V}_{II}\end{pmatrix},\ \mathbf{V}_I = \sigma_e^2\mathbf{I}_{nm} + \sigma^{*2}\left(\mathbf{I}_n \otimes \mathbf{J}_m\right) - \frac{\sigma^2}{N}\mathbf{J}_{nm},$$

$$\mathbf{V}_{I,II} = \mathbf{V}'_{II,I} = \sigma^{*2}\left(\mathbf{I}_n \otimes \underset{m\times(M-m)}{\mathbf{J}} \;\vdots\; \underset{nm\times(N-n)M}{\mathbf{0}}\right) - \frac{\sigma^2}{N}\underset{nm\times(NM-nm)}{\mathbf{J}}, \text{ and}$$

$$\mathbf{V}_{II} = \sigma_e^2\mathbf{I}_{NM-nm} + \sigma^{*2}\begin{pmatrix}\mathbf{I}_n \otimes \mathbf{J}_{M-m} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N-n} \otimes \mathbf{J}_M\end{pmatrix} - \frac{\sigma^2}{N}\mathbf{J}_{NM-nm}.$$

6. TARGET PARAMETERS, RANDOM VARIABLES, AND TERMINOLOGY

We assume that there is an interest in a ‘target’ that can be defined as a linear combination of the expected value (over response error) of the random variables. We use the sample data to estimate/predict the target.

We only consider inference for a single ‘target’ (not joint targets).

We represent the random variables by $\begin{pmatrix}\mathbf{Y}_I^* \\ \mathbf{Y}_{II}^*\end{pmatrix}$. We define a ‘target’ as $P = \mathbf{g}'\begin{pmatrix}\mathbf{Y}_I \\ \mathbf{Y}_{II}\end{pmatrix}$ where $E_{\xi_1\xi_2\xi_3}\begin{pmatrix}\mathbf{Y}_I^* \\ \mathbf{Y}_{II}^*\end{pmatrix} = \begin{pmatrix}\mathbf{X}_I \\ \mathbf{X}_{II}\end{pmatrix}\boldsymbol{\alpha}$ and partition $\mathbf{g}$ representing $\mathbf{g}' = \begin{pmatrix}\mathbf{g}'_I & \mathbf{g}'_{II}\end{pmatrix}$ so that $P = \mathbf{g}'_I\mathbf{Y}_I + \mathbf{g}'_{II}\mathbf{Y}_{II}$.

Since $\mathbf{Y}_I$ will be realized after selecting the sample and observing the results, the basic inference problem is prediction of $\mathbf{g}'_{II}\mathbf{Y}_{II}$. This is true when P is a fixed constant, and when P is a random variable. For this reason, we use the term ‘predictor’ when discussing inference.

For the random variables that are rearranged (see a) above), we define $\mathbf{g}' = \mathbf{b}' \otimes \frac{\mathbf{1}'_M}{M}$ where $\mathbf{b} = (b_1, \cdots, b_N)'$ are constants. In particular, we limit discussion to $\mathbf{b} = \mathbf{e}_i$, where $\mathbf{e}_i$ denotes an $N \times 1$ vector with a value of one in position i, and zero elsewhere. When random variables are collapsed, $\mathbf{g}' = \mathbf{e}'_i$.

7. DEVELOPING THE BEST LINEAR UNBIASED PREDICTOR (BLUP)

We define the BLUP of P as $\hat{P}$, where $\hat{P}$ satisfies the following criteria (see Royall, 1976):

Linear in the sample: $\hat{P} = \left(\mathbf{g}'_I + \mathbf{a}'\right)\mathbf{Y}_I^*$

Unbiased: $E_{\xi_1\xi_2\xi_3}\left(\hat{P} - P\right) = 0$

Minimum MSE: $\text{var}_{\xi_1\xi_2\xi_3}\left(\hat{P} - P\right)$ is minimized.

