

Technical Paper 91-1

**SAMPLING METHODOLOGY  
FOR  
FEDERAL WAGE SYSTEM  
WAGE SURVEYS**

*Prepared by the Information Systems Division*

**DoD Wage Fixing Authority  
Technical Staff**

2 October 1991

**(Supersedes Technical Paper 90-1)**

## CONTENTS

I. INTRODUCTION . . . . .	1
A. Scope . . . . .	1
B. Wage Area . . . . .	1
C. Types of Surveys . . . . .	1
D. Bureau of Labor Statistics Function . . . . .	2
E. Purpose . . . . .	2
II. SAMPLING METHODOLOGY . . . . .	3
A. Source of Universe . . . . .	3
B. Universe Definition . . . . .	3
1. Location. . . . .	3
2. Industry. . . . .	3
3. Number of Employees. . . . .	6
C. Sampling Process . . . . .	6
1. Definition of certainty units. . . . .	7
2. Definition of additional industries. . . . .	7
3. Determination of sample size. . . . .	7
4. Stratification. . . . .	7
5. Allocation of the sample over the strata. . . . .	7
6. Measure of size. . . . .	8
7. Sample selection. . . . .	9
8. Computation of the weight. . . . .	9
9. Establishment Survey List. . . . .	9
III. WEIGHT ADJUSTMENTS FOR NONRESPONSE . . . . .	10
A. General . . . . .	10
B. Noncertainty Units . . . . .	10
C. Certainty Units . . . . .	12
D. Unique Units . . . . .	12
E. Combined Returns . . . . .	12
1. Recognizable Units. . . . .	12
2. Hidden Units. . . . .	13
APPENDIX A . . . . .	1
Appropriated Fund SIC Groupings . . . . .	1
Non-Appropriated Fund CT SIC Groupings . . . . .	2
Non-Appropriated Fund AS/PS SIC Groupings . . . . .	3
APPENDIX B . . . . .	1
Allocation of Establishments to Strata . . . . .	1

## I. INTRODUCTION

### A. Scope

The Federal Government uses the Federal Wage System (FWS) to fix and adjust the rates of pay for prevailing rate employees. These employees are in trade, craft, and laboring type jobs, that is, blue collar workers. In 1968, the Civil Service Commission (now the Office of Personnel Management (OPM)) issued the basic policies of the FWS as Chapter 532 of the Federal Personnel Manual (FPM). The FWS provides common policies, systems, practices, and job grading standards for uniform application by all agencies subject to section 5342 of Title 5, United States Code. The system covers most trade, craft, and laboring employees in the executive branch of the U.S. Government who are paid from appropriated funds (AF). In 1972, the system was extended by Public Law 92-392 to cover nonappropriated fund (NAF) craft and trade employees. To assure equitable treatment for all Department of Defense (DOD) NAF hourly paid employees, the principles of the law pertaining to the fixing and adjusting of pay rates were extended by DOD to hourly paid employees engaged in Administrative Support (AS) and Patron Services (PS) positions. Such positions as sales clerks, secretaries, display assistants, and cashiers fall into this group. OPM issued regulation 5 CFR Part 532, dated May 11, 1981, to fulfill its responsibility for the overall administration of the Prevailing Rate System under sections 5342 and 5346 of Title 5, United States Code. OPM determines the basic policies, practices, and procedures with the advice of the Federal Prevailing Rate Advisory Committee, other executive agencies, and labor organizations.

### B. Wage Area

Since rates of pay are based upon prevailing rates in a locality, OPM has defined geographical areas (wage areas) for which rates of pay are set. The wage areas are defined in terms of state, county, and town boundaries. There are 135 AF wage areas and 141 NAF wage areas. The NAF wage areas are defined separately from the AF areas. For each wage area, OPM designates an agency as the lead agency responsible for fixing and administering rates of pay. This designation is generally given to the agency which has the largest number of employees covered by the FWS in the wage area. OPM has designated the Department of Defense (DOD) as the lead agency for 109 AF wage areas and for all NAF wage areas. Within DOD, the Department of Defense Wage Fixing Authority Technical Staff (DODWFATS) executes this authority with the advice of the DOD Wage Committee. The Department of Veterans Affairs (VA) is the lead agency for 25 wage areas and the National Aeronautics and Space Administration (NASA) is the lead agency for one wage area.

