Estimating the First Selected PSU Mean in a Two Stage Cluster Sample with Unequal Cluster Sizes based on an Expanded PSU Model

Ed Stanek

1. Introduction

Scott and Smith (1969) develop estimators for linear functions from a two stage cluster sample, with their discussion repeated in many places. This discussion was reviewed in c01ed13.doc. It is noteworthy that Scott and Smith allow expressions for the variance to depend on the cluster. A similar development is given by Vallient et al. We repeated the derivation of Vallient et al using the finite population variance in c01ed15.doc. These results were extended to settings with response error in c01ed25.doc. Both derivations are based on a ‘collapsed’ version of a vector of random variables, similar to the vector of random variables usually considered.

In the absence of response error, we developed results like those in c01ed15.doc in the context of an expanded PSU model with equal cluster sizes in c01ed27.doc. The results in c01ed27.doc have equal cluster sizes in the population, and produce the same estimates as the results in c01ed15.doc. However, new ideas are introduced in c01ed27.doc that we anticipate make it possible to expand the development to unequal size PSUs. The new ideas involve somewhat ‘arbitrary’ projections of terms in estimating equations to make the resulting equations of full rank and invertable. The projections are selected to make the estimating equations reduce to the estimating equations in c01ed15.doc. The purpose of this document is to extend the results in c01ed27.doc to unequal size PSUs.

The documents c01ed15.doc and c01ed27.doc assume that the number of SSUs in a PSU is equal for all PSUs. In an additional document, the expected value and variance is developed for a two stage problem with unequal size PSUs with response error (see c01ed26.doc). However, the document c01ed26.doc stops short of estimation. In order to clearly construct the estimating equations, the un-balance in PSU sizes has to be accounted for. It seems that this accounting will be more straightforward in an expanded population framework.

With this in mind, we develop an estimator for a cluster mean under two stage sampling with unequal size PSUs in an expanded population framework. We consider here the simpler setting where we have no response error. We expand the representation of the PSUs, but not of the SSUs.

The basic problem that we consider is estimation of a linear combination of population values based on the observed response for subjects and SSUs selected via two stage cluster sampling. The problem can be summarized with a model for the vector of random variables that represents a random permutation of the population. Detailed development and notation is given elsewhere, in particular in documents c00ed57.doc, c01ed15.doc, and c01ed27.doc.

2. A Superpopulation Framework for Inference of Linear Functions: Two Stage Cluster Sampling
2.1 Definition of the Population and the Superpopulation

2.1.1 The Population

The finite population is a labeled population of \( s = 1, \ldots, N \) primary sampling units (PSUs) where PSU \( s \) has \( M_s \) secondary sampling units (SSUs) \( (t = 1, \ldots, M_s) \). The non-stochastic values of the units in the population are represented by

\[ y' = (y'_1 \mid y'_2 \mid \cdots \mid y'_N) \]

where \( y'_s = (y_{s1} \ y_{s2} \ \cdots \ y_{sM_s}) \) for \( s = 1, \ldots, N \). These are population parameters.

2.1.2 The Superpopulation

We define the superpopulation to be a vector of random variables that arise from the two stage random permutations that are the basis of the two stage sampling. Normally, such random variables are represented by a linear combination of the underlying indicator random variables and unit values, such that

\[ Y_{ij} = \sum_{s=1}^{N} U_{is} \sum_{t=1}^{M_s} U_{jt}^{(s)} y_{st} \]

In this expression, \( U_{is} \) is an indicator random variable whose realized value is one if the \( i^{th} \) selected PSU is PSU \( s \), and zero otherwise. Similarly, \( U_{jt}^{(s)} \) is an indicator random variable whose realized value is one if the \( j^{th} \) selected SSU in PSU \( s \) is SSU \( t \), and zero otherwise.

We expand the expression for the random variables so as to represent each product of random variables in the linear combination uniquely. The basic random variables are \( U_{ist} y_{st} \). We arrange these random variables in a vector of dimension

\[ \left( \sum_{s=1}^{N} NM_s^2 \right) \times 1 \]

which we define as \( \mathbf{R}^* \) such that

\[ \mathbf{R}' = \left( \mathbf{R}_{1} \mathbf{R}_{12} \cdots \mathbf{R}_{1M_1} \mid \mathbf{R}_{21} \mathbf{R}_{22} \cdots \mathbf{R}_{2M_2} \mid \cdots \mid \mathbf{R}_{N1} \mathbf{R}_{N2} \cdots \mathbf{R}_{NM_N} \right) \]

where the individual vectors \( \mathbf{R}_{st} \) are of dimension \( NM_s \times 1 \) and are given by

\[ \mathbf{R}_{st} = \left( \mathbf{R}_{1st} \mathbf{R}_{2st} \cdots \mathbf{R}_{Nst} \right) \]

where \( \mathbf{R}_{ist} = y_{st} \left( U_{ist} U_{1t}^{(s)} U_{ist} U_{2t}^{(s)} \cdots U_{ist} U_{Mt}^{(s)} \right) \) such that

\[ \mathbf{R}_{st}' = y_{st} \left( U_{1st} U_{1t}^{(s)} U_{1st} U_{2t}^{(s)} \cdots U_{1st} U_{Mt}^{(s)} \right) \left( U_{Nst} U_{1t}^{(s)} U_{Nst} U_{2t}^{(s)} \cdots U_{Nst} U_{Mt}^{(s)} \right) \].

The superpopulation spans a larger dimensional space (ie. \( \sum_{s=1}^{N} (N-1)(M_s-1)^2 \)) than the population (which is a point in \( \sum_{s=1}^{N} M_s \) space).

2.1.3 Projection of the superpopulation SSUs onto a Single SSU random variable per selection.
We define a projection of SSU random variables onto a single SSU random variable per selection. The projection matrix is given by \( P = \bigoplus_{s=1}^{N} \left( I_{M_s} \otimes I_{NM_s} \right) \). After the projection, the basic random variables are of the form \( U_{\alpha} \sum_{i=1}^{M} U_{ji}^{(s)} y_{st} \). Hence \( R = P R^* = \left\{ \sum_{i=1}^{M} R_{1i}, \ldots, \sum_{i=1}^{M} R_{2i}, \ldots, \sum_{i=1}^{M} R_{Nt} \right\} \) of dimension \( \left( \sum_{i=1}^{N} NM_s \right) \times 1 \) where elements of \( \sum_{i=1}^{M} R_{st} \) are given by \( U_{\alpha} \sum_{i=1}^{M} U_{ji}^{(s)} y_{st} \).

Note that

\[
\begin{bmatrix}
\sum_{i=1}^{M} U_{11}^{(s)} y_{st} \\
\sum_{i=1}^{M} U_{12}^{(s)} y_{st} \\
\vdots \\
\sum_{i=1}^{M} U_{M_s}^{(s)} y_{st}
\end{bmatrix} =
\begin{bmatrix}
U_{11}^{(s)} & U_{12}^{(s)} & \cdots & U_{1M_s}^{(s)} \\
U_{21}^{(s)} & U_{22}^{(s)} & \cdots & U_{2M_s}^{(s)} \\
\vdots & \vdots & \ddots & \vdots \\
U_{M_s1}^{(s)} & U_{M_s1}^{(s)} & \cdots & U_{M_sM_s}^{(s)}
\end{bmatrix}
\begin{bmatrix}
y_{s1} \\
y_{s2} \\
\vdots \\
y_{sM_s}
\end{bmatrix} = U^{(s)} y_s \text{ so that the elements of } R^* \text{ can be represented as } R^* = \left( \left( \begin{bmatrix} U_{1s} \\ U_{2s} \\ \vdots \\ U_{Ns} \end{bmatrix} \otimes U^{(s)} y_s \right) \right).
\]

2.1.4 The relationship between the Projected Superpopulation and the Population

There is a relationship between the superpopulation and some of the parameters in the population. The projected random superpopulation vector can not be projected onto the full set of \( \sum_{s=1}^{N} M_s \) population parameters. It is possible to project \( R^* \) onto the \( N \times 1 \) vector of PSU means, \( \mu \), however. For example, note that \( \sum_{i=1}^{N} U_{is} \sum_{i=1}^{M} U_{ji}^{(s)} y_{st} = \sum_{i=1}^{M} U_{ji}^{(s)} y_{st} \) since \( \sum_{i=1}^{N} U_{is} = 1 \).

\[
\frac{\sum_{j=1}^{M} \sum_{i=1}^{M} U_{ji}^{(s)} y_{st}}{M_s} = \mu_s. \text{ Thus, if } C' = \bigoplus_{s=1}^{N} \frac{1_{NM_s}}{M_s}, \text{ then } C' R^* = \mu.\]
2.2 Population Parameters and Models for the Projected Superpopulation

The objective is estimation of parameters that can be expressed as linear functions of the values in the finite population. The functions correspond to individual PSU means. We first define these parameters in the population. Next, we define a model for the projected superpopulation and relate the projected superpopulation parameters to the population parameters. In addition, we specify any additional unbiased constraints and assumptions. Finally, we develop expressions for the expected value and variance of the projected superpopulation random variables.

