Variance Component Estimation in Multistage Sampling

Richard Valliant, Jill A. Dever, and Frauke Kreuter

2020-07-28

To allocate a sample among different stages of sampling, the contributions of the different stages to the variance of an estimator must be considered. These components of variance generally depend on the analysis variable and also on the form of the estimator. This vignette covers some basic variance results for linear estimators in two-stage and three-stage sampling and how the components can be estimated with functions in `PracTools`. Technical background is in Valliant, Dever, and Kreuter (2018), ch.9. First, the package must be loaded with

```r
library(PracTools)
```

Alternatively, `require(PracTools)` can be used.

Two-stage Sampling

Consider a two-stage sample design in which the first-stage units are selected using πps sampling, i.e., with varying probabilities and without replacement. We will also refer to this as *ppswor* sampling. Elements are selected at the second stage via simple random sampling without replacement (*srswor*). Quite a bit of notation is needed, even in this fairly simple case:

U = universe of PSUs

M = number of PSUs in universe

U_i = universe of elements in PSU i

N_i = number of elements in the population for PSU i

$N = \sum_{i \in U} N_i$ is the total number of elements in the population

π_i = selection probability of PSU i

π_{ij} = joint selection probability of PSUs i and j

m = number of sample PSUs

n_i = number of sample elements in PSU i

s = set of sample PSUs

s_i = set of sample elements in PSU i

y_k = analysis variable for element k in PSU i (subscript i is implied)

$\bar{y}_U$ = mean per element in the population

$\bar{y}_{Ui}$ = mean per element in the population in PSU i

The π-estimator of the population total, $t_U = \sum_{i \in U} \sum_{k \in U_i} y_k$, of an analysis variable y is

$$\hat{t}_\pi = \sum_{i\in s} \frac{\hat{t}_i}{\pi_i}$$

where $\hat{t}_i = (N_i/n_i) \sum_{k\in s_i} y_k$, which is the estimate of the total for PSU i with a simple random sample. The design variance of the estimated total can be written as the sum of two components:

$$V\left(\hat{t}_\pi\right) = \sum_{i\in U}\sum_{j\in U} (\pi_{ij} - \pi_i\pi_j) \frac{t_i}{\pi_i}\frac{t_j}{\pi_j} + \sum_{i\in U} \frac{N_i^2}{\pi_i n_i}\left(1 - \frac{n_i}{N_i}\right) S_{U2i}^2 \tag{1}$$

where

$$S_{U2i}^2 = \sum_{k\in U_i} \left(y_k - \bar{y}_{Ui}\right)^2 / \left(N_i - 1\right)$$

is the unit variance of y among the elements in PSU i.

Formula (1) is difficult or impossible to use for sample size computations because the number of PSUs in the sample is not exposed. Another is to analyze *srswor* sampling of PSUs and SSUs as in Example 1 below. Determining sample sizes this way does not mean that you are necessarily locked into selecting PSUs and elements within PSUs via *srswor* or *srswr*. Basing sample sizes on a design that is less complicated than the one that will actually be used is a common approach, although it can be deceptive for some analysis variables.

Special case: *srswor* at first and second stages

Suppose the first stage is an *srswor* of m out of M PSUs and the second stage is a sample of n_i elements selected by *srswor* from the population of N_i. The π-estimator is

$$\hat{t}_\pi = \frac{M}{m}\sum_{i\in s}\frac{N_i}{n_i}\sum_{k\in s_i} y_k$$

Its variance is equal to

$$V\left(\hat{t}_\pi\right) = \frac{M^2}{m}\frac{M-m}{M}S_{U1}^2 + \frac{M}{m}\sum_{i\in U}\frac{N_i^2}{n_i}\frac{N_i - n_i}{N_i}S_{U2i}^2 \tag{2}$$

where $S_{U1}^2 = \frac{\sum_{i\in U}(t_i - \bar{t}_U)^2}{M-1}$ with t_i being the population total of y in PSU i and $\bar{t}_U = \sum_{i\in U} t_i/M$ is the mean total per PSU.