### C. Types of Surveys

The lead agencies conduct wage surveys of private enterprise establishments in their respective wage areas annually to determine the level of prevailing wages paid for survey jobs that are representative by occupation or grade level of local Government employment. The scope of the survey is defined to include establishments in certain industries with a minimum number of employees. Two types of wage surveys are conducted on an alternating cycle. A full-scale survey involves the listing of establishments in the scope of the survey for the wage area and collecting pay data by personally visiting the establishments. A wage-change survey is conducted the second year in a wage area including only those establishments that participated in the previous full-scale survey. Wage-change data may be collected by telephone, mail, or personal visit. Thus, in any one year, about half of the wage areas have a full-scale survey. AF surveys are conducted in 109 DOD areas, 25 VA areas and one NASA area. NAF surveys are conducted in 141 areas. In universe areas, all in-scope establishments are included in the survey. In sample wage areas, a sample of the

in-scope establishments is statistically selected to be in the survey since including all establishments in the survey would be too costly and require too much time. Sample surveys are conducted in over half of the wage areas.

#### D. Bureau of Labor Statistics Function

In the beginning of the FWS, the Bureau of Labor Statistics (BLS) of the Department of Labor had the responsibility to furnish the lead agencies with the list of establishments to be included in the wage survey for each survey area. The list contains the name, address, number of employees, and main industry of the establishments. The list also has a numerical weight for each establishment on the list, which shows the numerical value of the establishment in representing establishments not selected to be on the list. In a universe survey, of course, all initial weights are one since all in-scope establishments are on the list. After the close of the survey data collection, BLS adjusted the establishment weights for establishment nonresponse. BLS used statistical sampling methodology to select establishments for wage areas having sample surveys and used a post-survey weighting procedure to adjust weights for nonresponse in both universe and sample surveys. In April 1989, OPM transferred these responsibilities from BLS to DODWFATS.

#### E. Purpose

The purpose of this technical paper is to describe the following methodologies used by DODWFATS:

- generating sample area establishment lists
- adjusting weights for non-response.

## II. SAMPLING METHODOLOGY

### A. Source of Universe

When BLS started providing lists of establishments to the lead agencies for their surveys, it used information obtained from state unemployment insurance agencies. The information was supplemented by data from trade associations, trade directories, and labor organizations. In 1983, BLS had to change the source of the information from state unemployment insurance agencies because of the Privacy Act and confidentiality concerns. Dun and Bradstreet was chosen as the source for the establishment lists, since it maintains records on all business establishments in the United States. Each record contains the name, address, number of employees, and the primary industry of the company. The industry is identified by the Standard Industrial Classification (SIC) code as defined by the U.S. Office of Management and Budget (OMB). In April 1989, the DODWFATS officially took over this BLS function. We obtained the computer programs that BLS used to produce the list of establishments for wage surveys. Therefore, we continued to use Dun and Bradstreet as the data source for establishment lists. We now obtain establishment records from Dun and Bradstreet periodically, usually semiannually. DODWFATS maintains a database of in-scope establishments for all wage areas. The database is updated by DODWFATS wage specialists with information collected on survey as well as by updates from Dun and Bradstreet.

### B. Universe Definition

Three factors determine if an establishment is in the definition of the universe for a wage area: the physical location of the establishment, the primary industry of the establishment, and the number of employees it has.