2.2.1 Linear Functions of Population Parameters of Interest

The parameters of interest correspond to the PSU means given by

\[ \mu_s = \frac{\sum_{i=1}^{M_s} y_{si}}{M_s} \quad \text{for } s = 1, \ldots, N. \]

These parameters are linear functions of the population parameters. We represent these parameters in a vector, \( \mu = C'y \) where

\[ \mu' = (\mu_1, \mu_2, \ldots, \mu_N) \quad \text{and} \quad C' = \bigoplus_{s=1}^{N} \frac{1}{M_s}. \]

We can express the population parameters as the sum of the parameters of interest (the PSU means) plus a residual, such that \( y = X'\mu + e \), where \( X^* = \bigoplus_{s=1}^{N} \frac{1}{M_s} \) and \( e = y - X'\mu \).

2.2.2 A Model for the Projected Superpopulation

We described a model for the superpopulation in c00ed49.doc as \( R^* = X\beta + E^* \). With unequal numbers of SSUs per PSU, \( X \) is a \( \left( \sum_{s=1}^{N} NM_s^2 \right) \times N \) matrix of constants, \( \beta \) is a \( N \times 1 \) parameter vector. Projecting this model onto the sub-space results in the model

\[ P'R^* = P'X\beta + P'E^* \]

\[ R^* = X'\beta + E^*. \]

The projected superpopulation model contains a deterministic and random component. Without additional assumptions, there is no explicit connection between the parameters in the projected superpopulation \( \beta \), and the population parameters \( \mu \). However, in the next section, we define the superpopulation parameters \( \beta \) to be equal to \( \mu \).
2.2.3  The Relationship between Projected Superpopulation Parameters and Population Parameters

We relate the parameters in the model for $\mathbf{R}^*$ to parameters in the population. First, note that the projected superpopulation $\mathbf{R}^*$ is of dimension $\left(\sum_{s=1}^{N} NM_s\right) \times 1$, whereas the population $\mathbf{y}$ is of dimension $\left(\sum_{s=1}^{N} M_s\right) \times 1$. As indicated in section 2.1.4, even though the space spanned by $\mathbf{R}^*$ is of higher dimension than that of $\mathbf{y}$, the vector space spanned by $\mathbf{y}$ is not a sub-space of $\mathbf{R}^*$. The two vector spaces do intersect, however, and this intersection enables us to relate $\mu$ to $\beta$.

Using the projection $\mathbf{C}' = \bigoplus_{s=1}^{N} \frac{1_{NM_s}}{M_s}$, $\mathbf{C}' \mathbf{R}^* = \mu$. Also, $\mathbf{C}' \mathbf{R}^* = \mathbf{C}' \mathbf{X} \beta + \mathbf{C}' \mathbf{e}$. The non-stochastic portion of the projected superpopulation model when projected onto the population parameters of interest (given by the PSU means) is given by $\mathbf{C}' \mathbf{X} \beta$. In the population, $\mathbf{C}' \mathbf{y} = \mu$ where $\mathbf{C}' = \bigoplus_{s=1}^{N} \frac{1_{M_s}}{M_s}$. Also, the population model is given by

$$\mathbf{y} = \mathbf{X}' \mu + \mathbf{e} \quad \text{where} \quad \mathbf{X}' = \bigoplus_{s=1}^{N} 1_{M_s}.$$

Then $\mathbf{C}' \mathbf{y} = \mathbf{C}' \mathbf{X} \mu + \mathbf{C}' \mathbf{e}$ where

$$\mathbf{C}' \mathbf{X}' = \bigoplus_{s=1}^{N} \frac{1_{M_s}}{M_s} \left( \bigoplus_{s=1}^{N} 1_{M_s} \right) = \mathbf{I}_N.$$  
We define the superpopulation model parameters $\beta$ such that $\mathbf{C}' \mathbf{X} \beta = \mathbf{C}' \mathbf{X}^* \mu$ and without loss of generality, define $\beta = \mu$. This implies that $\mathbf{C}' \mathbf{X}^* = \mathbf{C}' \mathbf{X}^* = \mathbf{I}_N$. Even with this assumption, the design matrix $\mathbf{X}'$ is not uniquely defined.

One design matrix that will satisfy this constraint is given by $\mathbf{X}' = \bigoplus_{s=1}^{N} \frac{1_{NM_s}}{N}$. In the two stage superpopulation model that was not projected onto the SSU random variables, (see c00ed49.doc, p4), the design matrix $\mathbf{X}$ was also not unique, and was equivalent (with unequal size PSUs) to $\mathbf{X} = \bigoplus_{s=1}^{N} \frac{1_{NM_s^2}}{NM_s}$. The choice of $\mathbf{X}^* = \bigoplus_{s=1}^{N} \frac{1_{NM_s}}{N}$ in the superpopulation model projected on the SSU random variables is compatible with the choice of $\mathbf{X} = \bigoplus_{s=1}^{N} \frac{1_{NM_s^2}}{NM_s}$, since these two design matrices satisfy the requirement that $\mathbf{P}' \mathbf{X} = \mathbf{X}^*$ (where $\mathbf{P}' = \bigoplus_{s=1}^{N} \left( 1_{M_s} \otimes 1_{NM_s} \right)$). We use this design matrix in subsequent developments. Since $\beta = \mu$ and $\mathbf{C}' \mathbf{R}^* = \mu$, we can express the superpopulation parameters as a linear function of the projected superpopulation vector. Thus, $\beta = \mathbf{C}' \mathbf{R}^*$. 

C01ED29.DOC  7/5/01 3:23 PM  
5
2.3 Expected Value and Variance of the Projected Superpopulation

We construct expressions for the expected value and variance of the superpopulation using the subscript $\xi_1$ to indicate expectation with respect to selection of PSUs and the subscript $\xi_2$ to indicate expectation with respect to selection of SSUs. Expressions for the expected value and variance are developed for a superpopulation with unequal numbers of SSUs in c00ed9.doc, with similar developments for projected superpopulation vectors in c01ed26.doc. These expressions are simplified for the setting where there are equal numbers of SSUs in each PSU in c00ed26.doc. The expression for the variance is further simplified in c00ed38.doc. These results are summarized in c00ed49.doc. We use the results summarized in c00ed49.doc (p5-6) to obtain expressions for

\[
EE_{s,s}^{1212} = P^r\left[\varvar_{s,s}^{1212} (R^*)\right]P.
\]

2.3.1 Expected Value of the Projected Superpopulation Vector

Generalizing the results from c00ed49.doc (p5) to unequal size PSUs,

\[
E_{s,s}^{1212} (R^*) = \left(\oplus_{s=1}^{N} \left[ I_{M_s} \otimes \frac{1_{NM_s}}{NM_s} \right] \right) y. \quad \text{Since } \quad P^r = \oplus_{s=1}^{N} \left( I_{M_s} \otimes I_{NM_s} \right),
\]

\[
E_{s,s}^{1212} (R^*) = P^r E_{s,s}^{1212} (R^*) = \left(\oplus_{s=1}^{N} \left[ I_{M_s} \otimes \frac{1_{NM_s}}{NM_s} \right] \right) y = \left[ \oplus_{s=1}^{N} \left( I_{NM_s} / N \right) \right] \mu.
\]

