NATIONAL SURVEY OF OLDER AMERICANS ACT PARTICPANTS (NSOAAP)

THIRTEENTH NATIONAL SURVEY (2018)

1. SAMPLE SELECTION, WEIGHTING, AND VARIANCE ESTIMATION

The survey employed a two-stage sample design, first selecting a sample of Area Agencies on Aging (AAAs) in stage one and, in the second stage, a sample of clients for each service within each sampled AAA. The thirteenth national survey covered six services – Home Delivered Meals, Homemaker Services, Transportation, the Family Caregiver Support Program, Congregate Meals and Case Management.

Weighting of each service record was done separately. Initially, base weights were computed by taking the inverse of the selection probability for each sampled client. Then the base weights were adjusted for nonresponse, followed by trimming of the extreme weights. Finally, a poststratification adjustment was made using available control totals. Fay's modified Balanced Repeated Replication (BRR) method was used for computation of the sampling variances of survey estimates.

Agency Selection

At the first stage of the two-stage design for the national survey, a stratified sample of 325 AAAs (allowing for a 20% non-response) was selected from the frame of 629 agencies. This sample size was increased from 316 in the twelfth survey in order to provide greater precision for the sample estimates. The 2018 frame is identical to the 2017 frame.

The sampling frame was essentially the same as that used for the sixth through –twelfth national surveys, except for an agency added or removed from year to year. The agency measures of size were completely updated in 2011 using new budget figures based on the most recent reports from the AAAs at the time of the sixth survey. These same budget figures were also used for the thirteenth survey.

The AAA sample was selected independently within five budget size strata, which were created based on the square root of the total budget sizes of the AAAs. The AAA and client samples were proportionally allocated to the total of the square root of the budget sizes in each stratum. However, within a stratum the sample of AAAs was selected with equal probability, but sorted by Census region and within region by the measure of size variable, MOS18, in serpentine order. Note that the measure of size variable, MOS18, is the square root of the budget size for the given AAA. This method was used instead of direct probability proportional to size (PPS) sampling because in the earlier national surveys it was found that budget as a measure of size was not well correlated with the total number of clients in each agency for every service. In the absence of any other information, budget size was still used in sample selection, but with less importance. First, the square root of the budget size (instead of budget size itself) was used to reduce the effect of large variation in budget sizes. Second, the sample was allocated at the stratum level proportional to the overall total of the square root of the budget size. This procedure gave a higher probability of selection to agencies with larger budget sizes, but all the agencies within a budget size stratum received the same probability of selection. As in the prior surveys, some agencies were selected with certainty. The total sample size was allocated to the five strata as shown in the following table:

Table 1 Sampling strata and allocation of agencies into strata for the national sample.

STRATUM	Square Root of Budget Size	AAA Frame	Probability of Selection	Allocation of AAA Sample
Certainty	Greater than or equal to $4,737	41	1.0000	41
Non-certainty Stratum 1	$2,648 - $4,736	75	0.9467	71
Non-certainty Stratum 2	$1,873 - $2,647	116	0.6034	70
Non-certainty Stratum 3	$1,480 - $1,872	157	0.4586	72
Non-certainty Stratum 4	Less than $1,480	240	0.2958	71
TOTAL		629		325

The 41 agencies with the largest budget sizes were selected with certainty for the AAA sample. The remaining sample was then selected independently within each of the non-certainty strata. The implicit stratification (sorting) variables in the selection process were the four Census Regions (Northeast, Midwest, South and West), and within region by the measure of size variable, MOS18, using a serpentine sort for MOS18. As a result, the number of agencies in each Region was selected roughly in proportion to the total of the square root of budget of the Region, while providing the additional sort by measure of size within Region. Table 2 shows the agency distribution in the frame and in the originally-selected sample by Census Region.

Table 2 Distributions of agencies in the universe and in the original sample by region.