In order to develop the BLUP of P, we first present expressions for $\hat{P}-P$ and its variance. First, note that

$$\hat{P}-P=\left(\left(\mathbf{g}_I'+\mathbf{a}'\right) \mid -\mathbf{g}_I' \mid -\mathbf{g}_{II}'\right)\begin{pmatrix}\mathbf{Y}_I^* \\ \hline \mathbf{Y}_I \\ \hline \mathbf{Y}_{II}\end{pmatrix}.$$

We represent $\text{var}_\xi\begin{pmatrix}\mathbf{Y}_I \\ \mathbf{Y}_{II}\end{pmatrix}=\begin{pmatrix}\mathbf{V}_I & \mathbf{V}_{I,II} \\ \mathbf{V}_{II,I} & \mathbf{V}_{II}\end{pmatrix}$. Then

$$\text{var}_{\xi_1\xi_2\xi_3}\begin{pmatrix}\mathbf{Y}_I^* \\ \hline \mathbf{Y}_I \\ \hline \mathbf{Y}_{II}\end{pmatrix}=\begin{pmatrix}\mathbf{V}_I^* & \mathbf{V}_I & \mathbf{V}_{I,II} \\ \mathbf{V}_I & \mathbf{V}_I & \mathbf{V}_{I,II} \\ \mathbf{V}_{II,I} & \mathbf{V}_{II,I} & \mathbf{V}_{II}\end{pmatrix}.$$

Let us define $\mathbf{V}_I^*=\mathbf{V}_I+\sigma_r^2\mathbf{I}_{nm}$ and $\mathbf{V}_{RI}^*=\sigma_r^2\mathbf{I}_{nm}$. Then

$$\text{var}_{\xi_1\xi_2\xi_3}\begin{pmatrix}\mathbf{Y}_I^* \\ \hline \mathbf{Y}_I \\ \hline \mathbf{Y}_{II}\end{pmatrix}=\begin{pmatrix}1 & 0 \\ 1 & 0 \\ 0 & 1\end{pmatrix}\begin{pmatrix}\mathbf{V}_I & \mathbf{V}_{I,II} \\ \mathbf{V}_{II,I} & \mathbf{V}_{II}\end{pmatrix}\begin{pmatrix}1 & 1 & 0 \\ 0 & 0 & 1\end{pmatrix}+\begin{pmatrix}\mathbf{V}_{RI}^* & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0}\end{pmatrix}.$$

<u>The Unbiased Constraint</u>:

We can expand the unbiased constraint: $E_{\xi_1\xi_2\xi_3}\left(\hat{P}-P\right)=\left(\mathbf{a}' \quad -\mathbf{g}_{II}'\right)\begin{pmatrix}\mathbf{X}_I \\ \mathbf{X}_{II}\end{pmatrix}\boldsymbol{\alpha}$, such that $E_{\xi_1\xi_2\xi_3}\left(\hat{P}-P\right)=\left(\mathbf{a}'\mathbf{X}_I-\mathbf{g}_{II}'\mathbf{X}_{II}\right)\boldsymbol{\alpha}$. In order for this expression to equal zero for any value of $\boldsymbol{\alpha}$, the unbiased constraint will be always be satisfied when $\mathbf{a}'\mathbf{X}_I-\mathbf{g}_{II}'\mathbf{X}_{II}=\mathbf{0}$. This is introduced as a constraint using Lagrangian multipliers when minimizing the MSE.

<u>Finding the Minimum MSE</u>:

Notice that

$$\text{var}_{\xi R}\left(\hat{P}-P\right)=\left(\left(\mathbf{g}_I'+\mathbf{a}'\right) \mid -\mathbf{g}_I' \mid -\mathbf{g}_{II}'\right)\begin{pmatrix}\mathbf{V}_I^* & \mathbf{V}_I & \mathbf{V}_{I,II} \\ \mathbf{V}_I & \mathbf{V}_I & \mathbf{V}_{I,II} \\ \mathbf{V}_{II,I} & \mathbf{V}_{II,I} & \mathbf{V}_{II}\end{pmatrix}\begin{pmatrix}\mathbf{g}_I+\mathbf{a} \\ \hline -\mathbf{g}_I \\ \hline -\mathbf{g}_{II}\end{pmatrix}, \text{ or}$$

$$\text{var}_{\xi R}\left(\hat{P}-P\right)=\left(\mathbf{a}' \quad -\mathbf{g}_{II}'\right)\begin{pmatrix}\mathbf{V}_I & \mathbf{V}_{I,II} \\ \mathbf{V}_{II,I} & \mathbf{V}_{II}\end{pmatrix}\begin{pmatrix}\mathbf{a} \\ -\mathbf{g}_{II}\end{pmatrix}+\left(\mathbf{g}_I'+\mathbf{a}'\right)\mathbf{V}_{RI}^*\left(\mathbf{g}_I+\mathbf{a}\right).$$