**1. Location.** A wage area consists of two parts, the survey area and the area of application, which are defined geographically in terms of counties or independent cities, or in New England states, for AF wage areas, of entire townships or cities. Survey data are collected from in-scope establishments located in the survey area. The area of application includes the survey area plus that part of the wage area from which no data are collected, but to which the wage schedule applies. In Washington, D.C., for example, the AF survey area includes the city Washington, D.C.; the independent cities Alexandria, Fairfax, and Falls Church in Virginia; the counties Charles, Frederick, Montgomery, and Prince Georges in Maryland; and the counties Arlington, Fairfax, Loudoun, and Prince William in Virginia. The area of application includes all the survey area plus the counties Calvert and St. Marys in Maryland; and the counties Facquier, King George, and Stafford in Virginia. NAF wage areas are usually smaller than AF wage areas. For example, the NAF wage area for Washington, D.C. is defined as the survey area being the city Washington, D.C. and the area of application being the same as the survey area.

**2. Industry.** Establishments engaged in the following industries are included in all AF wage surveys for regular wage schedules as defined in FPM Supplement 532-1, subchapter S5.

The industries are identified by 1987 SIC codes:

SIC 20 through 26 and SIC 28 through 38	All manufacturing classes except SIC 27 (printing) and SIC 39 (miscellaneous)
SIC 40	Railroad transportation
SIC 41 (except 412)	Local and suburban transit and interurban highway passenger transportation except taxicabs (SIC 412)
SIC 42	Motor freight transportation and warehousing
SIC 45	Transportation by air
SIC 48	Communications
SIC 49	Electric, gas, and sanitary services
SIC 50	Wholesale trade—durable goods
SIC 51	Wholesale trade—nondurable goods

Other industries may be added to the standard industry coverage when those industries account for significant proportions of local area employment or when it is expected that sizeable numbers of job matches might be found. For example, SIC 1011, iron ore mining, is added to the scope of the wage survey in Northwestern Michigan and SIC 8062, private hospitals, is added to the scope in many wage surveys.

Establishments engaged in the following industries are included in all NAF wage surveys for regular wage schedules as defined in the FPM Supplement 532-2, subchapter S5. The industries are identified by 1987 SIC codes:

Wholesale companies - durable goods

SIC 5013	Motor vehicle supplies and new parts
SIC 5021	Furniture
SIC 5023	Home furnishings
SIC 5043	Photographic equipment and supplies
SIC 5064	Electrical appliances, television and radio sets
SIC 5065	Electronic parts and equipment, not elsewhere classified
SIC 5072	Hardware
SIC 5091	Sporting and recreational goods
SIC 5092	Toys and hobby goods and supplies
SIC 5094	Jewelry, watches, precious stones
SIC 5099	Durable goods, not elsewhere classified

Wholesale companies - nondurable goods

SIC 5111	Printing and writing paper
----------	----------------------------

## Wholesale companies - nondurable goods

SIC 5112	Stationery and office supplies
SIC 5113	Industrial and personal service paper
SIC 5122	Drugs, drug proprietaries, and druggists' sundries
SIC 5131	Piece goods, notions, and other dry goods
SIC 5136	Men's and boys' clothing and furnishings
SIC 5137	Women's, children's, and infants' clothing
SIC 5139	Footwear
SIC 5145	Confectionary
SIC 5159	Farm-product raw materials, not elsewhere classified
SIC 5171	Petroleum bulk stations and terminals
SIC 5172	Petroleum and petroleum products wholesalers, except bulk stations and terminals
SIC 5191	Farm supplies
SIC 5192	Books, periodicals, and newspapers
SIC 5193	Flowers, nursery stock, and florists' supplies
SIC 5194	Tobacco and tobacco products
SIC 5198	Paints, varnishes, and supplies
SIC 5199	Nondurable goods not elsewhere included

## Retail companies

SIC 5311	Department stores
SIC 5331	Variety stores
SIC 5541	Gasoline service stations
SIC 5812	Eating places
SIC 5813	Drinking places (alcoholic beverages)
SIC 5962	Automatic merchandising machine operators

## Service companies

SIC 7011	Hotels and motels
SIC 7933	Bowling centers
SIC 7997	Membership sports and recreation clubs (golf and country clubs only)

In addition to the above industries, defined in the FPM supplement, on-base food contractors (SIC 7363) are included in NAF wage surveys.