2.3.1 Variance of the Projected Superpopulation Vector

The variance $R^*$ is given in c00ed49.doc (p6) for equal sized PSUs, and simplified in c00ed57.doc (pages 8-9) for $R^*$. For unequal size PSUs, the variance is derived in c00ed09.doc. It is given by

\[
\text{var}_{s,s}^{1212} (R^*) = P^r \left[ \text{var}_{s,s}^{1212} (R^*) \right] P.
\]
\[ \text{var}_{\text{R}}(\mathbf{R}) = \sum_{s=1}^{S} \left[ \frac{M_s}{M_s - 1} \left( \sum_{r=1}^{M_s} y_{st}^2 \right) - \frac{M_s}{M_s - 1} \frac{y_{s} y_{t}'}{M_s} \right] \left( \frac{I_{N}}{N} \otimes \mathbf{v}_{s} \right) \]

\[
\begin{bmatrix}
\mathbf{y}_1 y_1' \otimes \left( \mathbf{v}_{s_1} \otimes \frac{J_{M_1}}{M_1^2} \right) \\
\mathbf{y}_2 y_2' \otimes \left( \mathbf{v}_{s_1} \otimes \frac{J_{M_2}}{M_2^2} \right) \\
\vdots \\
\mathbf{y}_N y_N' \otimes \left( \mathbf{v}_{s_1} \otimes \frac{J_{M_N}}{M_N^2} \right)
\end{bmatrix}
\]

where \( \mathbf{v}_{s_1} = \frac{1}{N} \left( I_N - \frac{J_N}{N} \right) \) and \( \mathbf{v}_{s_2} = \frac{1}{M_s} \left( I_{M_s} - \frac{J_{M_s}}{M_s} \right) \).

The expression for the variance of \( \mathbf{R} \) is a matrix of dimension \( \left( \sum_{s=1}^{S} NM_s^2 \right) \times \left( \sum_{s=1}^{S} NM_s^2 \right) \). We express the variance of the projected vector of random variables, \( \mathbf{R}' = \mathbf{P}' \mathbf{R} \), as \( \text{var}_{\text{R}}(\mathbf{R}') = \mathbf{P}' \text{var}_{\text{R}}(\mathbf{R}) \mathbf{P} \) where \( \mathbf{P}' = \bigoplus_{s=1}^{N} \left( I_{M_s} \otimes I_{NM_s} \right) \). We simplifying this expression next. First, note that

\[
\mathbf{P}' = \bigoplus_{s=1}^{N} \left( I_{M_s} \otimes I_{NM_s} \right) \bigoplus_{r=1}^{M_s} \left( M_s - 1 \right) \left( \sum_{s=1}^{S} \frac{y_{st}^2}{M_s} - \frac{y_{s} y_{t}'}{M_s} \right) \otimes \left( \frac{I_{N}}{N} \otimes \mathbf{v}_{s_2} \right) \bigoplus_{s=1}^{N} \left( I_{M_s} \otimes I_{NM_s} \right)
\]

\[
= \bigoplus_{s=1}^{N} \left( I_{M_s} \otimes I_{NM_s} \right) \left( M_s - 1 \right) \left( \sum_{s=1}^{S} \frac{y_{st}^2}{M_s} - \frac{y_{s} y_{t}'}{M_s} \right) \otimes \left( \frac{I_{N}}{N} \otimes \mathbf{v}_{s_2} \right) \left( I_{M_s} \otimes I_{NM_s} \right)
\]

\[
= \bigoplus_{s=1}^{N} \left( \frac{M_s}{M_s - 1} I_{M_s} \otimes \frac{y_{st}^2}{M_s} - \frac{y_{s} y_{t}'}{M_s} \right) I_{M_s} \otimes \left( \frac{I_{N}}{N} \otimes \mathbf{v}_{s_2} \right)
\]
Now \( \sigma_s^2 = \frac{1}{M_s - 1} \left( M_s \left( \bigoplus_{r=1}^{M_s} y_{s, r} \right)^2 - \left( \bigoplus_{r=1}^{M_s} y_{s, r} \right) \right) \). As a result, the first term simplifies to

\[
P^N_{s=1} \left[ \frac{M_s}{M_s - 1} \left( \bigoplus_{r=1}^{M_s} y_{s, r} \right)^2 - \left( \bigoplus_{r=1}^{M_s} y_{s, r} \right) \right] \left( \frac{1}{N} \otimes v_{s_i} \right)
\]

\[
P = \frac{N}{N} \left[ \frac{M_s \sigma_s^2}{N} \otimes \left( \frac{1}{N} \otimes v_{s_i} \right) \right]
\]

or

\[
P^N_{s=1} \left[ \frac{M_s}{M_s - 1} \left( \bigoplus_{r=1}^{M_s} y_{s, r} \right)^2 - \left( \bigoplus_{r=1}^{M_s} y_{s, r} \right) \right] \left( \frac{1}{N} \otimes v_{s_i} \right)
\]

\[
P = \frac{N}{N} \left[ \sigma_s^2 \frac{1}{N} \otimes \left( \frac{1}{N} \otimes v_{s_i} \right) \right]
\]

We simplify the second term next. The second term is given by

\[
P' = \begin{pmatrix}
    y_1 y_1' \otimes \left( v_{s_i} \otimes \frac{J_{M_1}}{M_1^2} \right) \\
y_1 y_2' \otimes \left( -\frac{v_{s_i}}{N-1} \otimes \frac{M_1}{M_2} \right) \\
y_2 y_1' \otimes \left( -\frac{v_{s_i}}{N-1} \otimes \frac{M_1}{M_1} \right) \\
y_2 y_2' \otimes \left( v_{s_i} \otimes \frac{J_{M_2}}{M_2^2} \right) \\
y_2 y_2' \otimes \left( -\frac{v_{s_i}}{N-1} \otimes \frac{J_{M_2}}{M_2^2} \right) \\
y_2 y_2' \otimes \left( -\frac{v_{s_i}}{N-1} \otimes \frac{J_{M_2}}{M_2^2} \right) \\
y_3 y_1' \otimes \left( -\frac{v_{s_i}}{N-1} \otimes \frac{M_1}{M_1} \right) \\
y_3 y_2' \otimes \left( -\frac{v_{s_i}}{N-1} \otimes \frac{M_2}{M_2} \right) \\
y_N y_1' \otimes \left( v_{s_i} \otimes \frac{J_{M_N}}{M_N^2} \right) \\
y_N y_2' \otimes \left( -\frac{v_{s_i}}{N-1} \otimes \frac{M_N}{M_N} \right) \\
y_N y_2' \otimes \left( -\frac{v_{s_i}}{N-1} \otimes \frac{J_{M_N}}{M_N^2} \right) \\
y_N y_2' \otimes \left( -\frac{v_{s_i}}{N-1} \otimes \frac{J_{M_N}}{M_N^2} \right)
\end{pmatrix}
\]

which simplifies to
2.4 Sampling, Re-arranging, and Partitioning

We specify a re-arrangement of the projected superpopulation vector so that it can be partitioned into the sampled and remaining portions. We assume that simple random without replacement sampling is used to select \( i = 1, \ldots, n \) PSUs, and from each selected PSU, simple random without replacement sampling is used to select \( j = 1, \ldots, m_i \) SSUs. The sampled portion of the superpopulation we consider to be realized and non-stochastic. We focus attention on estimation of the remainder of the population. We specify a criteria for estimation, and then derive estimators (or estimating equations) that optimize this criteria.