If $\bar{n}$ elements are selected in each PSU and the sampling fractions of PSUs and elements within PSUs are all small, then the relvariance can be written as

$$\frac{V(\hat{t}_\pi)}{t_U^2} \doteq \frac{B^2}{m} + \frac{W^2}{m\bar{n}} \tag{3}$$

where $B^2 = S_{U1}^2/\bar{t}_U^2 = M^2 S_{U1}^2/t_U^2$ is the unit relvariance among PSU totals and $W^2 = M\sum_{i\in U} N_i^2 S_{U2i}^2/t_U^2$. The term B^2 is called the "between (PSU) component" while W^2 is the "within component". Expression (3) is the form used in the R function, `BW2stageSRS`. Textbooks often list a specialized form of (3) that requires that all PSUs have the same size, $N_i \equiv \bar{N}$, and that $\bar{n}$ elements are selected in each. In that case, the second-stage sampling fraction is $\bar{n}/\bar{N}$. This implies that the sample is self-weighting: $\pi_i\pi_{k|i} = m\bar{n}/M\bar{N}$. The relvariance based on (2) then simplifies to the less general form

$$\frac{V(\hat{t}_\pi)}{t_U^2} = \frac{1}{m}\frac{M-m}{M}B^2 + \frac{1}{m\bar{n}}\frac{\bar{N}-\bar{n}}{\bar{N}}W^2$$

where $W^2 = \frac{1}{M\bar{y}_U^2}\sum_{i\in U} S_{U2i}^2$.

Assuming that $\bar{n}$ elements are selected in each sample PSU, and m/M and $\bar{n}/N_i$ are both small, the more general form of the relvariance in (3) can also be written in terms of a measure of homogeneity δ as follows:

$$\frac{V\left(\hat{t}_\pi\right)}{t_U^2} \doteq \frac{\tilde{V}}{m\bar{n}} k\left[1+\delta\left(\bar{n}-1\right)\right] \tag{4}$$

where $\tilde{V} = S_U^2/\bar{y}_U^2$, $k = (B^2+W^2)/\tilde{V}$, and

$$\delta = \frac{B^2}{B^2+W^2}. \tag{5}$$

With some effort, it can be shown that when $N_i = \bar{N}$ and both M and $\bar{N}$ are large,

$$\frac{S_U^2}{\bar{y}_U^2} = \frac{1}{\bar{y}_U^2}\frac{\sum_{i\in U}\sum_{k\in U_i}\left(y_k-\bar{y}_U\right)^2}{(N-1)} \doteq B^2+W^2$$

i.e., the population relvariance can be written as the sum of between and within relvariances. If $k=1$, (4) equals the expression found in many textbooks. However, when the population count of elements per cluster varies, k may be far from 1, as will be illustrated in an example below. In those cases, (4) with an estimate of the actual k should be used for determining sample sizes and computing advance estimates of coefficients of variation.

Expressions (3) and (4) are useful for sample size calculation since the number of sample PSUs and sample units per PSU are explicit in the formula. Equation (4) also connects the variance of the estimated total to the variance that would be obtained from a simple random sample since $\tilde{V}/m\bar{n}$ is the relvariance of the estimated total in an *srswor* of size $m\bar{n}$ when the sampling fraction is small. The product $k[1+\delta(\bar{n}-1)]$ is a type of design effect. When $k=1$, the term $1+\delta\left(\bar{n}-1\right)$ is the approximate design effect found in many textbooks.

The next example uses the `MDarea.pop` from `PracTools`. This dataset is based on the U.S. Census counts from the year 2000 for Anne Arundel County in the US state of Maryland. The geographic divisions used in this dataset are called tracts and block groups. Tracts are constructed by the US Census Bureau to have a desired population size of 4,000 people. Block groups (BGs) are smaller with a target size of 1,500 people. Counts of persons in the dataset are the same for most tracts and block groups as in the 2000 Census.

- **Example. Between and within variance components in *srs/srs* design** The R function `BW2stageSRS` will calculate the unit relvariance of a population, B^2+W^2 for comparison, the ratio $k=(B^2+W^2)/(S_U^2/\bar{y}_U^2)$, and the full version of δ in (5). The function assumes that the entire sampling frame is an input. The full R code for this example is in the file `Example 9.2.R`, available at bookfiles. We first compute the results using the `PSU` and `SSU` variables as clusters. These fields are created so that all `PSU`s have the same size; likewise, all `SSU`s have the same size. For the variable `y1` in the Maryland population, the code is

```
require(PracTools)
data(MDarea.pop)
BW2stageSRS(MDarea.pop$y1, psuID=MDarea.pop$PSU)
#>          B2          W2 unit relvar       B2+W2           k       delta
#> 0.007859299 1.455263645 1.462741186 1.463122944 1.000260988 0.005371592
BW2stageSRS(MDarea.pop$y1, psuID=MDarea.pop$SSU)
#>          B2          W2 unit relvar       B2+W2           k       delta
#>  0.03653178  1.42770886  1.46274119  1.46424065  1.00102510  0.02494930
```