The following industries are included in all NAF surveys for the AS and PS schedules as defined by DODWFA. The industries are identified by 1987 SIC codes:

## Retail companies

SIC 5611	Men's and boys' clothing and accessory stores
SIC 5621	Women's clothing stores
SIC 5651	Family clothing stores
SIC 5912	Drug stores and proprietary stores
SIC 5963	Direct selling establishments

## Banking and insurance companies

SIC 6021	National commercial banks
SIC 6022	State commercial banks
SIC 6029	Commercial banks
SIC 6311	Life insurance
SIC 6331	Fire, marine, and casualty insurance

## Service companies

SIC 7032	Sporting and recreational camps
SIC 8351	Child day care services

**3. Number of Employees.** The minimum number of employees an establishment must have to be in the universe for an AF wage survey is 50. The Alaska wage survey is an exception with a minimum number of 20. For NAF wage surveys the minimum number is 20 employees except for establishments in industries SIC 5541, 5962, 7933, and 7997, for which the minimum number is eight. For the NAF AS/PS schedules, the minimum number of employees is also 20 except for industries SIC 5963, 7032, and 8351, for which eight is the minimum number.

**C. Sampling Process**

The sampling design used in the wage surveys is sampling proportionate to establishment employment. The universe is divided into groups based upon the industry in which the establishment is engaged. These groups are in turn divided into subgroups based upon the establishment employment. These subgroups are called strata. Within each stratum, a sample of establishments is selected independently from the other strata. The number of selected establishments is allocated over the strata based upon total stratum employment. The basic steps for the selection of sample establishments are:

- Definition of certainty units, if any.
- Definition of additional industries to be in the scope of the universe for the wage survey, if any, such as private hospitals, SIC 8062.
- Determination of sample size.
- Stratification of the universe by industry and employment size.
- Allocation of the sample over the strata based upon the number of employees.
- Computation of the measure of size for each establishment to determine its probability of selection.
- Selection of establishments within each stratum.
- Computation of the weight for each selected unit.
- Listing the selected establishments.



**1. Definition of certainty units.** FPM Supplements 532-1 and 532-2 permit the inclusion of individual establishments in the survey with certainty where such inclusion is considered highly important and the establishment is within scope for the survey area. The automated system allows definition of establishments individually by an identification number and also collectively by industry grouping using the SIC code. For example, on-base food contractors are included with certainty for NAF wage surveys.

**2. Definition of additional industries.** FPM Supplements 532-1 and 532-2 allow a lead agency to define the addition of other industry classes to a regular survey in a wage area where those industries account for significant proportions of local private employment of the kinds and levels found in local Federal employment. These added industries are defined in terms of the appropriate SIC code(s). For example, in the Orlando, Florida wage area, SIC 7996, amusement parks, is added to the universe.

**3. Determination of sample size.** The number of units to be in the sample is determined by the user or may be calculated as a function of the sample size of the last full scale survey. This number is the number of establishments to be selected in the sampling process and does not include the number of establishments that are to be included with certainty. Presently, the sample size is determined by the user and is based upon the size of the sample for the wage area in the last full scale survey. The automated system will compute the sample size if the user does not enter one. The necessary data file from the last full scale sample selection must exist and be input to the sampling process. In this case the sample size,  $n$ , will be calculated as:

$$n = \frac{1}{\frac{(M-m)}{Mm} + \frac{1}{N}}$$

$M$  = number of establishments in universe for last full scale  
excluding establishments defined as certainty.

$m$  = number of establishments in sample for last full scale  
excluding certainty establishments.

$N$  = number of establishments in universe currently excluding  
certainty establishments.

**4. Stratification.** The universe of establishments is determined by three factors: establishment physical location, its primary industry, and its number of employees. Each establishment in the scope of the universe for the wage area is put into a group based upon its primary industry and employment size. Establishments are first grouped by industry. Then, each industry grouping is divided into size groupings based upon its number of employees. Since the universe has been divided into groups called strata, we say it has been stratified. Establishments identified to be selected with certainty are grouped into separate strata called certainty strata. The certainty strata have the same size groupings as the sampled strata. See Appendix A for the definitions of the industry groupings and size groupings.