Census Region	Number of AAAs in the Frame	Number of AAAs in the Sample
Northeast	171	88
Midwest	104	64
South	229	110
West	125	63
Total	**629**	**325**

Client Selection

Client samples by service type (Home Delivered Meals, Homemaker, Transportation, Caregiver Service, Congregate Meals, and Case Management) were drawn randomly within each sampled AAA. Before selecting the sample of clients, Westat obtained the total number of clients receiving each service within an agency by contacting either the sampled agency or the State Unit on Aging (SUA) for the state in which the sampled agency is located. Based on the total number of clients, line numbers from client master lists were sampled using a Westat software application that started with the total number of clients in each service by agency and randomly selected the matching line numbers for the sampled clients. The number of clients selected from a service within each agency is such that the expected overall probability of selection of a client within a service is roughly the same for all clients within each sampling stratum. Also, to allow for a nonresponse or ineligibility rate (e.g., due to mortality, nursing home placement, or other service termination), the goal was to increase the number of clients selected by the inverse of the rates observed in the previous cycle of the national survey in order to meet the required sample size for each service. However, to continue to do so for the 13th National Survey would have resulted in the proposed numbers getting to be too large for some services. Thus, for the 13th National Survey, there are slight increases in the within-AAA sampling target coupled with an increase in the number of AAAs sampled to achieve the goal of 6,000 respondents.

In the certainty agencies, the number of clients selected in each agency varied depending on the budget sizes of the agencies. However, in the non-certainty agencies, fixed-size client samples were selected from each agency for each service as indicated in Table 3 below.

Table 3 Within-AAA sample sizes by stratum type for the six target services

Service	Certainty Stratum*	Non-certainty Stratum**
Family Caregiver**	(325*38*MOS18)/SUM(MOS18)	38
Home Delivered Meals	(325*12*MOS18)/SUM(MOS18)	12
Homemaker Service	(325*8*MOS18)/SUM(MOS18)	8
Transportation	(325*22*MOS18)/SUM(MOS18)	22
Congregate Meals	(325*16*MOS18)/SUM(MOS18)	16
Case Management	(325*14*MOS18)/SUM(MOS18)	14

* In the formulas for the certainty strata above, the quantity MOS18 is the square root of the budget size for the given AAA, and the expression SUM (MOS18) is the sum of the size measures over all AAAs on the frame. Thus, the formula for the client sample size for a certainty AAA is the ratio of the individual measure of size to the sum of all the measures of size times 325 times the fixed sample size for the given service. The result is then rounded up to the next largest integer.

Selection Probability

The probability of selection of a client within a service can be mathematically expressed as follows. First, for non-certainty agencies, let

$P_{i \in h}$ = Probability of selection of agency i in stratum h,

$$= \frac{\text{Number of non-certainty agencies selected from the stratum}}{\text{Total number of non-certainty agencies in the stratum}}$$

$= \frac{m_h}{M_h}$, for agencies in a non-certainty stratum.

For certainty agencies, the probability of selection was 1 (that is, $P_{h=c} = 1$). Next, let

P_{ijs} = Probability of selection of client j in service s within agency i,

$$= \frac{\text{Number of clients selected from service } s \text{ in agency } i}{\text{Total number of clients in service } s \text{ in agency } i} = \frac{n_{is}}{N_{is}}.$$

Recall that n_{is} was fixed in advance for non-certainty agencies by service, as shown in Table 3.

Thus, the overall probability of selection of client j in service s within agency i in stratum h was

$$\pi_{ijs} = P_{i \in h} \times P_{ijs} = \frac{m_h}{M_h} \times \frac{n_{is}}{N_{is}} \qquad \text{for the clients within non-certainty agencies,}$$

$$= 1 \times \frac{n_{is}}{N_{is}} = \frac{n_{is}}{N_{is}} \qquad \text{for the clients within certainty agencies.}$$

Weighting

Weighting was done in four steps: calculation of base weights, nonresponse adjustment, trimming of extreme weights, and poststratification adjustments to known population control totals.

Base Weights

The base weight is the inverse of the overall selection probability of a client. The base weight of a client can be obtained by calculating the base weight for an agency and multiplying that weight by the within-agency-level base weight of a client in a service within that agency.

The base weight for an agency i can be expressed as

$$a_{i,i \in h} = \frac{1}{P_h} = \frac{M_h}{m_h} \qquad \text{for non-certainty agencies,}$$

$$= 1 \qquad \text{for certainty agencies,}$$

and the base weight for a client in a service within an agency can be expressed as

$$v_{ijs} = \frac{1}{P_{ijs}} = \frac{N_{is}}{n_{is}},$$

= the within-agency base weight of client j in service s within agency i.