The values of δ are 0.005 for `PSU` and 0.025 for `SSU`. Next, to illustrate the dramatic effect that varying sizes of clusters can have, we compute the same statistics as above using tracts and block groups (BGs) within tracts as clusters. These vary substantially in the number of persons in each cluster. A new variable called `trtBG` is computed since the values of the variable, `BLKGROUP`, are nested within each tract:

```
  trtBG <- 10*MDarea.pop$TRACT + MDarea.pop$BLKGROUP
  BW2stageSRS(MDarea.pop$y1, psuID=MDarea.pop$TRACT)
#>          B2          W2 unit relvar        B2+W2            k        delta
#>   0.2604683   1.8390286   1.4627412   2.0994969   1.4353167   0.1240623
  BW2stageSRS(MDarea.pop$y1, psuID=trtBG)
#>          B2          W2 unit relvar        B2+W2            k        delta
#>   0.3488622   1.9498600   1.4627412   2.2987221   1.5715167   0.1517635
```

The value of δ is 0.124 `TRACT`s are clusters and 0.152 when `trtBG` defines clusters. The measures of homogeneity increase substantially when tracts or BGs are clusters compared to the `PSU` and `SSU` results. This is entirely due to the increase in B^2 when units with highly variable sizes are used and an *srs* is selected. For example, $B^2 = 0.0079$ for `y1` when `PSU` is a cluster but is 0.2605 when `TRACT` is a cluster.

More General Two-stage Designs

Variances of estimators in two-stage designs more complicated than simple random sampling at each stage can be written as a sum of components. However, these have limited usefulness in determining sample sizes for the same reason that (1) is not. A more convenient formulation is the case where PSUs are selected with varying probabilities but with replacement, and the sample within each PSU is selected by *srswor*. With-replacement designs may not often be used in practice but have simple variance formulae. The *pwr*-estimator of a total (Särndal, Swensson, and Wretman 1992) is

$$\hat{t}_{pwr} = \frac{1}{m} \sum_{i \in s} \frac{\hat{t}_i}{p_i}$$

where $\hat{t}_i = \frac{N_i}{n_i} \sum_{k \in s_i} y_k$ is the estimated total for PSU i from a simple random sample and p_i is the one-draw selection probability of PSU i. The variance of $\hat{t}_{pwr}$ is

$$V\left(\hat{t}_{pwr}\right) = \frac{1}{m} \sum_{i \in U} p_i \left(\frac{t_i}{p_i} - t_U\right)^2 + \sum_{i \in U} \frac{N_i^2}{m p_i n_i} \left(1 - \frac{n_i}{N_i}\right) S_{U2i}^2. \tag{6}$$

Making the assumption that $\bar{n}$ elements are selected in each PSU, the variance reduces to

$$V\left(\hat{t}_{pwr}\right) = \frac{S_{U1(pwr)}^2}{m} + \frac{1}{m\bar{n}} \sum_{i \in U} \left(1 - \frac{\bar{n}}{N_i}\right) \frac{N_i^2 S_{U2i}^2}{p_i}$$

where, in this case, $S_{U1(pwr)}^2 = \sum_{i \in U} p_i \left(\frac{t_i}{p_i} - t_U\right)^2$. Dividing this by t_U^2 and assuming that the within-PSU sampling fraction, $\bar{n}/N_i$, is negligible, we obtain the relvariance of $\hat{t}_{pwr}$ as, approximately,

$$\frac{V\left(\hat{t}_{pwr}\right)}{t_U^2} \doteq \frac{B^2}{m} + \frac{W^2}{m\bar{n}} = \frac{\tilde{V}}{m\bar{n}} k \left[1 + \delta\left(\bar{n} - 1\right)\right] \tag{7}$$

with $\tilde{V} = S_U^2 / \bar{y}_U^2$, $k = (B^2 + W^2)/\tilde{V}$,

$$B^2 = \frac{S_{U1(pwr)}^2}{t_U^2}, \tag{8}$$

$$W^2 = \frac{1}{t_U^2} \sum_{i \in U} N_i^2 \frac{S_{U2i}^2}{p_i}, \tag{9}$$

$$\delta = B^2 \big/ \left(B^2 + W^2\right) \tag{10}$$

Expression (7) has the same form as (4) but with different definitions of B^2 and W^2. Expression (7) also has the interpretation of an *srs* variance of an unclustered variance, $\tilde{V}/m\bar{n}$, times a design effect, $k[1+\delta(\bar{n}-1)]$, in the same way that (4) did.