2.4.1 Re-arranging and Partitioning the Projected Superpopulation

The two stage sampling corresponds to selection of a simple random without replacement sample of \( n \) PSU’s, and from each selected PSU selecting a simple random without replacement sample of \( m \) SSUs. We re-arrange terms in the projected superpopulation model into a sampled and remainder vector by pre-multiplying \( R^* \) by an \( (N^2M) \times (N^2M) \) permutation matrix \( K^* \). The matrix is defined as

\[
K^* = \begin{pmatrix}
I_N \otimes \left( I_n \begin{pmatrix} 0 \\ n \times (N-n) \end{pmatrix} \right) \otimes I_m \begin{pmatrix} 0 \\ m \times (M-m) \end{pmatrix} \\
I_N \otimes \left( I_n \begin{pmatrix} 0 \\ n \times (N-n) \end{pmatrix} \right) \otimes \left( I_M \begin{pmatrix} 0 \\ (M-m) \times m \end{pmatrix} \right)
\end{pmatrix}
\]

We define \( K^* R^* = \begin{pmatrix} K^* R^* \\\nK^* R^* \end{pmatrix} = \begin{pmatrix} R^*_r \\ R^*_r \end{pmatrix} \).
resulting in the model \( \begin{pmatrix} R_x \cr R_r \end{pmatrix} = \begin{pmatrix} X_x \cr X_r \end{pmatrix} \beta + \begin{pmatrix} E_x \cr E_r \end{pmatrix} \). The vector \( R_x \) is of dimension \((nM) \times 1\), with terms given by \( \sum_{i=1}^{M} R^x_{si} \) of dimension \( Nmn \times 1 \) where

\[
\begin{pmatrix}
U_{1s} \sum_{i=1}^{M} U_{1t}^{(s)} y_{st} \\
U_{1s} \sum_{i=1}^{M} U_{2t}^{(s)} y_{st} \\
\vdots \\
U_{ns} \sum_{i=1}^{M} U_{1t}^{(s)} y_{st} \\
U_{ns} \sum_{i=1}^{M} U_{2t}^{(s)} y_{st} \\
\vdots \\
U_{ns} \sum_{i=1}^{M} U_{mt}^{(s)} y_{st}
\end{pmatrix}
\]

where \( \sum_{i=1}^{M} R^x_{st} \) is of dimension \( mn \times 1 \). The vector \( R^x \) is of dimension \((N^2M - nNm) \times 1\).

Since \( X^x = I_N \otimes \frac{1}{N} I_{nm} \), \( K^x X^x = \begin{pmatrix} K^x X_x \\
K^x X_r \end{pmatrix} = \begin{pmatrix} X_x \\
X_r \end{pmatrix} = \begin{pmatrix}
I_N \otimes \frac{1}{N} I_{nm} \\
I_N \otimes \frac{1}{N} I_{N(M-n)} \\
I_N \otimes \frac{1}{N} I_{N(n-M)}
\end{pmatrix} \).

Note that since \( K^x \) is a permutation matrix, \( K^x K^x = I \). We partition the variance of the projected superpopulation in a similar manner. After re-arranging random variables into the
sample and remaining portions, \[ \text{var}_{\bar{s}\bar{r}2} \left[ \begin{bmatrix} K^1 R^1 \\ \vdots \\ K^d R^d \end{bmatrix} \right] = \text{var}_{\bar{s}\bar{r}2} \left[ \begin{bmatrix} R^\circ_{\bar{s}} \\ \vdots \\ R^\circ_{\bar{r}} \end{bmatrix} \right], \]
which we represent as
\[ \text{var}_{\bar{s}\bar{r}2} \begin{bmatrix} \frac{R^\circ_{\bar{s}}}{R^\circ_{\bar{r}}} \end{bmatrix} = \begin{bmatrix} V^\circ_{\bar{s}} \\ V^\circ_{\bar{r}} \end{bmatrix}. \]

The expression for \[ V^\circ_{\bar{s}} \] is given by
\[ V^\circ_{\bar{s}} = K^\circ_{\bar{s}} \left[ \text{var}_{\bar{s}\bar{r}2} \left( \begin{bmatrix} R^\circ \end{bmatrix} \right) \right] K^\circ_{\bar{s}}'. \]

Since \[ K^\circ_{\bar{s}} = I_N \otimes \begin{bmatrix} I_n \\ 0 \end{bmatrix} \odot \left( \begin{bmatrix} I_m \\ 0 \end{bmatrix} \right) \] and
\[ \text{var}_{\bar{s}\bar{r}2} \left( \begin{bmatrix} R^\circ \end{bmatrix} \right) = (M-1) \left( \bigoplus_{s=1}^{N} \sigma^2 \right) \otimes \begin{bmatrix} I_n \\ 0 \end{bmatrix} \otimes \begin{bmatrix} I_m \\ 0 \end{bmatrix} \] we get
\[ \text{var}_{\bar{s}\bar{r}2} \left( \begin{bmatrix} R^\circ \end{bmatrix} \right) = \begin{bmatrix} D_{\mu^2} - \frac{\mu_0 \mu'}{N} \end{bmatrix} \otimes \left( \begin{bmatrix} v^\circ_{\bar{s}_1} \otimes \begin{bmatrix} J \end{bmatrix} \right) \right), \]
\[ V^\circ_{\bar{s}} = K^\circ_{\bar{s}} \left( \begin{bmatrix} (M-1) \left( \bigoplus_{s=1}^{N} \sigma^2 \right) \otimes \begin{bmatrix} I_n \\ 0 \end{bmatrix} \otimes \begin{bmatrix} I_m \\ 0 \end{bmatrix} \right) + \begin{bmatrix} D_{\mu^2} - \frac{\mu_0 \mu'}{N} \end{bmatrix} \otimes \left( \begin{bmatrix} v^\circ_{\bar{s}_1} \otimes \begin{bmatrix} J \end{bmatrix} \right) \right) \right] K^\circ_{\bar{s}}'. \]

This simplifies further to
\[ V^\circ_{\bar{s}} = \frac{1}{N-1} \begin{bmatrix} I_n - \frac{J}{N} \end{bmatrix}, \]
and
\[ v^\circ_{\bar{s}_1} = \frac{1}{N-1} \begin{bmatrix} I_n - \frac{J}{N} \end{bmatrix}. \]

Also,
\[ V^\circ_{sr} = K^\circ_{sr} \left( \begin{bmatrix} (M-1) \left( \bigoplus_{s=1}^{N} \sigma^2 \right) \otimes \begin{bmatrix} I_n \\ 0 \end{bmatrix} \otimes \begin{bmatrix} I_m \\ 0 \end{bmatrix} \right) + \begin{bmatrix} D_{\mu^2} - \frac{\mu_0 \mu'}{N} \end{bmatrix} \otimes \left( \begin{bmatrix} v^\circ_{\bar{s}_1} \otimes \begin{bmatrix} J \end{bmatrix} \right) \right) \right] K^\circ_{sr}'. \]

The expressions for \[ K^\circ_{sr} \] and \[ K^\circ_{sr}' \] are:
\[ K^\circ_{sr} = I_N \otimes \begin{bmatrix} I_n \\ 0 \end{bmatrix} \odot \left( \begin{bmatrix} I_m \\ 0 \end{bmatrix} \right), \]
\[ K^\circ_{sr}' = I_N \otimes \begin{bmatrix} I_n \\ 0 \end{bmatrix} \odot \left( \begin{bmatrix} I_m \\ 0 \end{bmatrix} \right). \]

As a result,
\[ K^\circ_{sr} \left( \begin{bmatrix} (M-1) \left( \bigoplus_{s=1}^{N} \sigma^2 \right) \otimes \begin{bmatrix} I_n \\ 0 \end{bmatrix} \otimes \begin{bmatrix} I_m \\ 0 \end{bmatrix} \right) + \begin{bmatrix} D_{\mu^2} - \frac{\mu_0 \mu'}{N} \end{bmatrix} \otimes \left( \begin{bmatrix} v^\circ_{\bar{s}_1} \otimes \begin{bmatrix} J \end{bmatrix} \right) \right) \right] K^\circ_{sr}'. \]

