Variance Estimation Requirements for Summary Totals and Rates for the Annual Survey of Occupational Injuries and Illnesses

Notation

Let

h = reported size class (1, 2, 3, 4, 5, all)

i = original size class (1, 2, 3, 4, 5)

j = establishment

k = TEI

K = aggregate industry

$\hat{}$ = an estimate

E = reported employment

H = hours worked

R = rate

V = variance

$\%RSE$ = percent relative standard error

W = original weight

$F_{kij} = D_{kij} * A_{kij} * P_{kij} * B_{kij}$

Where

D = Reaggregation Factor

A = Unit Nonresponse Adjustment Factor

P = Outlier Adjustment Factor

B = State Benchmark Factor

$Z_{kij} = W_{kij} F_{kij} \sqrt{\frac{W_{kij}-1}{W_{kij}}}$

(Note: if $W_{kij} = 0$, then $Z_{kij} = 0$)

n_{ki} = total sample size for i^{th} original size class

m_{ki} = usable sample size for i^{th} original size class

δ_{khij} = 1 if reported size class for the j^{th} unit is equal to original size class ($i = h$)

= 0 otherwise

μ = value from Table of Means for Hours, Employment, and the 22 Variables of Interest. The SMG will provide these values.

Y_g = total of g^{th} variable of interest, where g is one of the following:

Injuries and Illnesses Combined

Total recordable cases (TRC)
Lost workday cases (LWDC)
Cases with days away from work (DAFWC)
Cases with days of restricted work activity (DRWAC)
Nonfatal cases without lost workdays (CWOLWD)

Injuries Only

Total recordable cases (TRC)
Lost workday cases (LWDC)
Cases with days away from work (DAFWC)
Cases with days of restricted work activity (DRWAC)
Nonfatal cases without lost workdays (CWOLWD)

Illness Only

Total recordable cases (TRC)
Lost workday cases (LWDC)
Cases with days away from work (DAFWC)
Cases with days of restricted work activity (DRWAC)
Nonfatal cases without lost workdays (CWOLWD)

Types of Illnesses

Skin diseases
Dust diseases of the lung
Respiratory conditions due to toxic agents
Poisoning
Disorders due to physical agents
Disorders associated with repeated trauma
All other trauma

Note: if any calculated value for a variance (denoted by V) is less than zero, set the variance equal to zero.

1. For each area and for each TEI, k,

   (a) For each reported size class h and for each original size class $i = 1$ to 4:

      i. Compute:

$$SZ_{khi} = \sum_{j=1}^{m_{khi}} Z_{khij}\delta_{khij}$$

      ii. For Reported Employment E and Hours Worked H, compute:

$$SE_{khi} = \sum_{j=1}^{m_{khi}} \left(Z_{khij}E_{khij}\delta_{khij}\right)$$

$$SEE_{khi} = \sum_{j=1}^{m_{khi}} \left(Z_{khij}E_{khij}\delta_{khij}\right)^2$$

$$SH_{khi} = \sum_{j=1}^{m_{khi}} \left(Z_{khij}H_{khij}\delta_{khij}\right)$$

$$SHH_{khi} = \sum_{j=1}^{m_{khi}} \left(Z_{khij}H_{khij}\delta_{khij}\right)^2$$

$$SHE_{khi} = \sum_{j=1}^{m_{khi}} \left(Z_{khij}^2 H_{khij}E_{khij}\delta_{khij}\right) = \sum_{j=1}^{m_{khi}} \left(Z_{khij}H_{khij}\delta_{khij}\right)\left(Z_{khij}E_{khij}\delta_{khij}\right)$$

      iii. For each variable of interest g, compute:

$$SY_{kghi} = \sum_{j=1}^{m_{khi}} \left(Z_{khij}Y_{kghij}\delta_{khij}\right)$$

$$SYY_{kghi} = \sum_{j=1}^{m_{khi}} \left(Z_{khij}Y_{kghij}\delta_{khij}\right)^2$$

$$SYE_{kghi} = \sum_{j=1}^{m_{khi}} \left(Z_{khij}^2 Y_{kghij}E_{khij}\delta_{khij}\right) = \sum_{j=1}^{m_{khi}} \left(Z_{khij}Y_{kghij}\delta_{khij}\right)\left(Z_{khij}E_{khij}\delta_{khij}\right)$$

$$SYH_{kghi} = \sum_{j=1}^{m_{khi}} \left(Z_{khij}^2 Y_{kghij}H_{khij}\delta_{khij}\right) = \sum_{j=1}^{m_{khi}} \left(Z_{khij}Y_{kghij}\delta_{khij}\right)\left(Z_{khij}H_{khij}\delta_{khij}\right)$$