- **Example. Between and within variance components in *ppswr/srs* design** This example repeats the calculations in the example above for the variables in the Maryland area population. Assume that clusters will be selected proportional to the count of persons in each cluster. The function `BW2stagePPS` computes the population values of B^2, W^2, and δ shown in (8), (9), and (10) which are appropriate for *ppswr* sampling of clusters. The code for `y1` using `PSU` or `SSU` as clusters is shown below. The variables, `pp.PSU` and `pp.SSU`, hold the one-draw probabilities p_i that appear in (6):

```
pp.PSU <- table(MDarea.pop$PSU) / nrow(MDarea.pop)
pp.SSU <- table(MDarea.pop$SSU) / nrow(MDarea.pop)
BW2stagePPS(MDarea.pop$y1, pp=pp.PSU, psuID=MDarea.pop$PSU)
#>          B2          W2 unit relvar       B2+W2           k       delta
#> 0.007762335 1.455263403 1.462741186 1.463025738 1.000194534 0.005305672
BW2stagePPS(MDarea.pop$y1, pp=pp.SSU, psuID=MDarea.pop$SSU)
#>          B2          W2 unit relvar       B2+W2           k       delta
#>  0.03644120  1.42770995  1.46274119  1.46415115  1.00096392  0.02488896
```

The code for PSUs that are tracts and block groups is

```
pp.trt <- table(MDarea.pop$TRACT) / nrow(MDarea.pop)
pp.BG <- table(trtBG) / nrow(MDarea.pop)
BW2stagePPS(MDarea.pop$y1, pp=pp.trt, psuID=MDarea.pop$TRACT)
#>          B2          W2 unit relvar       B2+W2           k       delta
#> 0.009171403 1.453908596 1.462741186 1.463079999 1.000231629 0.006268559
BW2stagePPS(MDarea.pop$y1, pp=pp.BG, psuID=trtBG)
#>          B2          W2 unit relvar       B2+W2           k       delta
#>  0.01602891  1.44780622  1.46274119  1.46383513  1.00074787  0.01094994
```

The between term when clusters are defined by `PSU` is about the same as when clusters are selected by *srs* because `PSU`'s all have the same size. With PSUs being either tracts or block groups in the *ppswr/srswor* design, the between term is much smaller than the within, compared to the results in the *srs/srs* example. For example, with `y1` and *srs* sampling of tracts, $B^2 = 0.2604$ but for *pps* sampling of tracts $B^2 = 0.0091$.

When clusters are selected by *srs*, S^2_{U1} is the variance of the cluster totals around the average cluster total. In contrast, with *pps* sampling of clusters, $S^2_{U1(pwr)}$ is the variance of the estimated population totals, $t_i\,/p_i$ around the population total, t_U. When clusters are selected with probability proportional to N_i, then $t_i\,/p_i = N_i\bar{y}_{Ui}$. If these one-cluster estimates of the population total are fairly accurate, as they are here, the B^2 term can be quite small. This leads to much smaller values of δ in *pps* sampling of clusters. This implies that the negative effect of clustering on the variance is lessened for a design that selects clusters with $pp(N_i)$. This kind of comparison explains most practitioners' preference for *pps* sampling of clusters, especially when the clusters vary in population size.

General Three-stage Designs

In the case of with-replacement sampling of PSUs with varying probabilities and srswor at the second and third stages, the relvariance can be written (with a few assumptions) in a form useful for sample size calculations. Treating the case where SSUs are selected via srs (either with or without replacement) is not too unrealistic since SSUs (like block groups) are often created to have about the same population sizes.