**5. Allocation of the sample over the strata.** The number of establishments to be selected from each noncertainty stratum is calculated by allocating the sample size over these noncertainty strata. The allocation is based upon the number of employees in the stratum. The allocation procedure is described in Appendix B. The number of establishments to be selected from a certainty stratum is by definition all the establishments in the stratum.

6. **Measure of size.** A measure of size is calculated for each establishment. It determines the probability of selection for the establishment. The measures are calculated independently for each stratum. The magnitude of the measure is related to status of the establishment in the previous full scale survey. Establishments that were selected in the previous full scale survey have higher measures of size, ones that were in the universe previously but not selected have small measures of size, and establishments not in the universe at all previously have measures of size somewhere in between. This is done to make the survey more efficient since there is less cost and time involved in contacting a previously interviewed establishment. [redacted] as first formulated in an article in the American Statistical Association Journal, March 1951. The actual formulae used are given in the internal BLS memorandum written by Phil D. Gilliland to Thomas N. Wakin, subject: Automated Sample Selection Procedure for the Federal Wage System. The formulae are shown below.

To calculate the measure of size for establishments in stratum  $r$  the following quantities are needed:

- $r$  = stratum that establishment is currently in.
- $p$  = stratum that establishment was in last survey, if any.
- 5.  $N_r$  = total number of establishments in current stratum  $r$ .
- 4.  $N_{rp}$  = number of establishments in current stratum  $r$  which were in stratum  $p$  in last survey.
- 5.  $n_p$  = number of establishments in sample last survey for stratum  $p$ .
- 3.  $n_{rp}$  = number of establishments in current stratum  $r$  which were in sample last survey for stratum  $p$ .

The measures of size for establishments in stratum  $r$  are:

- a. For establishments that were in the sample last survey in stratum  $p$ :

$$\frac{N_{rp} + n_p - n_{rp}}{n_p N_r} = \frac{6}{25} = .24$$

- b. For establishments that were in stratum  $p$  last survey but not in the sample:

$$\frac{n_p - n_{rp}}{n_p N_r} = \frac{5 - 3}{25} = \frac{2}{25} = .08$$

- c. For all other establishments, i.e., establishments not in the universe for the last survey:

$$\frac{1}{N_r} = \frac{1}{5} = .20$$

It is possible that the measure of size may be so large that an establishment will be selected twice. This can happen if the measure of size is greater than or equal to the sampling interval. See formula a above. The sampling interval is  $\frac{1}{n_r}$  where  $n_r$  is the number of allocated establishments for stratum  $r$ . If this is the case, the establishments in the stratum are reselected with equal probability by assigning all establishments a measure of size of  $\frac{1}{N_r}$  where  $N_r$  is the number of

establishments in the stratum. The sampling process is fully automated and uses data files with records of the establishments. Therefore, if the data file from the last full-scale survey for the wage area is not available, all establishments within the stratum are assigned the same measure of size using formula c above.

**7. Sample selection.** The selection of establishments within each stratum is done independently. Each stratum has  $N$  units from which  $n$  units are to be selected for the sample as determined by the allocation procedure above. The units in a stratum are numbered in a certain order, i.e. by sic code and number of employees. The first of the  $n$  units to be selected is selected at random and then every  $k$ th unit on the list after that is selected where  $k = \frac{N}{n}$ . Thus  $n$  units are selected. This method is called systematic sampling. The probability of a unit being selected is its measure of size as described above.

**8. Computation of the weight.** Each establishment selected has a weight assigned. The weight is a whole number and indicates the number of establishments in the same stratum that the selected establishment represents. For example, if there are three companies in the universe in a particular stratum and one is selected for the survey, it will be given a weight of three because it represents itself and two other companies. In a universe survey, of course, all companies have a weight of one since each represents itself. The sum of the weights for selected establishments in a stratum must equal the number of establishments in the stratum.