Expanding this expression, $\text{var}\left(\hat{P}-P\right)=\mathbf{a}'\mathbf{V}_I^*\mathbf{a}+2\left(\mathbf{g}_I'\mathbf{V}_{RI}^*-\mathbf{g}_{II}'\mathbf{V}_{II,I}\right)\mathbf{a}+\mathbf{g}_{II}'\mathbf{V}_{II}\mathbf{g}_{II}+\mathbf{g}_I'\mathbf{V}_{RI}^*\mathbf{g}_I$.
Including the constraint via a Lagrangian multiplier, we seek to find the value of $\mathbf{a}$ that will minimize

$$f(\mathbf{a},\boldsymbol{\lambda}) = \mathbf{a}'\mathbf{V}_I^*\mathbf{a} + 2\left(\mathbf{g}_I'\mathbf{V}_{RI}^* - \mathbf{g}_{II}'\mathbf{V}_{II,I}\right)\mathbf{a} + 2\left(\mathbf{a}'\mathbf{X}_I - \mathbf{g}_{II}'\mathbf{X}_{II}\right)\boldsymbol{\lambda} + \mathbf{g}_{II}'\mathbf{V}_{II}\mathbf{g}_{II} + \mathbf{g}_I'\mathbf{V}_{RI}^*\mathbf{g}_I .$$

Differentiating with respect to $\mathbf{a}$ and $\boldsymbol{\lambda}$,

$$\frac{\partial f(\mathbf{a},\boldsymbol{\lambda})}{\partial \mathbf{a}} = 2\mathbf{V}_I^*\mathbf{a} + 2\left(\mathbf{V}_{RI}^*\mathbf{g}_I - \mathbf{V}_{I,II}\mathbf{g}_{II}\right) + 2\mathbf{X}_I\boldsymbol{\lambda} \text{ and}$$

$$\frac{\partial f(\mathbf{a},\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = 2\left(\mathbf{X}_I'\mathbf{a} - \mathbf{X}_{II}'\mathbf{g}_{II}\right).$$

To find the value of $\mathbf{a}$ that minimizes $f(\mathbf{a},\boldsymbol{\lambda})$, we set these derivatives to zero simultaneously, and solve for $\mathbf{a}$. The estimating equations are given by

$$\frac{1}{2}\begin{pmatrix} \dfrac{\partial f(\hat{\mathbf{a}},\hat{\boldsymbol{\lambda}})}{\partial \mathbf{a}} \\ \dfrac{\partial f(\hat{\mathbf{a}},\hat{\boldsymbol{\lambda}})}{\partial \boldsymbol{\lambda}} \end{pmatrix} = \begin{pmatrix} \mathbf{V}_I^* & \mathbf{X}_I \\ \mathbf{X}_I' & \mathbf{0} \end{pmatrix}\begin{pmatrix} \hat{\mathbf{a}} \\ \hat{\boldsymbol{\lambda}} \end{pmatrix} - \begin{pmatrix} -\mathbf{V}_{RI}^*\mathbf{g}_I + \mathbf{V}_{I,II}\mathbf{g}_{II} \\ \mathbf{X}_{II}'\mathbf{g}_{II} \end{pmatrix} = \begin{pmatrix} \mathbf{0}_n \\ \mathbf{0} \end{pmatrix}$$

or equivalently, by $\begin{pmatrix} \mathbf{V}_I^* & \mathbf{X}_I \\ \mathbf{X}_I' & \mathbf{0} \end{pmatrix}\begin{pmatrix} \hat{\mathbf{a}} \\ \hat{\boldsymbol{\lambda}} \end{pmatrix} = \begin{pmatrix} -\mathbf{V}_{RI}^*\mathbf{g}_I + \mathbf{V}_{I,II}\mathbf{g}_{II} \\ \mathbf{X}_{II}'\mathbf{g}_{II} \end{pmatrix}$.

