

Technical Paper 2011-1

SAMPLING METHODOLOGY
FOR
FEDERAL WAGE SYSTEM
WAGE SURVEYS

Civilian Personnel Management Service
Wage and Salary Division

1 March 2011

Supercedes Technical Paper 91-1

CONTENTS

I. INTRODUCTION	
A. Scope	1
B. Wage Area	1
C. Types of Surveys	2
D. Bureau of Labor Statistics Function	2
E. Purpose	2
II. SAMPLING METHODOLOGY	
A. Source of Universe	3
B. Universe Definition	3
1. Location	3
2. Industry	4
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	8
6. Measure of size	8
7. Sample selection	9
8. Computation of the weight	9
9. Establishment Survey List	10
III. WEIGHT ADJUSTMENTS FOR NON-RESPONSE	
A. General	11
B. Noncertainty Units	11
C. Certainty Units	13
D. Unique Units	13
E. Combined Returns	14
1. Recognizable Units	14
2. Hidden Units	14
APPENDIX A	
Appropriated Fund NAICS Groupings	
Non-Appropriated Fund CT NAICS Groupings	
Non-Appropriated Fund AS/PS NAICS Groupings	
APPENDIX B	
Allocation of Establishments to Strata	

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 unskilled labor type jobs, more commonly referred to as blue collar. 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 unskilled labor employees in the executive branch of the federal government who are paid from appropriated funds (ApF). 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 title 5 Code of Federal Regulations (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 on prevailing rates in a locality, OPM has defined geographical areas, referred to as wage areas, for which rates of pay are set. The wage areas are defined in terms of state, county (or parish), and town boundaries. There are 132 ApF wage areas and 121 NAF wage areas. The NAF wage areas are defined separately from the ApF 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. Prior to 1995, the Department of Veterans Affairs was the lead agency for 25 wage areas and the National Aeronautics and Space Administration was the lead agency for one wage area. In 1995 OPM designated the Department of Defense (DoD) as the lead agency for all ApF wage areas and for all NAF wage areas. Within DoD, the Civilian Personnel Management Service, Wage and Salary Division (WAGE) executes this authority with the advice of the DoD Wage Committee.

C. Types of Surveys

The lead agency conducts 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 Federal 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. ApF surveys are conducted in 132 areas. NAF surveys are conducted in 121 areas. In universe survey wage areas, all in-scope establishments are included in the survey. In sample survey 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 more than 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, point of contact, 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, 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 non-response. 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 non-response in both universe and sample surveys. In April 1989, OPM transferred these responsibilities from BLS to the Department of Defense Wage Fixing Authority Technical Staff (DODWFATS). In 1994 DODWFATS became the Civilian Personnel Management Service, Wage and Salary Division (WAGE).

E. Purpose

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

- 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, the information was obtained from state unemployment insurance agencies and 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 other 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 was identified by the Standard Industrial Classification (SIC) code. Beginning in 2005, the SIC code was replaced by the North American Industry Classification System (NAICS) code. In April 1989, the DODWFATS officially took over this BLS function. The computer programs used to produce the list of establishments for wage surveys were obtained from BLS. The establishment records are updated annually with new information from Dun and Bradstreet. WAGE maintains a database of in-scope establishments for all wage areas. The database is updated by 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 in the establishment.

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, parishes, or independent cities. In New England states, for ApF wage areas, definitions may consist 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 ApF survey area includes the city of Washington, D.C.; the independent cities of Alexandria, Fairfax, Falls Church, Manassas, and Manassas Park in Virginia; the counties of Charles, Frederick, Montgomery, and Prince George's in Maryland; and the counties of Arlington, Fairfax, Loudoun, and Prince William in Virginia. The area of application includes all the survey area plus the counties of Calvert and St. Mary's in Maryland; the city of Fredericksburg in Virginia; the counties of Clark, Facquier, King George, Spotsylvania, Stafford, and Warren in Virginia; and the county of Jefferson in West Virginia. NAF wage areas are usually smaller than ApF wage areas. For example, the NAF wage area for Washington, D.C. is defined as the survey area being the city of 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 ApF wage surveys for regular wage schedules as defined in title 5 CFR Part 532.