The expressions for \[ K^\circ_{sr} \] and \[ K^\circ_{sr}' \] are:
\[ K^\circ_{sr} = I_N \otimes \begin{bmatrix} I_n \\ 0 \end{bmatrix} \odot \left( \begin{bmatrix} I_m \\ 0 \end{bmatrix} \right), \]
\[ K^\circ_{sr}' = I_N \otimes \begin{bmatrix} I_n \\ 0 \end{bmatrix} \odot \left( \begin{bmatrix} I_m \\ 0 \end{bmatrix} \right). \]
\[ K_s = \left[ D_{\mu^2} - \frac{\mu\mu}{N} \right] \otimes (v_s \otimes J_M) = \]
\[ \left( I_N \otimes \left( I_n \mid 0_{n\times(N-n)} \right) \otimes \left( I_m \mid 0_{m\times(M-m)} \right) \right) \left[ D_{\mu^2} - \frac{\mu\mu}{N} \right] \otimes \frac{1}{N-1} \left( I_N - \frac{J_N}{N} \right) \otimes J_M = . \]

Then

\[ K_s = \left[ (M-1) \left( \sum_{s=1}^{N} \sigma_s^2 \right) \otimes \frac{I_N}{N} \otimes \frac{1}{M-1} \left( I_M - \frac{J_M}{M} \right) \right] K_s' = \]
\[ \left( \sum_{s=1}^{N} \sigma_s^2 \right) \otimes \frac{I_n}{N} \otimes \frac{J}{M} = . \]

Also,

\[ K_s = \left[ D_{\mu^2} - \frac{\mu\mu}{N} \right] \otimes (v_s \otimes J_M) = \]
\[ \left( D_{\mu^2} - \frac{\mu\mu}{N} \right) \otimes \frac{1}{N-1} \left( I_n \mid 0_{n\times(N-n)} \right) \otimes J_{m\times M} = \]
\[ \left( D_{\mu^2} - \frac{\mu\mu}{N} \right) \otimes \frac{1}{N-1} \left( I_n \mid 0_{n\times(N-n)} \right) \otimes J_{m\times M} = . \]

Thus,
Expressions for the variance have been evaluated as ‘unconditional’ expressions for the variance. Recall that these expressions are formed by evaluating

\[ \text{var}(\mathbf{R}) = \mathbf{E}_s \left[ \text{var}_{\mathbf{E}_s} \left( \mathbf{R}^e \right) \right] + \text{var}_{\mathbf{E}_s} \left[ \mathbf{E}_s \left( \mathbf{R}^e \right) \right] \mathbf{K}^e \]

or

\[ \text{var}_{\mathbf{E}_s} \left( \mathbf{R}^e \right) = \mathbf{K}^e \mathbf{E}_s \left[ \text{var}_{\mathbf{E}_s} \left( \mathbf{R}^e \right) \right] \mathbf{K}^e + \mathbf{K}^e \text{var}_{\mathbf{E}_s} \left[ \mathbf{E}_s \left( \mathbf{R}^e \right) \right] \mathbf{K}^e \]

We evaluate expressions for the variance conditional on \( \mathbf{U}_s = \mathbf{u}_s \) in c01ed28.doc, and do not pursue the partially conditional expansion here.


We described a model for the superpopulation in section 2.2.2 given by

\[ \mathbf{R}^e = \mathbf{X} \mathbf{\beta} + \mathbf{E}^e \]

The design matrix is given by \( \mathbf{X}^e = \mathbf{I}_N \otimes \frac{1}{NM} \). Since \( \mathbf{K}^e \mathbf{R}^e = \left( \mathbf{K}^e \mathbf{R}^e \right) \), we have

\[ \mathbf{K}^e \mathbf{X}^e = \left( \mathbf{K}^e \mathbf{X}^e \right) = \left( \mathbf{X}^e \right) \]

The projected superpopulation model contains a deterministic and random component. We define the \( N \times 1 \) parameter vector \( \mathbf{\beta} \) to be equal to the PSU means, \( \mathbf{\mu} \).

Consider a linear function of the random variables given by \( \mathbf{T} = \mathbf{g}^e \mathbf{R}^e = \mathbf{g}^e \mathbf{R}^e + \mathbf{g}^e \mathbf{R}^e \), and define a predictor of this random variable that is a linear function of the potentially observed sample data given by \( \mathbf{\hat{T}} = \mathbf{L}^e \mathbf{R}^e \). We focus here on a linear function that is equal to
the mean response for a selected PSU. Once the random variable representing the particular selection of a PSU is realized, the mean response for the realized PSU is a parameter, not a random variable. The term estimator is used for parameters, as opposed to the term predictors, for random variables.

As noted in section 2.2.3, certain linear combinations of the random variables are not random. However, we focus on linear combinations of $R^s$ that are stochastic, such as the PSU mean corresponding to the first selected PSU. For such an example,

$$g^r = 1_N^\prime \otimes \left( \begin{array}{c} 1 \\ 0 \\ \vdots \\ 0 \\ \end{array} \right) \frac{1}{M}, \text{ such that } g^r = 1_N^\prime \otimes \left( \begin{array}{c} 1 \\ 0 \\ \vdots \\ 0 \\ \end{array} \right) \frac{1}{M} \text{ and }$$

$$g_r^r = \left( \begin{array}{c} 1 \\ 0 \\ \vdots \\ 0 \\ \end{array} \right) \otimes \frac{1}{M-m} \begin{array}{c} 0 \\ \vdots \\ \vdots \\ 0 \\ \end{array}_{N(N-n)M}.$$

The estimation error is given by $\hat{T} - T = \left( L' - g_s^r \right) R_s - g_r^r R_r^r$. Let us define $a' = \left( L' - g_s^r \right)$ so that $\hat{T} - T = \left( a' - g_r^r \right) \left( R_s^e R_r^e \right)$. Then our goal is to estimate $a'$. We choose an estimator that is unconditionally unbiased, such that $E_{\hat{a}g} \left( \hat{T} - T \right) = 0$. The unbiased constraint simplifies to $a'X_r^r - g_r^r X_r^r = 0$. We also require the estimator to have minimum unconditional variance, such that $\text{var}_{\hat{a}g} \left( \hat{T} - T \right)$ is a minimum. The variance can be expanded such that $\text{var}_{\hat{a}g} \left( \hat{T} - T \right) = a'V_s^o a - 2a'V_s^o g_r^r + g_r^r V_r^o g_r^r$. As a result, the Lagrange function to be minimized is given by $\Phi = a'V_s^o a - 2a'V_s^o g_r^r + g_r^r V_r^o g_r^r + 2 \left( a'X_r^r - g_r^r X_r^r \right) \lambda$, where $\lambda$ has dimension $N \times 1$.

We minimize this function next. Taking the derivative with respect to $a$, and setting the derivative to zero results in the equation, $V_s a - V_s g_r^r + g_r^r V_r g_r^r = 0$; taking the derivative with respect to $\lambda$ and setting the derivative to zero results in the equation given by $\left( a'X_r^r - g_r^r X_r^r \right) = 0$. These equations can be expressed simultaneously as

$$\left( \begin{array}{c} V_s^o & X_s^r \\ 0_{N \times N} & \end{array} \right) \left( \begin{array}{c} \hat{a} \\ \hat{\lambda} \end{array} \right) = \left( \begin{array}{c} V_s^r \\ X_r^r \end{array} \right) g_r^r.$$

These estimating equations are based on a partially expanded representation of the superpopulation. Prior to solving these equations, we review the dimension of the terms in
these equations. First, the matrix $V_s^r$ has dimension $Nnm \times Nnm$, while the matrix $V_{sr}^r$ has dimension $Nnm \times \left[ Nn (M - m) + N (N - n) M \right]$. Also, $X_i^r$ has dimension $Nnm \times N$, $X_i'$ has dimension $N \times \left[ Nn (M - m) + N (N - n) M \right]$, $\hat{a}$ has dimension $Nnm \times 1$, $\hat{\lambda}$ has dimension $N \times 1$ and $g_r^r$ has dimension $\left[ Nn (M - m) + N (N - n) M \right] \times 1$.

We reduce the dimension of these equations prior to solving them. The dimension reduction corresponds to the same reduction that occurs when we project the partially expanded superpopulation onto the usual set of $NM$ random variables used to represent the population. We first outline the general strategy. The estimating equations are of the form $AB = C$. We would like to make the square matrix $A$ non-singular by reducing the dimension. To reduce the dimension, we pre and post multiply $A$ by a matrix $M$, resulting in the equations: $(M'AM)(M'B) = (M'C)$. We choose the matrix $M$ so that the equations that result after the multiplication match the equations based on a collapsed version of the PSU random variables. We can think of this process as involving two steps. The first step is a reduction in the dimension of the equations. This results in equations given by $M'AB = M'C$. The second step in the process introduces a matrix between $A$ and $B$. Although I have represented this matrix as $MM'$, it is possible that other matrices could be used to simplify the equations.