To solve this equation for $\hat{\mathbf{a}}$, we make use of results on the inverse of a partitioned matrix $\mathbf{X} = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$, such that $\mathbf{X}^{-1} = \left(\begin{array}{c|c} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}\mathbf{Q}^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{Q}^{-1} \\ \hline -\mathbf{Q}^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{Q}^{-1} \end{array}\right)$ where $\mathbf{Q} = \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}$. As a result,

$$\begin{pmatrix} \mathbf{V}_I^* & \mathbf{X}_I \\ \mathbf{X}_I' & \mathbf{0} \end{pmatrix}^{-1} = \left(\begin{array}{c|c} \mathbf{V}_I^{*-1} - \mathbf{V}_I^{*-1}\mathbf{X}_I\left(\mathbf{X}_I'\mathbf{V}_I^{*-1}\mathbf{X}_I\right)^{-1}\mathbf{X}_I'\mathbf{V}_I^{*-1} & \mathbf{V}_I^{*-1}\mathbf{X}_I\left(\mathbf{X}_I'\mathbf{V}_I^{*-1}\mathbf{X}_I\right)^{-1} \\ \hline \left(\mathbf{X}_I'\mathbf{V}_I^{*-1}\mathbf{X}_I\right)^{-1}\mathbf{X}_I'\mathbf{V}_I^{*-1} & -\left(\mathbf{X}_I'\mathbf{V}_I^{*-1}\mathbf{X}_I\right)^{-1} \end{array}\right),$$

$$\hat{\mathbf{a}} = \left[\mathbf{V}_I^{*-1} - \mathbf{V}_I^{*-1}\mathbf{X}_I\left(\mathbf{X}_I'\mathbf{V}_I^{*-1}\mathbf{X}_I\right)^{-1}\mathbf{X}_I'\mathbf{V}_I^{*-1}\right]\left(-\mathbf{V}_{RI}^*\mathbf{g}_I + \mathbf{V}_{I,II}\mathbf{g}_{II}\right) + \mathbf{V}_I^{*-1}\mathbf{X}_I\left(\mathbf{X}_I'\mathbf{V}_I^{*-1}\mathbf{X}_I\right)^{-1}\mathbf{X}_{II}'\mathbf{g}_{II}.$$

We can now express the best linear unbiased predictor. Recall that $P = \mathbf{g}_I'\mathbf{Y}_I + \mathbf{g}_{II}'\mathbf{Y}_{II}$, and we predict P by $\hat{P} = \left(\mathbf{g}_I' + \mathbf{a}'\right)\mathbf{Y}_I^*$, where the best predictor replaces $\mathbf{a}$ by $\hat{\mathbf{a}}$. Let us define $\hat{\alpha} = \left(\mathbf{X}_I'\mathbf{V}_I^{*-1}\mathbf{X}_I\right)^{-1}\mathbf{X}_I'\mathbf{V}_I^{*-1}\mathbf{Y}_I^*$. Then

$$\begin{aligned}\hat{\mathbf{a}}'\mathbf{Y}_I^* &= \left(-\mathbf{g}_I'\mathbf{V}_{RI}^* + \mathbf{g}_{II}'\mathbf{V}_{II,I}\right)\left[\mathbf{V}_I^{*-1}\mathbf{Y}_I^* - \mathbf{V}_I^{*-1}\mathbf{X}_I\left(\mathbf{X}_I'\mathbf{V}_I^{*-1}\mathbf{X}_I\right)^{-1}\mathbf{X}_I'\mathbf{V}_I^{*-1}\mathbf{Y}_I^*\right] + \mathbf{g}_{II}'\mathbf{X}_{II}\left(\mathbf{X}_I'\mathbf{V}_I^{*-1}\mathbf{X}_I\right)^{-1}\mathbf{X}_I'\mathbf{V}_I^{*-1}\mathbf{Y}_I^* \\ &= -\mathbf{g}_I'\mathbf{V}_{RI}^*\mathbf{V}_I^{*-1}\left(\mathbf{Y}_I^* - \mathbf{X}_I\hat{\alpha}\right) + \mathbf{g}_{II}'\left[\mathbf{X}_{II}\hat{\alpha} + \mathbf{V}_{II,I}\mathbf{V}_I^{*-1}\left(\mathbf{Y}_I^* - \mathbf{X}_I\hat{\alpha}\right)\right]\end{aligned}$$

.