The computation of the weight is as follows:

- a. The quantity  $\frac{N}{n}$  is calculated:  $\frac{N}{n} = W\frac{r}{n}$  where  $N$  is the number of establishments in the stratum;  $n$  is the number of selected establishments in the stratum;  $W$  is the integer, whole number, part of the quotient;  $r$  is the remainder part of the quotient.
- b. Since the sum of the weights must equal the number of establishments in the stratum,  $r$  establishments are given a weight of  $W+1$  and  $n-r$  establishments are given a weight of  $W$ . The weights are assigned randomly among the selected establishments.

**9. Establishment Survey List.** The selected establishments are printed in a specific format showing the name of the company, address, employment size, industry (SIC), weight assigned, and other information. The certainty establishments appear first in ascending order by industry (SIC) and increasing employment within industry, then the sampled establishments are listed in the same order. This list is called the Establishment Survey List (ESL).

### III. WEIGHT ADJUSTMENTS FOR NONRESPONSE

#### A. General

After data collection is complete, the appropriate DODWFA Division enters the data into the survey database. This data includes the total employment and the collection status (DAC, OOB, OOA, WSC, REF ...) for each establishment on the ESL. Information Systems Division uses this data to adjust the weights. An annotated copy of the original ESL is also sent to Information Systems Division (ISD) to serve as backup. Weights are adjusted based on all of this information.

#### B. Noncertainty Units

The procedures for noncertainty units (establishments selected through a random sample and usually having a weight greater than one) are designed to ensure that weights for refusing establishments are randomly and evenly distributed among survey participants.

Nonresponding unique establishments may not be reweighted for since some aspect of the unique establishment, such as employment size or industry type, may preclude an assumption of comparability to another unit. For noncertainty establishments, the ISD staff will determine whether the number of responding units in a given stratum is equal to or greater than the number of refusals in that stratum.

-If yes, proceed to 1.

-If no, go to 7.

1. Add the weights of the refusing establishments and the responding establishments in the stratum. Divide by the number of respondents. The result is the Adjusted Sampling Ratio Reciprocal (ASRR).

$$ASRR = \frac{\text{Weights of Refusals and Respondents}}{\text{No. of Responding Establishments}}$$

Example:

Suppose the weights of the refusal and responding establishments totaled 34, and there were 8 responding establishments:

$$\frac{34}{8} = 4\frac{2}{8} \text{ (do not reduce fraction to lowest common denominator)}$$

2. If the ASRR is a whole number, assign it as the new weight of each responding establishment in the stratum.

If the ASRR is not a whole number, continue to 3.

3. The ASRR is expressed as a compound number (whole number plus fraction). Use the remainder (the numerator of the fraction) as the number of establishments that get the larger weight (the whole number plus one).

Example:

The ASRR is  $4\frac{2}{8}$ . Since the remainder is 2, 2 units will get the larger weight of 5 and the other 6 units will get the smaller weight of 4.

4. Assign the new weights randomly (using random numbers selected by the computer) among the responding establishments in the stratum.
5. Verify that the new weights are correct by adding them up. The sum should equal the total original weights of the refusing and responding establishments.

$$2 \times 5 + 6 \times 4 = 34 \text{ total weights}$$

6. Verify also that the new weights in the stratum are within one of each other. In this example, five is indeed within one of four.
7. If the number of refusing establishments is significantly greater than the number of responding establishments in a cell size grouping, the stratum is collapsed into either the next highest employment stratum or the next lowest to account for all employment in both strata. This procedure will allow the weights of the refusing establishments to be distributed as evenly as possible between responding establishments in both strata.

Example:

Cell Size	Weight	Establishments	
		Responding	Refusing
50 - 99     4	3	2	4
100 - 249   5	2	5	1

To collapse the strata, first compute the weighted total employment. This is done by multiplying each within scope unit (responding or refusing establishments) in each stratum by its weight and then by the

a.  $(2 \times 3 + 4 \times 3) \times \blacksquare = 1,350$       $(5 \times 2 + 1 \times 2) \times \blacksquare = 2,100$

$1,350 + 2,100 = 3,450$  (total weighted employment)

- b. The total weighted employment is then divided by the number of responding units, which in this case is 7.