Therefore, the overall base weight of a client within a service is

$$w_{ijs} = a_i \times v_{ijs} = \frac{1}{\pi_{ijs}},$$

$$= \frac{M_h}{m_h} \times \frac{N_{is}}{n_{is}} \qquad \text{for non-certainty agencies,}$$

$$= 1 \times \frac{N_{is}}{n_{is}} \qquad \text{for certainty agencies.}$$

Nonresponse Adjustment

Since not all sampled agencies and clients responded to the survey, the base weights had to be adjusted for nonresponse. The nonresponse adjustment was done in two steps by performing separate adjustments for agency-level and client-level nonresponse. The nonresponse adjustments were applied specific to each service group within cells defined by Agency size and Census region.

If m_{hs}^r denotes the number of agencies in stratum h that responded to the survey for service s, then the agency-level nonresponse adjustment was calculated as follows:

$$a_{is,i \in h}^r = \frac{M_h}{m_h} \times \frac{m_h}{m_{hs}^r} = \frac{M_h}{m_{hs}^r}$$

= the nonresponse adjusted weight of agency i for service s.

If n_{is}^r denotes the number of clients that responded for service s within agency i, then the client-level nonresponse adjustment was calculated as follows:

$$v_{ijs}^r = \frac{N_{is}}{n_{is}} \times \frac{n_{is}}{n_{is}^r} = \frac{N_{is}}{n_{is}^r},$$

= the nonresponse adjusted weight for client j for service s within agency i.

Therefore, the overall nonresponse-adjusted weight of client j for service s within agency i is

$$w_{ijs}^r = a_{is}^r \times v_{is}^r = \frac{M_h}{m_{hs}^r} \times \frac{N_{is}}{n_{is}^r}.$$

Trimming of Weights

To keep the variance of the survey estimates within an acceptable level, extreme weights were trimmed. The sample design was set up to select clients within a service with equal probability so that the base weights of all clients within a service would be roughly equal. This would have been the case if the measure of size used in selecting the agencies (i.e., the square root of each agency's annual budget) was perfectly correlated with the number of clients in a service and if there had been no nonresponse. But in reality, this correlation was not high, and there was some nonresponse. Some agencies had larger budgets due to larger numbers of clients in some services but smaller numbers of clients in other services. Similarly, some agencies had smaller budgets but relatively larger numbers of clients in a particular service. This contributed to increased variability in the selection probabilities and subsequently in the base weights. Moreover, the variability in weights was increased further due to the adjustment of client nonresponse rates that varied substantially from agency to agency. Since variability in the weights increases the variances of the survey estimates, those weights which were too high compared to the median base weight over all clients within a given service were trimmed to acceptable upper limits to reduce the variance of the survey estimates.

Initially, the acceptable upper limits were determined by using the median base weight within a service group such that weights larger than 4 times the median base weight in the service group were trimmed to be equal to 4 times the median base weight in the group. However, for all six services, this trimming rule was empirically shown to over-trim with respect to the percentiles of the distribution of all weights for that service. Thus, for Homemaker and Home Delivered Meals the weights were trimmed at the 99th percentile. For Family Caregiver, Congregate Meals, and Transportation, the weights were trimmed at the 98th percentile. For Case Management the weights were trimmed at the 95th percentile. One effect of trimming weights is that estimated totals are reduced from what they would have been, had trimming not been applied to the weights. This loss in the sum of weights due to the trimming was adjusted in the final poststratification step described below. The trimmed, nonresponse adjusted weights will be denoted by w_{ijs}^{θ} in the following sections.

Poststratification Adjustment

The final step of weighting involved the benchmarking of the estimated number of clients in a service (based on the trimmed, nonresponse-adjusted weights) to the known total number of clients (control total) obtained from the AoA State Program Reports (SPR). The poststratification adjustment, or benchmarking, was done at the regional level, since reliable control totals were available at the regional level.