   (b) For each reported size class h:

      i. For Hours Worked H, compute:

$$L_{kHh} = \frac{\sum_{i=1}^{5}\sum_{j=1}^{m_{khi}} W_{khij}F_{khij}H_{khij}\delta_{khij}}{\sum_{i=1}^{5}\sum_{j=1}^{m_{ki}} W_{khij}F_{khij}E_{khij}}$$

      ii. For each variable of interest g, compute:

$$L_{kgh} = \frac{\sum_{i=1}^{5}\sum_{j=1}^{m_{ki}} W_{khij}F_{khij}Y_{kghij}\delta_{khij}}{\sum_{i=1}^{5}\sum_{j=1}^{m_{ki}} W_{kij}F_{kij}E_{kij}}$$

(c) For each reported size class h and for each original size class $i = 1$ to 4:

i. If $n_{ki} = 0$ or $m_{ki} = 0$, then set

$$V\left(\hat{E}_{khi}\right) = 0$$
$$V\left(\hat{H}_{khi}\right) = 0$$
$$C\left(\hat{H}\hat{E}_{khi}\right) = 0$$

For all variables of interest g, then set

$$V\left(\hat{Y}_{kghi}\right) = 0$$
$$C\left(\hat{Y}\hat{E}_{kghi}\right) = 0$$
$$C\left(\hat{Y}\hat{H}_{kghi}\right) = 0$$

ii. For Reported Employment E:

A. If $n_{ki} > 1$ and $m_{ki} > 0$, compute:

$$V\left(\hat{E}_{khi}\right) = \frac{1}{(n_{ki}-1)}\left[n_{ki}SEE_{khi} - \left(SE_{khi}\right)^2\right]$$

B. If $n_{ki} = 1$ and $m_{ki} = 1$, compute:

$$V\left(\hat{E}_{khi}\right) = \left(SE_{khi} - SZ_{khi}\mu_{i,employment}\right)^2$$

iii. For Hours Worked H:

A. If $n_{ki} > 1$ and $m_{ki} > 0$, compute:

$$V\left(\hat{H}_{khi}\right) = \frac{1}{(n_{ki}-1)}\left[n_{ki}SHH_{khi} - \left(SH_{khi}\right)^2\right]$$
$$C\left(\hat{H}\hat{E}_{khi}\right) = \frac{1}{(n_{ki}-1)}\left[n_{ki}SHE_{khi} - SH_{khi}SE_{khi}\right]$$

B. If $n_{ki} = 1$ and $m_{ki} = 1$, compute:

$$V\left(\hat{H}_{khi}\right) = \left(SH_{khi} - SZ_{khi}\mu_{i,hours}\right)^2$$
$$C\left(\hat{H}\hat{E}_{khi}\right) = SHE_{khi} - \left(SZ_{khi}\right)^2\mu_{i,hours}\mu_{i,employment}$$

iv. For each variable of interest Y_g

A. If $n_{ki} > 1$ and $m_{ki} > 0$, compute:

$$V\left(\hat{Y}_{kghi}\right) = \frac{1}{(n_{ki}-1)}\left[n_{ki}SYY_{kghi} - \left(SY_{kghi}\right)^2\right]$$
$$C\left(\hat{Y}\hat{H}_{kghi}\right) = \frac{1}{(n_{ki}-1)}\left[n_{ki}SYH_{kghi} - SY_{kghi}SH_{khi}\right]$$
$$C\left(\hat{Y}\hat{E}_{kghi}\right) = \frac{1}{(n_{ki}-1)}\left[n_{ki}SYE_{kghi} - SY_{kghi}SE_{khi}\right]$$