$\frac{3,450}{7} = 492$  (Truncate the answer.)

- c. This figure is then divided by the average employment size in each cell to obtain the new weights. Round to the appropriate fractional ASRR:

$\frac{492}{75} \rightarrow 6\frac{1}{2}$ : therefore in the stratum with size grouping 50 - 99, the new weights would be 1 with weight of 7 and 1 with weight of 6.

$\frac{492}{175} \rightarrow 2\frac{4}{5}$ : therefore in the stratum with size grouping 100 - 249, the new weights would be 4 with a weight of 3 and 1 with a weight of 2.

### C. Certainty Units

In the case of a certainty establishment, the following procedure is used so that the comparable establishment receives a new weight that accounts for the total employment of both the comparable and the refusing establishment.

1. Sum the weighted assigned (expected) employments of the certainty establishment and the actual (observed) weighted employment of the comparable establishment.
2. Divide this sum by the unweighted actual employment of the comparable establishment.
3. Round the result to the nearest whole number. This is the revised weight of the comparable establishment.

Example:

Certainty Establishment: assigned employment 1,222

Comparable Establishment: actual employment 600, weight 2

Sum weighted employment

$$1,222 + (600 \times 2) = 2,422$$

Divide by comparable establishment's unweighted employment:  $\frac{2,422}{600} = 4.03$

Round to the nearest integer:  $4.03 \rightarrow 4$

Comparable establishment's revised weight is 4.

4. In some cases the comparable establishment's weight will not change. This may occur if the comparable establishment's unweighted employment is much larger than that of the certainty establishment.

### D. Unique Units

If the establishment is a unique establishment (i.e., there is something about the company such as employment size or industry that sets it apart from other companies), reweighting is generally not recommended.

### E. Combined Returns

A combined return (CMB) refers to a situation where data are collected for multiple units from one responding unit.

1. **Recognizable Units.** When a participating establishment represents a combined return, it is necessary to reweight for shifts or additions to the represented populations. In a situation where data are collected from a unit clearly identified on the ESL that includes one or more other clearly recognizable unit(s) appearing on the ESL, reweighting is accomplished by computing the total weighted employment for these units then dividing this number by the actual employment for which data were collected. The result is the new weight for the establishment from which data were collected.

**2. Hidden Units.** If a combined return includes data from an establishment on the ESL plus establishments that are not recognizable on the ESL (hidden units in weights):

1. Identify the strata that contain the hidden establishments and randomly reduce the weights of establishments on the ESL in those strata since the hidden establishments are accounted for in the reweighting of the combined participating establishment.
2. Compute the new weight for the combined participating establishment.
  - a) multiply its assigned weight by its assigned employment.
  - b) add the total assigned employment of the hidden establishments.
  - c) divide the sum by the total actual collected employment.

**Note:** A combined return for hidden establishments can be reweighted only if all establishments which the combined return represents are clearly within survey scope and their individual employments are known.

APPENDIX A

Appropriated Fund SIC Groupings

Industry Groupings		
<i>Industry</i>	<i>SIC Code</i>	<i>Grouping Number</i>
Manufacturing - nondurables	SIC 20-23, 26-31	1
Manufacturing - durables	SIC 24-25, 32-38	2
Railroad transportation	SIC 40	3
Local and suburban transit and interurban highway passenger transportation except taxicabs	SIC 41 (except 412)	4
Motor freight transportation and warehousing	SIC 42	5
Transportation by air	SIC 45	6
Communications	SIC 48	7
Electric, gas, and sanitary services	SIC 49	8
All wholesale trade	SIC 50-51	9

Size class groupings:

50-99; 100-249; 250-499; 500-999; 1,000-2,499; 2,500 and more.