The post-stratified weights (w_{ijs}^{p}) for service s were calculated by multiplying the trimmed, nonresponse-adjusted weights (w_{ijs}^{θ}) by the ratio of the known control total (N_s) to the estimated total $(\sum_{ij} w_{ijs}^{\theta})$ as follows:

$$w_{ijs}^{p} = w_{ijs}^{\theta} \times \frac{N_s}{\sum_{ij} w_{ijs}^{\theta}}$$

The poststratification adjustment described in this paragraph was applied to Home-delivered Meals, Homemaker Services, Congregate Meals, Case Management, and Family Caregiver. The adjustments for Transportation services were calculated somewhat differently and are described below.

Poststratification Adjustment for Transportation Service

For the Transportation service, control totals for the number of clients were not available. However, State Units on Aging (SUAs) did provide the number of one-way passenger trips in the State Program Reports (SPR). These SPR regional level trip counts were used for the purpose of estimating control totals for the number of clients receiving transportation services by region. The following summarizes the methodology used for constructing these estimated transportation client counts:

- The national survey asked respondents how many one-way trips per month they usually took using the AAA transportation service.
- An average annual per-person trip count by region was estimated from the survey data file using the trimmed, nonresponse-adjusted weights.
- By dividing the total trip count by the per-person average annual number of trips, Westat estimated the total number of persons who received transportation services by region.

The method of estimation explained above can be mathematically expressed as follows:

$$\hat{N}_s = \sum_g \hat{N}_{gs} = \sum_g \frac{T_g}{\bar{t}_g} = \sum_g \frac{T_g}{\dfrac{\sum_{ij} t_{ij} w_{ijs}^{\boldsymbol{\theta}}}{\sum_{ij} w_{ijs}^{\boldsymbol{\theta}}}} = \sum_g \frac{T_g}{\hat{T}_{gw}} \times \hat{N}_{gw} \,,$$

where

$\hat{N}_s$ is the final estimate of transportation client count,

$\hat{N}_{gs}$ is the final estimate of transportation client count in region g,

T_g is the total number of one-way trips reported by the SUAs in region g,

$\bar{t}_g = \dfrac{\sum_{ij,i\in g} t_{ij} w_{ijs}^{\boldsymbol{\theta}}}{\sum_{ij,i\in g} w_{ijs}^{\boldsymbol{\theta}}}$ is the per-person weighted average of annual number of trips in region g,

t_{ij} is the number of annual one-way trips made by client j in agency i,

$\hat{T}_{gw} = \sum_{ij,i\in g} t_{ij} w_{ijs}^{\boldsymbol{\theta}}$ is an initial estimate of the total number of one-way trips in region g based on the trimmed, nonresponse-adjusted weights;

$\hat{N}_{gw} = \sum_{ij,i\in g} w_{ijs}^{\boldsymbol{\theta}}$ is an initial estimate of the total number of transportation clients in region g based on the trimmed, nonresponse-adjusted weights.

The above estimator is widely known as a *Ratio Estimator* in the sample survey literature because the initial estimate of the total number of transportation clients ($\hat{N}_w$) is adjusted by the ratio of actual to estimated total number of one-way trips ($\dfrac{T}{\hat{T}_w}$).

Variance Estimation

Westat routinely uses replication-based variance estimation methods for computing sampling variances of the survey estimates derived from complex multi-stage sample designs. Westat's variance computation software, WesVar, is designed for this purpose. A version of balanced repeated replication (BRR) referred to as "Fay's method" was used to calculate the variances (and their square roots, the standard errors) of estimates derived from theNSOAAP. Implementation of BRR methods for variance estimation requires the use of a series of "replicate weights," each of which provides an alternative (replicate-specific) estimate of a characteristic of interest. The variability of the replicate estimates about the full-sample estimate of the same characteristic is then used to obtain the variance or standard error of the characteristic.