The variance formulae for a three-stage design with *ppswor* selection of first-stage units is complex enough that it is not useful for sample size planning. See Valliant, Dever, and Kreuter (2018), sec. 9.2.4 for details. To obtain a simpler formula, suppose that $\bar{n}$ SSUs are sampled in each sample PSU, the sampling fractions of

SSUs in each PSU, $\bar{n}\,/N_i$, are small, and $\bar{\bar{q}}$ elements are selected in each sample SSU. The relvariance of the *pwr*-estimator is then

$$\frac{V\left(\hat{t}_{pwr}\right)}{t_U^2} = \frac{B^2}{m} + \frac{W_2^2}{m\bar{n}} + \frac{W_3^2}{m\bar{n}\bar{\bar{q}}}, \tag{11}$$

where $B^2 = S_{U1(pwr)}^2\,/t_U^2$ is given by (8),

$$W_2^2 = \frac{1}{t_U^2}\sum_{i\in U} N_i^2 S_{U2i}^2\,/p_i\ ; \tag{12}$$

$$W_3^2 = \frac{1}{t_U^2}\sum_{i\in U}\frac{N_i}{p_i}\sum_{j\in U_i} Q_{ij}^2 S_{U3ij}^2. \tag{13}$$

The relvariance can also be written in terms of two measures of homogeneity:

$$\frac{V\left(\hat{t}_{pwr}\right)}{t_U^2} = \frac{\tilde{V}}{m\bar{n}\bar{\bar{q}}}\left\{k_1\delta_1\bar{n}\bar{\bar{q}} + k_2\left[1 + \delta_2\left(\bar{\bar{q}} - 1\right)\right]\right\} \tag{14}$$

where

$k_1 = (B^2 + W^2)/\tilde{V}$ with $\tilde{V} = \frac{1}{Q-1}\sum_{i\in U}\sum_{j\in U_i}\sum_{k\in U_{ij}}\left(y_k - \bar{y}_U\right)^2/\bar{y}_U^2$ is the unit relvariance of y in the population.

$k_2 = (W_2^2 + W_3^2)/\tilde{V}$

$\delta_1 = B^2/(B^2 + W^2)$

$W^2 = \frac{1}{t_U^2}\sum_{i\in U} Q_i^2 S_{U3i}^2\,/p_i$ with $S_{U3i}^2 = \frac{1}{Q_i-1}\sum_{j\in U_i}\sum_{k\in U_{ij}}\left(y_k - \bar{y}_{Ui}\right)^2$ and $\bar{y}_{Ui} = \sum_{j\in U_i}\sum_{k\in U_{ij}} y_k\,/Q_i$, i.e., S_{U3i}^2 is the element-level variance among all elements in PSU i

$\delta_2 = W_2^2/(W_2^2 + W_3^2)$

Note that the term W^2 in δ_1 does not enter the variance in (11) but is defined by analogy to the term in two-stage sampling. If elements were selected directly from the sample PSUs (instead of first sampling SSUs), then W^2 above would be the appropriate within-PSU component.

The term δ_1 is a measure of the homogeneity among the PSU totals. If the estimate of the population total from each PSU total, $t_i\,/p_i$, was exactly equal to the population total, t_U, then $B^2 = 0$ and $\delta_1 = 0$. That is, if the variation within PSUs is much larger than the variation among PSU totals, then δ_1 will be small; this is the typical situation in household surveys *if PSUs all have about the same number of elements.* As we saw in the earlier example, the condition of equal-sized PSUs can be critically important to insure that B^2 is small.

If the SSUs all have about the same totals, t_{ij}, then W_2^2 will be small and $\delta_2 \doteq 0$. Although attempts may be made to create SSUs that have about the same number of elements Q_{ij}, the totals t_{ij} of other variables tend to vary, leading to values of δ_2 that are larger than those of δ_1.

The R function, `BW3stagePPS`, will calculate B^2, W^2, W_2^2, W_3^2, δ_1, and δ_2 defined above for *ppswr/srs/srs* and *srswr/srs/srs* sampling. The function is appropriate if an entire frame is available and takes the following parameters:

Parameter	Description
`X`	data vector; length is the number of elements in the population.
`pp`	vector of one-draw probabilities for the PSUs; length is number of PSUs in population.