We develop one dimension reduction matrix here. Of course, since the matrix $A$ is singular, there are infinite possible solutions. Our choice of the dimension reduction matrix is arbitrary in that sense. The question arises as to why we should choose one reduction matrix over another. We don’t have a clear answer to this question. The dimension reduction matrix that we choose is motivated by the manner in which the superpopulation can be collapsed to form a set of unique estimating equations. This collapsing is no more arbitrary than the dimension reduction that we introduce here.

For the partially expanded superpopulation, we can reduce the superpopulation vector to the usual vector of random variables by the projection given by $\left( I_N' \otimes I_N \otimes I_M \right) R^s = Y_{NM \times 1}$.

In order to determine an expression for $M$, we need to know how to reduce $V_s^r$ to $V_r$. This involves various projections and re-arrangements which we summarize here. First, since $V_s^r = \text{var}_{s \in S} \left( R_s^r \right)$, we first obtain an expression that involves $R_s^r$. Since $R_s^r = K' K R_s^r = K' \left( \begin{bmatrix} R_s' \\ R_r' \end{bmatrix} \right)$, we have $\left( I_N' \otimes I_N \otimes I_M \right) R^s = Y_{NM \times 1}$.

In a similar manner, note that $V_s = \text{var}_{s \in S} (Y_s)$, and $Y = K' KY = K' \left( \begin{bmatrix} Y_s' \\ Y_r' \end{bmatrix} \right)$ and hence
\[
\begin{align*}
\left( \mathbf{I}_N \otimes \mathbf{I}_N \otimes \mathbf{I}_M \right) \mathbf{K}^{'} & \left( \begin{array}{c}
\mathbf{R}_s^r \\
\mathbf{R}_r^r 
\end{array} \right) = \mathbf{Y} \\
\left( \mathbf{I}_N \otimes \mathbf{I}_N \otimes \mathbf{I}_M \right) \mathbf{K}^{'} & \left( \begin{array}{c}
\mathbf{R}_s^r \\
\mathbf{R}_r^r 
\end{array} \right) = \mathbf{K} \left( \begin{array}{c}
\mathbf{Y}_s^r \\
\mathbf{Y}_r^r 
\end{array} \right),
\end{align*}
\]

Since \( \mathbf{K} = \left( \frac{\mathbf{K}_s}{\mathbf{K}_r} \right) \), we can express \( \left( \begin{array}{c}
\mathbf{K}_s^r \\
\mathbf{K}_r^r 
\end{array} \right) \left( \mathbf{I}_N \otimes \mathbf{I}_N \otimes \mathbf{I}_M \right) \mathbf{K}^{'} \left( \begin{array}{c}
\mathbf{R}_s^r \\
\mathbf{R}_r^r 
\end{array} \right) = \left( \begin{array}{c}
\mathbf{Y}_s^r \\
\mathbf{Y}_r^r 
\end{array} \right) \) which implies that
\[
\left[ \mathbf{K}_s \left( \mathbf{I}_N \otimes \mathbf{I}_N \otimes \mathbf{I}_M \right) \mathbf{K}^{'} \right] \left( \begin{array}{c}
\mathbf{R}_s^r \\
\mathbf{R}_r^r 
\end{array} \right) = \mathbf{Y}_s^r.
\]

TO HERE 6/26/01

We develop expressions in mode detail next. In terms of the re-arranged superpopulation, since \( \mathbf{K}^* \mathbf{R}^* = \left( \frac{\mathbf{K}_s^r \mathbf{R}_s^r}{\mathbf{K}_r^r \mathbf{R}_r^r} \right) = \left( \begin{array}{c}
\mathbf{R}_s^r \\
\mathbf{R}_r^r 
\end{array} \right) \), this projection is given by
\[
\left( \mathbf{I}_N \otimes \mathbf{I}_N \otimes \mathbf{I}_M \right) \mathbf{K}^{'} \left( \begin{array}{c}
\mathbf{R}_s^r \\
\mathbf{R}_r^r 
\end{array} \right).
\]

Let us define
\[
\left( \begin{array}{c}
\mathbf{P}_s^r \\
\mathbf{P}_r^r 
\end{array} \right) = \left( \begin{array}{c}
\mathbf{I}_N \otimes \frac{\mathbf{I}_n}{(N-n)n} \otimes \frac{\mathbf{I}_m}{(M-m)m} \\
\mathbf{I}_N \otimes \frac{\mathbf{I}_m}{(M-m)m} \otimes \frac{\mathbf{I}_n}{(N-n)n} 
\end{array} \right)
\]

such that \( \mathbf{P}_s^r = \mathbf{I}_N \otimes \frac{\mathbf{I}_n}{(N-n)n} \otimes \frac{\mathbf{I}_m}{(M-m)m} \) and
\[
\mathbf{P} = \left( \begin{array}{c}
\mathbf{I}_N \otimes \frac{\mathbf{I}_n}{(N-n)n} \otimes \frac{\mathbf{I}_{M-m}}{M-m} \\
\mathbf{I}_N \otimes \frac{\mathbf{I}_{M-m}}{M-m} \otimes \frac{\mathbf{I}_n}{(N-n)n} 
\end{array} \right).\]

Note that since \( \mathbf{K}^* \mathbf{K}^* = \mathbf{I}_{N^2M} \) and \( \left( \mathbf{I}_N \otimes \mathbf{I}_N \otimes \mathbf{I}_M \right) \mathbf{R}^* = \mathbf{Y}_{NMx1} \), then
\[
\left( \begin{array}{c}
\mathbf{P}_s^r \\
\mathbf{P}_r^r 
\end{array} \right) \left( \begin{array}{c}
\mathbf{R}_s^r \\
\mathbf{R}_r^r 
\end{array} \right) = \mathbf{Y}.
\]
\( Y \) can be re-arranged into a sampled and remaining portion by premultiplication by the matrix \( K = \begin{pmatrix} K_s \\ K_r \end{pmatrix} \) to form \( KY = \begin{pmatrix} K_s Y_s \\ K_r Y_r \end{pmatrix} \) where \( K_s = \begin{pmatrix} I_n & 0 \\ n_{(N-n)} & 0_{(M-m)\text{can}} \\ \end{pmatrix} \otimes \begin{pmatrix} I_m & 0 \\ m_{(N-n)} & 0_{(M-m)\text{can}} \end{pmatrix} \),

and \( K_r = \begin{pmatrix} I_n & 0_{(N-n)\text{can}} \\ 0_{(N-n)\text{can}} & I_{N-n} \otimes I_M \end{pmatrix} \). We make use of the permutation matrix here.

As a result, \( \begin{pmatrix} K_s \\ K_r \end{pmatrix} \begin{pmatrix} P_s' | P_r' \end{pmatrix} \begin{pmatrix} R_s' \\ R_r' \end{pmatrix} = \begin{pmatrix} Y_s \\ Y_r \end{pmatrix} \).