Let y_{ij} denote a survey characteristic (variable) for the jth respondent in the ith agency, and let w_{ij}^{p} denote the corresponding full-sample final weight. Further, let w_{ij}^{k} denote the kth replicate weight, where $k = 1, 2, ..., K$. The estimated total for the survey variable is given by the weighted sum

$$\hat{y} = \sum_{ij} w_{ij}^{p} y_{ij} \ .$$

The corresponding replicate estimates are given by the weighted sums

$$\hat{y}_k = \sum_{ij} w_{ij}^{k} y_{ij} \text{ , for } k = 1, 2, ..., K$$

The variance of the estimate $\hat{y}$ is then computed as:

$$\text{var}(\hat{y}) = \frac{1}{(1-.30)^2} \sum_{k=1}^{K} (\hat{y}_k - \hat{y})^2 \ ,$$

where the 0.30 in the above formula is referred to as "Fay's factor." The corresponding standard error is simply the square root of $\text{var}(\hat{y})$ as computed above.

The replicate weights, w_{ij}^{k}, required for variance estimation were derived from replicate-specific base weights and include all of the adjustments (e.g., nonresponse, trimming, and poststratification) used to develop the final full-sample weights, w_{ij}^{p}.

Replicates were formed by first creating variance strata and variance units. For non-certainty AAAs, variance strata were formed with two or three AAAs in each stratum, and each AAA was treated as a variance unit. For certainty AAAs, each AAA was treated as a variance stratum, and random groups of clients were formed as variance units within the stratum. This difference in forming variance strata for certainty and non-certainty AAAs was necessary to account for the fact that there was no first stage sampling variance for certainty AAAs. Under BRR, the replicates are formed in a balanced way by taking one variance unit from each variance stratum. However, a modified version of BRR called Fay's method was used for the AoA survey. Under the modified approach, the full-sample weights are adjusted or "perturbed" to define the required replicates, rather than taking one variance unit from each stratum. Further details on BRR and Fay's method, or replication methods in general, can be found in WesVar 5.1 User's Guide. The User's Guide is available without charge by emailing wesvar_tech_support@westat.com; see this link: https://www.westat.com/capability/information-systems-software/wesvar/wesvar-documentation. Note that the User's Guide is for WesVar 4.2, with an addendum for what's new in WesVar 5.1.

WesVar, SUDAAN, STATA, SAS, SPSS and other complex sample survey software packages can use replicate weights to compute variance estimates that fully account for the complex design used in the AoA national surveys.

2. SIGNIFICANCE TESTING OF THE DIFFERENCE BETWEEN TWO SURVEY CHARACTERISTICS

The statistic given below can be used to test whether the observed difference between two estimated proportions is statistically significant. This test can be used to check the significance of the difference either between an agency level and a national level characteristic or between characteristics estimated for two different agencies. The test statistic is

$$z = \frac{|\hat{p}_1 - \hat{p}_2|}{\sqrt{SE^2(\hat{p}_1) + SE^2(\hat{p}_2)}}$$

where, $\hat{p}_1$ and $\hat{p}_2$ are estimates of the two survey characteristics to be compared, and $SE^2(\hat{p}_1)$ and $SE^2(\hat{p}_2)$ are squares of the corresponding standard errors of the two estimates.

When the sample size (i.e., the number of valid responses in each comparison group) is 30 or more, the above test statistic will approximately follow a statistical distribution called the *normal distribution* and the difference will be considered significant at the 5% level of significance if $z > 1.96$. The interpretation of such a result is that the probability of obtaining a difference as large as the observed difference by chance alone is less than 5%.

However, if the number of valid responses in one of the groups is less than 30, then the above test statistic will follow a different statistical distribution called the *t-distribution* with $(n_1 + n_2 - 2)$ degrees of freedom, where n_1 and n_2 are the number of valid responses in the two groups. In this case, the critical value for the significance of a difference will depend on $(n_1 + n_2 - 2)$. The following table presents a rough indication of the critical values of the t distribution for a 5% level of significance for different values of $(n_1 + n_2 - 2)$. The computed value of z must be greater than the corresponding critical value for the difference between the two estimates to be considered significant.

Degrees of freedom, $(n_1 + n_2 - 2)$	Critical value of t distribution at the 5% level of significance
>58	1.96
30-58	2.05
25-29	2.06
20-24	2.08
15-19	2.13

For interested readers, more detailed tables of critical values of the normal, t, and other statistical distributions are available in standard textbooks on statistical methods.