Non-Appropriated Fund CT SIC Groupings

Industry Groupings for CT Establishments		
<i>Industry</i>	<i>SIC Code</i>	<i>Grouping Number</i>
All wholesale trade	SIC 50-51	1
Department stores	SIC 5311	2
Variety stores	SIC 5331	3
Automatic merchandizing machine operators	SIC 5962	4
Gas stations	SIC 5541	5
Restaurants	SIC 58	6
Hotels	SIC 7011	7
Bowling alleys	SIC 7933	8
Private golf and country club	SIC 7997	9

Size class groupings:

For all above SIC's except 5541, 5962, 7933, and 7997: 20-49; 50-99; 100-249; 250-499; 500-999; 1,000-2,499; 2,500 and more.

For SIC's 5541, 5962, 7933, and 7997: Same size class groupings as above plus 8-19 grouping.

## Non-Appropriated Fund AS/PS SIC Groupings

Industry Groupings for AS/PS Establishments		
<i>Industry</i>	<i>SIC Code</i>	<i>Grouping Number</i>
Clothing stores	SIC 5611, 5621, 5651	1
Drug stores	SIC 5912	2
Direct selling establishments	SIC 5963	3
Banks	SIC 6021, 6022, 6029	4
Life, fire insurance	SIC 6311, 6331	5
Sporting and recreational	SIC 7032	6
Child day care services	SIC 8351	7

## Size class groupings:

For all above SIC's except 5963, 7032, and 8351: 20-49; 50-99; 100-249; 250-499; 500-999; 1,000-2,499; 2,500 and more.

For SIC's 5963, 7032, and 8351: Same size class groupings as above plus 8-19 grouping

APPENDIX B

Allocation of Establishments to Strata

This explains the procedure to allocate the number of desired sample units over the strata that are to be sampled. By definition any stratum defined as a certainty stratum has all establishments in the stratum in the sample and is excluded from this procedure.

1. Calculate survey sample size,  $n$ , for noncertainty strata only. Do not include the count for certainty strata, if any.
2. Calculate the number of employees,  $E_r$ , in each noncertainty stratum  $r$  and the total noncertainty universe employment,  $E$ , over all noncertainty strata.
3. Calculate the number of establishments,  $N_r$ , in each noncertainty stratum  $r$ .
4. Calculate  $F_0$ :

$$F_0 = \frac{n(\text{total noncertainty sample size})}{E(\text{total noncertainty universe employment})}$$

5. Calculate  $n_r$ , allocated sample size for each noncertainty stratum  $r$ :

$$n_r = F_0 E_r (E_r = \text{employment total in stratum } r)$$

6. If  $n_r$  is greater than or  $= N_r - 2$  then set  $n_r = N_r$ . That is, all units in stratum  $r$  are in sample. The stratum is called a certainty sampled stratum. If there are no such strata, skip to step 11.
7. Calculate the total number of units,  $N_c$ , allocated to certainty sampled strata (step 6). Calculate the total employment,  $E_c$ , in the certainty sampled strata.

8. Calculate  $F_1$

$$F_1 = \frac{n(\text{from step 1}) - N_c(\text{from step 7})}{E(\text{from step 2}) - E_c(\text{from step 7})}$$

9. For all strata that are not certainty sample, recalculate  $n_r$ :

$$n_r = F_1 e_r (e_r = \text{employment total in stratum } r)$$

10. Repeat steps 6 through 9, if  $N_c$  is greater than zero (step 7).
11. For all strata where  $n_r$  is greater than zero and  $n_r$  is less than one, allocate one unit. That is, set  $n_r = 1$ .
12. For all strata where  $n_r$  is less than two and  $N_r$  is greater than one, allocate two units. That is, set  $n_r = 2$ .
13. Otherwise, round  $n_r$ , computed above (in step 5, 6, or 9 as appropriate) to nearest whole number.

14. For each stratum compute  $W_r = \frac{N_r}{n_r}$ . For any stratum where  $W_r$  is greater than the maximum weight, set  $n_r = \frac{N_r}{\text{maximum weight}}$ , dropping any decimal portion.