As a result, the best linear unbiased predictor is given by

$$\hat{P} = \mathbf{g}_I'\left[\mathbf{Y}_I^* - \mathbf{V}_{RI}^*\mathbf{V}_I^{*-1}\left(\mathbf{Y}_I^* - \mathbf{X}_I\hat{\alpha}\right)\right] + \mathbf{g}_{II}'\left[\mathbf{X}_{II}\hat{\alpha} + \mathbf{V}_{II,I}\mathbf{V}_I^{*-1}\left(\mathbf{Y}_I^* - \mathbf{X}_I\hat{\alpha}\right)\right].$$

We can express the predictor in a slightly different manner by substituting $\mathbf{V}_I^* = \mathbf{V}_I^* - \mathbf{V}_I$, resulting in

$$\hat{P} = \mathbf{g}_I{}' \left[\mathbf{X}_I \hat{\alpha} + \mathbf{V}_I \mathbf{V}_I^{*-1} \left(\mathbf{Y}_I^* - \mathbf{X}_I \hat{\alpha} \right) \right] + \mathbf{g}'_{II} \left[\mathbf{X}_{II} \hat{\alpha} + \mathbf{V}_{II,I} \mathbf{V}_I^{*-1} \left(\mathbf{Y}_I^* - \mathbf{X}_I \hat{\alpha} \right) \right].$$

<u>Example 1</u>.

Target: $\sum_{s=1}^{N} U_{is} \mu_s$ when $i \le n$

Then $\hat{\alpha} = \dfrac{\mathbf{1}'_{nm}}{nm} \mathbf{Y}_I^* = \bar{Y}^*$, $\mathbf{Y}_I^* - \mathbf{X}_I \hat{\alpha} = \mathbf{P}_{nm} \mathbf{Y}_I^*$, $\mathbf{V}_I \mathbf{V}_I^{*-1} \mathbf{P}_{nm} = \left[\rho_t \mathbf{I}_{nm} + (1-\rho_t) k^* \left(\mathbf{I}_n \otimes \dfrac{\mathbf{J}_m}{m} \right) \right] \mathbf{P}_{nm}$,

and $\mathbf{V}'_{I,II} \mathbf{V}_I^{*-1} \mathbf{P}_{nm} = k^* \begin{pmatrix} \mathbf{I}_n \otimes \dfrac{\mathbf{J}_{(M-m)\times m}}{m} \\ \hline \mathbf{0}_{(N-n)M \times nm} \end{pmatrix} \mathbf{P}_{nm}$ where $\rho_t = \dfrac{\sigma_e^2}{\sigma_e^2 + \sigma_r^2}$, $k^* = \dfrac{m\sigma^{*2}}{m\sigma^{*2} + \left(\sigma_e^2 + \sigma_r^2 \right)}$, and

$\mathbf{P}_a = \mathbf{I}_a - \dfrac{\mathbf{J}_a}{a}$. Using these expressions, the predictor simplifies when $i \le n$ to

$$\hat{T} = \left(\frac{m}{M} \right) \mathbf{e}'_{iI} \left(\mathbf{1}_n \bar{Y}_I^* + \left[\rho_t + (1-\rho_t) k^* \right] \mathbf{P}_n \bar{\mathbf{Y}}_I^* \right) + \left(\frac{M-m}{M} \right) \mathbf{e}'_{iI} \left(\mathbf{1}_n \bar{Y}_I^* + k^* \mathbf{P}_n \bar{\mathbf{Y}}_I^* \right).$$