Parameter	Description
`psuID`	vector of PSU identification numbers. This vector must be as long as X. Each element in a given PSU should have the same value in psuID. PSUs must be in the same order as in X.
`ssuID`	vector of SSU identification numbers. This vector must be as long as X. Each element in a given SSU should have the same value in ssuID. PSUs and SSUs must be in the same order as in X. ssuID should have the form `psuID||(ssuID within PSU)`.

- **Example. Variance components in three stage *srswr/srs/srs* and *ppswr/srs/srs* designs.** In the Maryland population suppose that suppose that tracts and BGs within tracts are the first- and second-stage units, and that persons are elements in a three-stage design. All three stages are selected by *srs*. The call to `BW3stagePPS` for the variable `y1` in an *srswr/srs/srs* design is:

```
M <- length(unique(MDarea.pop$TRACT))
trtBG <- 10*MDarea.pop$TRACT + MDarea.pop$BLKGROUP
pp.trt <- rep(1/M,M)
BW3stagePPS(X=MDarea.pop$y1, pp=pp.trt,
      psuID=MDarea.pop$TRACT, ssuID=trtBG)
#>         B         W        W2        W3 unit relvar        k1 
#> 0.2577266 1.8390286 0.2698581 2.1083645 1.4627412 1.4334423 
#>        k2    delta1    delta2 
#> 1.6258670 0.1229169 0.1134705 
```

We repeat the calculation but assuming *ppswr* sampling of PSUs. The calculation for `y1` using tracts and block groups as the first- and second-stage sampling units is done via this call:

```
trtBG <- 10*MDarea.pop$TRACT + MDarea.pop$BLKGROUP
pp.trt <- table(MDarea.pop$TRACT) / nrow(MDarea.pop)
BW3stagePPS(X=MDarea.pop$y1, pp=pp.trt,
      psuID=MDarea.pop$TRACT, ssuID=trtBG)
#>           B           W          W2          W3 unit relvar          k1 
#> 0.009171403 1.453908596 0.249889887 1.687254100 1.462741186 1.000231629 
#>          k2      delta1      delta2 
#> 1.324324498 0.006268559 0.128999129 
```

Notice that $\delta_1 = 0.123$ with *srs* sampling of tracts but is 0.006 when tracts are sampled proportional to their population sizes.

An important practical, sample design problem that we do not cover in this vignette is how to estimate variance components and measures of homogeneity from a complex, multistage sample. This topic is covered in detail in section 9.4 of Valliant, Dever, and Kreuter (2018). The `PracTools` package includes a variety of other functions relevant to two- and three-stage sampling that are also not discussed in this vignette:

Function	Description
BW2stagePPSe	Estimate components of relvariance for a sample design where primary sampling units (PSUs) are selected with *pps* and elements are selected via *srs*. The input is a sample selected in this way.

Function	Description
BW3stagePPSe	Estimate components of relvariance for a sample design where primary sampling units (PSUs) are selected with probability proportional to size with replacement (ppswr) and secondary sampling units (SSUs) and elements within SSUs are selected via simple random sampling (srs). The input is a sample selected in this way.
clusOpt2	Compute the sample sizes that minimize the variance of the *pwr*-estimator of a total in a two-stage sample.
clusOpt2fixedPSU	Compute the optimum number of sample elements per primary sampling unit (PSU) for a fixed set of PSUs.
clusOpt3	Compute the sample sizes that minimize the variance of the *pwr*-estimator of a total in a three-stage sample.
clusOpt3fixedPSU	Compute the sample sizes that minimize the variance of the *pwr*-estimator of a total in a three-stage sample when the PSU sample is fixed.
CVcalc2	Compute the coefficient of variation of an estimated total in a two-stage design. Primary sampling units (PSUs) can be selected either with probability proportional to size (*pps*) or with equal probability. Elements are selected via simple random sampling (*srs*).
CVcalc3	Compute the coefficient of variation of an estimated total in a three-stage design. Primary sampling units (PSUs) can be selected either with probability proportional to size (*pps*) or with equal probability. Secondary units and elements within SSUs are selected via simple random sampling (*srs*).
deff	Compute the Kish, Henry, Spencer, or Chen-Rust design effects.

References

Särndal, C.-E., B. Swensson, and J. Wretman. 1992. *Model Assisted Survey Sampling.* New York: Springer-Verlag.

Valliant, R., J. A. Dever, and F. Kreuter. 2018. *Practical Tools for Designing and Weighting Survey Samples.* 2nd ed. New York: Springer-Verlag.