We simplify the product \( P' = \begin{pmatrix} K_s \\ K_r \end{pmatrix} \begin{pmatrix} P_s' | P_r' \end{pmatrix} = \begin{pmatrix} K_s P_s' | K_r P_r' \end{pmatrix} \). Terms in this expression are given by

\[
K_s P_s' = \begin{pmatrix} I_n & 0_{(N-n)\text{can}} \\ n_{(N-n)} & 0_{(M-m)\text{can}} \end{pmatrix} \begin{pmatrix} I_{N'} \otimes \begin{pmatrix} I_n & 0_{(N-n)\text{can}} \\ 0_{(N-n)\text{can}} & I_{M-m} \end{pmatrix} \otimes I_{m_{(N-n)}} \end{pmatrix} \\
= I_{N'} \otimes I_{m_{(N-n)}}
\]

\[
K_s P_r' = \begin{pmatrix} I_n & 0_{(N-n)\text{can}} \\ n_{(N-n)} & 0_{(M-m)\text{can}} \end{pmatrix} \begin{pmatrix} I_{N'} \otimes \begin{pmatrix} I_n & 0_{(N-n)\text{can}} \\ 0_{(N-n)\text{can}} & I_{M-m} \end{pmatrix} \otimes I_{m_{(N-n)}} \end{pmatrix} \\
= \begin{pmatrix} 0_{n_{(N-n)} (M-m)\text{can}} & 0_{n_{(N-n)} (M-m)\text{can}} \\ n_{(N-n)} (M-m)\text{can} & 0_{(N-n)\text{can} (M-m)\text{can}} \end{pmatrix}
\]

\[
K_s P_s' = \begin{pmatrix} I_n & 0_{(N-n)\text{can}} \\ n_{(N-n)} & 0_{(M-m)\text{can}} \end{pmatrix} \begin{pmatrix} I_{N'} \otimes \begin{pmatrix} I_n & 0_{(N-n)\text{can}} \\ 0_{(N-n)\text{can}} & I_{M-m} \end{pmatrix} \otimes I_{m_{(N-n)}} \end{pmatrix} \\
= \begin{pmatrix} 0_{n_{(N-n)} (M-m)\text{can}} & 0_{n_{(N-n)} (M-m)\text{can}} \\ n_{(N-n)} (M-m)\text{can} & 0_{(N-n)\text{can} (M-m)\text{can}} \end{pmatrix}
\]

and finally,
\[ K_j P_r' = \left[ \begin{array}{c|c|c} I \times (M-m) \\ \hline 0 \times (M-m) & 0 \times M \end{array} \right] \otimes \left[ \begin{array}{c|c|c} I \times (M-m) \\ \hline 0 \times (M-m) & 0 \times M \end{array} \right] \]

\[ = \left[ \begin{array}{c|c|c} I_n \otimes I_{n(M-m)} \\ \hline 0 \times n(M-n) & 0 \times n(N-n) \end{array} \right] \otimes \left[ \begin{array}{c|c|c} I_n \otimes I_{n(M-m)} \\ \hline 0 \times n(M-n) & 0 \times n(N-n) \end{array} \right] \]

. To summarize,

\[ \left( \frac{K_j}{K_r} \right) \begin{pmatrix} P_x' & P_y' \\ R_x' & R_y' \end{pmatrix} = \begin{pmatrix} Y_x' \\ Y_y' \end{pmatrix}, \]

where \( P' = \begin{pmatrix} P_x' & P_y' \\ P_x' & P_y' \end{pmatrix} \). Now we are ready to reduce the estimating equations to a dimension that is usually used for the problem. We do so by introducing the matrix \( \begin{pmatrix} P_x' & 0 \\ 0 & 1_n' \end{pmatrix} \), where we note that

\[ \begin{pmatrix} P_x' & 0 \\ 0 & 1_n' \end{pmatrix} = J_N \otimes I_{m+1}, \] and \( P_x' P_r' = J_N \otimes I_{M-n} \). Using these matrices, we project the terms in the estimating equations such that

\[ \begin{pmatrix} P_x' & 0 \\ 0 & 1_n' \end{pmatrix} \begin{pmatrix} V_x' X_s' \\ 0 \end{pmatrix} = J_N \otimes I_{m+1}, \] and \( P_x' P_r' = J_N \otimes I_{M-n} \). Using these matrices, we project the terms in the estimating equations such that

\[ \begin{pmatrix} P_x' & 0 \\ 0 & 1_n' \end{pmatrix} \begin{pmatrix} V_x' X_s' \\ 0 \end{pmatrix} = J_N \otimes I_{m+1}, \] and \( P_x' P_r' = J_N \otimes I_{M-n} \). Using these matrices, we project the terms in the estimating equations such that

\[ \begin{pmatrix} P_x' & 0 \\ 0 & 1_n' \end{pmatrix} \begin{pmatrix} V_x' X_s' \\ 0 \end{pmatrix} = J_N \otimes I_{m+1}, \] and \( P_x' P_r' = J_N \otimes I_{M-n} \). Using these matrices, we project the terms in the estimating equations such that

\[ \begin{pmatrix} P_x' & 0 \\ 0 & 1_n' \end{pmatrix} \begin{pmatrix} V_x' X_s' \\ 0 \end{pmatrix} = J_N \otimes I_{m+1}, \] and \( P_x' P_r' = J_N \otimes I_{M-n} \). Using these matrices, we project the terms in the estimating equations such that

\[ \begin{pmatrix} P_x' & 0 \\ 0 & 1_n' \end{pmatrix} \begin{pmatrix} V_x' X_s' \\ 0 \end{pmatrix} = J_N \otimes I_{m+1}, \] and \( P_x' P_r' = J_N \otimes I_{M-n} \). Using these matrices, we project the terms in the estimating equations such that

\[ \begin{pmatrix} P_x' & 0 \\ 0 & 1_n' \end{pmatrix} \begin{pmatrix} V_x' X_s' \\ 0 \end{pmatrix} = J_N \otimes I_{m+1}, \] and \( P_x' P_r' = J_N \otimes I_{M-n} \). Using these matrices, we project the terms in the estimating equations such that

\[ \begin{pmatrix} P_x' & 0 \\ 0 & 1_n' \end{pmatrix} \begin{pmatrix} V_x' X_s' \\ 0 \end{pmatrix} = J_N \otimes I_{m+1}, \] and \( P_x' P_r' = J_N \otimes I_{M-n} \). Using these matrices, we project the terms in the estimating equations such that

\[ \begin{pmatrix} P_x' & 0 \\ 0 & 1_n' \end{pmatrix} \begin{pmatrix} V_x' X_s' \\ 0 \end{pmatrix} = J_N \otimes I_{m+1}, \] and \( P_x' P_r' = J_N \otimes I_{M-n} \). Using these matrices, we project the terms in the estimating equations such that
and \( v_{\delta_1}^* = \frac{1}{N-1} \left( I_n - \frac{J_n}{N} \right) \),

\[
V^s_{sr} = -\frac{1}{NM} \left( \sum_{x=1}^{N} \sigma^2_x \otimes I_n \otimes J_{m \otimes (M-m)} \right) \left[ \begin{array}{c} 0 \\ \text{m} \otimes \text{M} \otimes \text{NM} \end{array} \right] + \frac{1}{N-1} \left( D_{\mu^2} - \frac{\mu\mu'}{N} \right) \left( I_n - \frac{J_n}{N} \right) \otimes J_{m \otimes (M-m)} \left( D_{\mu^2} - \frac{\mu\mu'}{N} \right) \left( I_n - \frac{J_n}{N} \right) \otimes J_{m \otimes M} \right).
\]

\[
\frac{1}{N-1} \left( D_{\mu^2} - \frac{\mu\mu'}{N} \right) \left( I_n - \frac{J_n}{N} \right) \otimes \left( I_m - \frac{J_m}{M} \right) \left( D_{\mu^2} - \frac{\mu\mu'}{N} \right) \left( I_n - \frac{J_n}{N} \right) \otimes \left( I_m - \frac{J_m}{M} \right) \right).
\]

Thus,

\[
P^s'V^s'P^s = (M-1) \left( \sum_{x=1}^{N} \sigma^2_x \right) \left( I_n \otimes v_{\delta_1}^* \right) + \left( I_n' \left[ D_{\mu^2} - \frac{\mu\mu'}{N} \right] I_n \right) \left( v_{\delta_1}^* \right) \left( J_m \right) = \sigma^2 \left( I_n \otimes \left( I_m - \frac{J_m}{M} \right) \right) + \sigma^2 \left( I_n' \left( I_n - \frac{J_n}{N} \right) \otimes J_m \right)
\]

since

\[
\sigma^2 = \frac{1}{N} \left( \sum_{x=1}^{N} \sigma^2_x \right).
\]