The industries are identified by 2007 NAICS codes:

311 through 339 (except 323)	All manufacturing classes except printing and related support activities (NAICS 323)
221	Utilities
481	Air transportation
482	Rail transportation
484	Truck transportation
485 (except 4853)	Transit and ground passenger transportation except taxi and limousine service (NAICS 4853)
487 (except 4872)	Scenic and sightseeing transportation except scenic and sightseeing transportation, water (NAICS 4872)
488 (except 4883 and 4884)	Support activities for transportation except support activities for water transportation (NAICS 4883) and support activities for road transportation (NAICS 4884)
492	Couriers and messengers
493	Warehousing and storage
515	Broadcasting (except Internet)
517	Telecommunications
5621	Waste collection
5622	Waste Treatment and Disposal
423	Merchant wholesalers, durable goods
424	Merchant wholesalers, 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, NAICS 212210, iron ore mining, is added to the scope of the wage survey in Northwestern Michigan and NAICS 622110, General Medical and Surgical 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 title 5 CFR Part 532. The industries are identified by 2007 NAICS codes:

42312	Motor Vehicle Supplies and New Parts Merchant Wholesalers
4232	Furniture and Home Furnishing Merchant Wholesalers
42362	Electrical and Electronic Appliance, Television, and Radio Set Merchant Wholesalers
42369	Other Electronic Parts and Equipment Merchant Wholesalers
42371	Hardware Merchant Wholesalers
42391	Sporting and Recreational Goods and Supplies Merchant Wholesalers
42399	Other Miscellaneous Durable Goods Merchant Wholesalers
4241	Paper and Paper Product Merchant Wholesalers
42421	Drugs and Druggists' Sundries Merchant Wholesalers
4243	Apparel, Piece Goods, and Notions Merchant Wholesalers
42445	Confectionary Merchant Wholesalers
4247	Petroleum and Petroleum Products Merchant Wholesalers
4249	Miscellaneous Nondurable Goods Merchant Wholesalers
44132	Tire Dealers
44311	Appliance, Television, and Other Electronics Stores
44411	Home Centers
44611	Pharmacies and Drug Stores
4471	Gasoline Stations
44814	Family Clothing Stores
4521	Department Stores
45299	All Other General Merchandise Stores
45321	Office Supplies and Stationery Stores
4542	Vending Machine Operators
71391	Golf Courses and Country Clubs
71395	Bowling Centers
72111	Hotels (except Casino Hotels) and Motels
7221	Full-service Restaurants
7222	Limited-service Eating Places
7224	Drinking Places (Alcoholic Beverages)

In addition to the above industries defined in title 5 CFR Part 532, on-base food contractors (NAICS 561320) are included in NAF wage surveys.

The following industries are included in all NAF surveys for the AS and PS schedules as defined by WAGE. The industries are identified by 2007 NAICS codes:

Finance and Insurance

522110	Commercial Banking
522120	Savings Institutions
522130	Credit Unions
522210	Credit Card Issuing
524113	Direct Life Insurance Carriers
524114	Direct Health and Medical Insurance Carriers
524126	Direct Property and Casualty Insurance Carriers

3. Number of Employees: The minimum number of employees an establishment must have to be in the universe for an ApF 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 NAICS codes 441320, 447110, 447190, 454210, 713910, and 713950, for which the minimum number is eight. For the NAF AS/PS schedules, the minimum number of employees is also 20.

C. Sampling Process

The sampling design used in the wage surveys is sampling proportionate to establishment employment. The universe is divided into groups based on the industry in which the establishment is engaged. These groups are in turn divided into subgroups based on the establishment employment. These subgroups are called strata. Within each stratum, a sample of establishments is selected independently from the other 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 on 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, such as General Medical and Surgical Hospitals, NAICS 622110
- Determination of sample size
- Stratification of the universe by industry and employment size
- Allocation of the sample over the strata based on 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 of the selected establishments

1. Definition of certainty units: title 5 CFR Part 532 permits 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 NAICS code. For example, on-base food service contractors are included with certainty for NAF wage surveys.