B. If $n_{ki} = 1$ and $m_{ki} = 1$, compute:

$$V\left(\hat{Y}_{kghi}\right) = \left(SY_{kghi} - SZ_{khi}\mu_{i,g}\right)^2$$
$$C\left(\hat{Y}\hat{H}_{kghi}\right) = SYH_{kghi} - \left(SZ_{khi}\right)^2 \mu_{i,g}\mu_{i,hours}$$
$$C\left(\hat{Y}\hat{E}_{kghi}\right) = SYE_{kghi} - \left(SZ_{khi}\right)^2 \mu_{i,g}\mu_{i,employment}$$

(d) For each reported size class h, compute:

i.

$$V\left(\hat{E}_{kh}\right) = \sum_{i=1}^{4} V\left(\hat{E}_{khi}\right)$$
$$C\left(\hat{Y}\hat{H}_{kgh}\right) = \sum_{i=1}^{4} V\left(\hat{Y}\hat{H}_{kghi}\right)$$
$$V\left(U\hat{H}_{kh}\right) = \sum_{i=1}^{4} V\left(\hat{H}_{khi}\right)$$
$$V\left(U\hat{Y}_{kgh}\right) = \sum_{i=1}^{4} V\left(\hat{Y}_{kghi}\right)$$

ii. If TEI k's BMF_FLAG(a, o, t) = NO or PRORATE_FLAG = YES, then

$$V\left(\hat{H}_{kh}\right) = \sum_{i=1}^{4} \left[\left(V\hat{H}_{khi}\right)\right], \text{and for all variables of interest } g,$$
$$V\left(\hat{Y}_{kgh}\right) = \sum_{i=1}^{4} \left[V\left(\hat{Y}_{kghi}\right)\right]$$

2. Within each Grantee State

(a) For Reported Employment, E:

$$V\left(\hat{E}_{Kh}\right) = \sum_{k} V\left(\hat{E}_{kh}\right)$$

(b) For Hours Worked, H:

$$V\left(\hat{H}_{Kh}\right) = \sum_{k} V\left(\hat{H}_{kh}\right)$$
$$V\left(U\hat{H}_{Kh}\right) = \sum_{k} V\left(U\hat{H}_{kh}\right)$$

(c) For each variable of interest, g:

$$V\left(\hat{Y}_{Kgh}\right) = \sum_k V\left(\hat{Y}_{kgh}\right)$$
$$C\left(\hat{Y}\hat{H}_{Kgh}\right) = \sum_k C\left(\hat{Y}\hat{H}_{kgh}\right)$$
$$V\left(U\hat{Y}_{Kgh}\right) = \sum_k V\left(U\hat{Y}_{kgh}\right)$$

(d) For each reported size class h, compute:

i.

$$V\left(\hat{E}_{kh}\right) = \sum_{i=1}^{4} V\left(\hat{E}_{khi}\right)$$
$$C\left(\hat{Y}\hat{H}_{kgh}\right) = \sum_{i=1}^{4} C\left(\hat{Y}\hat{H}_{kghi}\right)$$

ii. If TEI k's BMF_FLAG(a, o, t) = NO or PRORATE_FLAG = YES, then

$$V\left(\hat{H}_{kh}\right) = \sum_{i=1}^{4}\left[V\left(\hat{H}_{khi}\right)\right], \text{and for all variables of interest,} Y_g$$
$$V\left(\hat{Y}_{kgh}\right) = \sum_{i=1}^{4}\left[V\left(\hat{Y}_{kghi}\right)\right]$$

otherwise compute:

$$V\left(\hat{H}_{kh}\right) = L_{kHh}^2 V\left(\hat{E}_{k,all}\right) + \sum_{i=1}^{4}\left[V\left(\hat{H}_{khi}\right) - 2L_{kHh}C\left(\hat{H}\hat{E}_{khi}\right)\right], \text{and for all variables of interest,} g$$
$$V\left(\hat{Y}_{kgh}\right) = L_{kgh}^2 V\left(\hat{E}_{k,all}\right) + \sum_{i=1}^{4}\left[V\left(\hat{Y}_{khi}\right) - 2L_{kgh}C\left(\hat{Y}\hat{E}_{kghi}\right)\right]$$