Also, \( P^s'X^s'1_N = \left( I_n' \otimes I_{nm} \right) \left( I_n \otimes \frac{1_{nm}}{N} \right) = 1_{nm} = X_s \), and \( 1_N' \hat{\lambda} = \lambda \)

while,

\[
P^s'V^s'P^s = -\frac{1}{NM} \left( I_n' \left( \sum_{x=1}^{N} \sigma^2_x \right) \otimes I_n \otimes J_{m \otimes (M-m)} \right) \left[ \begin{array}{c} 0 \\ \text{m} \otimes \text{M} \otimes \text{NM} \end{array} \right] + \frac{1}{N-1} \left( D_{\mu^2} - \frac{\mu\mu'}{N} \right) \left( I_n - \frac{J_n}{N} \right) \otimes \left( I_m - \frac{J_m}{M} \right) \left( D_{\mu^2} - \frac{\mu\mu'}{N} \right) \left( I_n - \frac{J_n}{N} \right) \otimes \left( I_m - \frac{J_m}{M} \right)
\]

which simplifies to

\[
P^s'V^s'P^s = -\sigma^2 \left( I_n' \otimes \frac{J_n}{M} \right) \left[ \begin{array}{c} 0 \\ \text{m} \otimes \text{M} \otimes \text{NM} \end{array} \right] + \sigma^2 \left( I_n - \frac{J_n}{N} \right) \otimes \left( I_m - \frac{J_m}{M} \right) \left( -\frac{J}{N} \otimes \text{m} \otimes \text{M} \right) = V^v
\]

since \( \sigma^2 = \frac{1}{N-1} \left( D_{\mu^2} - \frac{\mu\mu'}{N} \right) I_N' \left[ \begin{array}{c} 1_N' \\ \end{array} \right] \). Also,
\[ 1_N X_r ^{\prime} P_r ^{\prime} = 1_N \left( \begin{array}{c} I_N \otimes \frac{1}{1N(M-m)} \\ I_N \otimes \frac{1}{1N(n-M)} N \end{array} \right)^{\prime} \left( \begin{array}{c} I_N \otimes I_N(M-m) \\ 0 \\ N(N-n)M \end{array} \right) \]

\[ = \frac{1}{N} \left( I_n(M-m) \otimes 1_{(N-n)M} ^{\prime} \right) \left( I_N \otimes I_N(M-m) \right) \left( \begin{array}{c} I_N \otimes I_N(M-m) \\ 0 \\ N(N-n)M \end{array} \right) \]

\[ = \left( I_{n(M-m)} \otimes 1_{(N-n)M} ^{\prime} \right) \]

\[ = 1_{NM-M} ^{\prime} \]

\[ = X_r ^{\prime} \]

Finally, using the expressions for that \( g_r ^{\prime} = \left( \begin{array}{c} 1_N ^{\prime} \otimes \frac{1}{1N(M-m)} \otimes \frac{1}{1N(M-m)} \\ 0 \end{array} \right) \). Then

\[ P_r ^{\prime} g_r ^{\prime} = \left( \begin{array}{c} I_N \otimes I_n(M-m) \\ 0 \\ \left( N-n \right) \otimes 0 \end{array} \right) \left( \begin{array}{c} 1_N \otimes I_N(M-m) \\ 0 \\ \left( N-n \right) \otimes 0 \end{array} \right) \]

\[ = \left( \begin{array}{c} 1 \\ 0 \\ \vdots \\ 0 \end{array} \right) \left( \begin{array}{c} 1 \\ 0 \\ \vdots \\ 0 \end{array} \right) \left( \begin{array}{c} 1 \\ 0 \\ \vdots \\ 0 \end{array} \right) \left( \begin{array}{c} 1 \\ 0 \\ \vdots \\ 0 \end{array} \right) \]

\[ = N g_r \]

We summarize these results. Upon projecting the terms in the estimating equations onto a lower dimensional space, we obtain a set of pseudo-estimating equations given by
The solution is given in c01ed15.doc. Since

\[
\begin{pmatrix}
V_s & X_s \\
X_s' & 0
\end{pmatrix}^{-1}
= \begin{pmatrix}
V_s^{-1} - V_s^{-1}X_s(X_s'V_s^{-1}X_s)^{-1}X_s'V_s^{-1} & V_s^{-1}X_s(X_s'V_s^{-1}X_s)^{-1} \\
(X_s'V_s^{-1}X_s)^{-1}X_s'V_s^{-1} & -(X_s'V_s^{-1}X_s)^{-1}
\end{pmatrix},
\]

\[
\hat{a}^* = \left[ V_s^{-1} - V_s^{-1}X_s(X_s'V_s^{-1}X_s)^{-1}X_s'V_s^{-1} \right] V_s g_r + V_s^{-1}X_s(X_s'V_s^{-1}X_s)^{-1} X_s' g_r.
\]

We can now express the best linear unbiased predictor. Recall that

\[ T = g \hat{g}' R = g \hat{g}' R + g \hat{g}' R. \]

and that we define an estimator of this random variable that is a linear function of the potentially observed sample data given by \( \hat{T} = L \hat{R}. \). Also, since

\[ a' = (L - g) , \quad \hat{r} = (g + a') \hat{R}, \]

where we predict \( g \hat{g}' R \) by \( \hat{a}' \hat{R} \). Note that \( \hat{a}' \) has dimension \( 1 \times N \times m \), while \( \hat{a}' \) has dimension \( 1 \times m \). Now

\[ \hat{a}' Y_s = \hat{a}' \left( \frac{J_N}{N} \otimes I_{nm} \right) R_s. \]

The matrix \( \frac{P P'}{N} = \frac{J_N}{N} \otimes I_{nm} \) is an ‘arbitrary’ matrix introduced to reduce the dimension of the matrix in the estimating equations so as to be able to form a ‘unique’ solution.

We form predictors by predicting \( g \hat{g}' R \) by \( \hat{a} \left( \frac{J_N}{N} \otimes I_{nm} \right) R_s \) instead of \( \hat{a} \hat{R} \). We use \( \hat{a}' Y_s \) as the predictor \( \hat{a}' \left( \frac{J_N}{N} \otimes I_{nm} \right) R_s. \) Let us define \( \beta = (X_s'V_s^{-1}X_s)^{-1} V_s^{-1}X_s Y_s. \) Then

\[ \hat{a}' Y_s = g \hat{g}' V_s' Y_s - g \hat{g}' V_s' X_s \beta + g \hat{X_s} \beta \]

\[ = g \hat{g}' X_s \beta + V_s' V_s^{-1} (Y_s - X_s \beta). \]
Making use of expression in c01ed15.doc, this expression simplifies to
\[ \hat{a}'Y_s = g_r' \left[ X, \hat{\beta} + V_{rr}^{-1} (Y_s - X, \hat{\beta}) \right] \]
\[ = \left( \frac{M - m}{M} \right) \left[ \bar{Y} + \left[ \frac{m}{(m \sigma^2 + f_m \sigma^2_e)} \left( \sigma^2 - \sigma^2_e \right) \frac{1}{M} \right] \left( \bar{Y}_i - \bar{Y} \right) \right] \]
\[ = \left( \frac{M - m}{M} \right) \left[ \bar{Y} + k \left( \bar{Y}_i - \bar{Y} \right) \right] \]

where \( k = \frac{m \sigma^2}{(m \sigma^2 + f_m \sigma^2_e)} \left( 1 - \frac{\sigma^2_e}{M \sigma^2} \right) \). The estimate of \( T = g_s' \mathbf{R}_s + g_r' \mathbf{R}_r \) is given by
\[ \hat{T} = g_s' \mathbf{R}_s + \hat{a}'Y_s \]. Since \( g_s' \mathbf{R}_s = \frac{m}{M} \bar{Y}_i \),
\[ \hat{T} = \frac{m}{M} \bar{Y}_i + \frac{M - m}{M} \left[ \bar{Y} + k \left( \bar{Y}_i - \bar{Y} \right) \right] \]. The first part of this estimator is a function of the sum of the realized SSU values for the first selected PSU. The second part of the estimator is the predictor of the remaining values of the SSUs corresponding to the realized PSU. Rather than using the realized SSU sample mean for the selected PSU as a predictor of the remaining SSUs, we use the shrinkage estimates (Stein) to predict these values.