3. Within each area and for each aggregate industry K and for each reported size class h, compute (where summation is over all the TEIs contained in K):

(a) For Reported Employment, E:

$$V\left(\hat{E}_{Kh}\right) = \sum_k V\left(\hat{E}_{kh}\right)$$

(b) For Hours Worked, H:

$$V\left(\hat{H}_{Kh}\right) = \sum_k V\left(\hat{H}_{kh}\right)$$

(c) For each variable of interest, g:

$$V\left(\hat{Y}_{Kgh}\right) = \sum_k V\left(\hat{Y}_{kgh}\right)$$
$$C\left(\hat{Y}\hat{H}_{Kgh}\right) = \sum_k C\left(\hat{Y}\hat{H}_{kgh}\right)$$

4. **Requirements for Database of Standard Estimates:** With each estimate $\left(\hat{E}, \hat{H}, \hat{Y}, \hat{R}\right)$ in the standard database, associate the estimate's Percent Relative Standard error as calculated below. These values should be available through the Extract System. The Percent Relative Standard Errors should have four integer places and be rounded to four decimal places. NOTE: If an estimate is zero, then its Percent Relative Standard Error is zero.

   (a) For TEI k, compute:

$$\%RSE\left(\hat{E}_{kh}\right) = \frac{\sqrt{V\left(\hat{E}_{kh}\right)}}{\hat{E}_{kh}} * 100$$

$$\%RSE\left(\hat{H}_{kh}\right) = \frac{\sqrt{V\left(\hat{H}_{kh}\right)}}{\hat{H}_{kh}} * 100$$

$$\%RSE\left(\hat{Y}_{kgh}\right) = \frac{\sqrt{V\left(\hat{Y}_{kgh}\right)}}{\hat{E}_{kgh}} * 100$$

$$\%RSE\left(\hat{R}_{kgh}\right) = \left[\sqrt{\frac{V\left(U\hat{Y}_{kgh}\right)}{\hat{Y}_{kgh}^2} + \frac{V\left(U\hat{H}_{kh}\right)}{\hat{H}_{kh}^2} - 2\frac{C\left(\hat{Y}\hat{H}_{kgh}\right)}{\hat{Y}_{kgh}\hat{H}_{kh}}}\right] * 100$$

   (b) For aggregate industry K, compute:

$$\%RSE\left(\hat{E}_{Kh}\right) = \frac{\sqrt{V\left(\hat{E}_{Kh}\right)}}{\hat{E}_{Kh}} * 100$$

$$\%RSE\left(\hat{H}_{Kh}\right) = \frac{\sqrt{V\left(\hat{H}_{Kh}\right)}}{\hat{H}_{Kh}} * 100$$

$$\%RSE\left(\hat{Y}_{Kgh}\right) = \frac{\sqrt{V\left(\hat{Y}_{Kgh}\right)}}{\hat{E}_{Kgh}} * 100$$

$$\%RSE\left(\hat{R}_{Kgh}\right) = \left[\sqrt{\frac{V\left(U\hat{Y}_{Kgh}\right)}{\hat{Y}_{Kgh}^2} + \frac{V\left(U\hat{H}_{Kh}\right)}{\hat{H}_{Kh}^2} - 2\frac{C\left(\hat{Y}\hat{H}_{Kgh}\right)}{\hat{Y}_{Kgh}\hat{H}_{Kh}}}\right] * 100$$

5. **Requirements for Abbreviated Sampling Errors Report:** For each area, generate a report containing the following:

   (a) Area Identification

   (b) TEI or Aggregate Industry Identification
   For each TEI and Aggregation Industry:

   (c) Total Employment %RSE

   (d) Total Hours Worked (in thousands) %RSE

   (e) Total TRC for Injuries and Illnesses Combined %RSE

   (f) TRC Rate for Injuries and Illnesses Combined %RSE

   (g) Total CWOLWD for Injuries and Illnesses Combined %RSE

   (h) Total LWDC for Injuries and Illnesses Combined %RSE

   (i) LWDC Rate for Injuries and Illnesses Combined %RSE

   (j) Total DAFWC for Injuries and Illnesses Combined %RSE

   (k) DAFWC Rate for Injuries and Illnesses Combined %RSE