2. Definition of additional industries: title 5 CFR Part 532 allows 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 the local private employment of the kinds and levels found in local Federal employment. These added industries are defined in terms of the appropriate NAICS code(s). For example, in the Orange County (Orlando), Florida wage area, NAICS code 713110, Amusement and Theme 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 on 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: (1) the physical location of the establishment; (2) the primary industry; and (3) the number of employees. Each establishment in the scope of the universe for the wage area is put into a group based on its primary industry and employment size. Establishments are first grouped by industry. Then each industry grouping is divided into size groupings based on the number of employees. These groupings are referred to as strata and the universe is considered 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 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 on 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 the 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. The measures of size are computed using formulae derived from probability theories of statistician Nathan Keyfitz as first formulated in an article in the *American Statistical Association Journal*, March 1951. The formulae were presented 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

N_r = total number of establishments in current stratum r

N_{rp} = number of establishments in current stratum r which were in stratum p in last survey

n_p = number of establishments in sample last survey for stratum p

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 :

$$\frac{N_{rp} + n_p - n_{rp}}{n_p N_r}$$

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

$$\frac{n_p - n_{rp}}{n_p N_r}$$

c. For all other establishments, i.e., establishments in the universe for the last survey:

$$\frac{1}{N_r}$$

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 NAICS 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 3 because it represents itself and two other companies. All companies in a universe survey have a weight of one since each represents only 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 (NAICS), weight assigned, and other information. The certainty establishments are listed first in ascending order by industry (NAICS) and then in increasing employment within industry; then the sampled establishments are listed next and in the same order. This list is called the Establishment Survey List (ESL).

III. WEIGHT ADJUSTMENTS FOR NON-RESPONSE

A. General

After data collection is complete, the appropriate WAGE regional office enters the data into the survey database. This data includes the total employment and the collection status (DAC, REF, OOB, etc.) for each establishment on the ESL. This data is then used to adjust the weights.

B. Noncertainty Units

The procedures for noncertainty units (establishments selected by 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 certainty establishments (found in universe surveys or specifically requested to appear on the sample survey list with a weight of one) are accounted for by increasing the weight of a comparable responding establishment to account for the total employment. 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, an analyst 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 the number of responding units is equal to or greater than the number of refusals, proceed to step 1. If not, proceed to step 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)*.

$$ASRR = \frac{\text{Weights of Refusals and Respondents}}{\text{Number of Responding Establishments}}$$

Example: The weights of the refusal and responding establishments totals 34 and there are 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 step 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, two units will get the larger weight of 5 and the other six 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 integer of each other, i.e. weights of 5 and 4 rather than weights of 6 and 3.

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	3	2	4
100 – 249	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 average employment size:

$$\text{a. } (2 \times 3 + 4 \times 3) \times 75 = 1,350 \qquad (5 \times 2 + 1 \times 2) \times 175 = 2,100$$

$$1,350 + 2,100 = 3,450 \text{ (total weighted employment)}$$

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

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

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*:

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

$\frac{492}{175} \rightarrow 2\frac{4}{5}$: therefore in the stratum with size grouping 100 – 249, the new weights would be four units with a weight of 3 and one unit 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 of 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: 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 weight):

a. 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.

b. 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 employment is known.

APPENDIX A

Appropriated Fund NAICS Groupings

Industry Groupings		
Industry	NAICS Code	Grouping Number
Manufacturing		
Nondurables	311 - 316, 322, 324 – 326	1
Durables	321, 327, 331 - 339	2
Transportation & Warehousing		
Air	481	6
Rail	482	3
Truck	484	5
All other transportation, except taxi & water	485 (except 4853) 487 (except 4872)	4
Support activities except water & road	488 (except 4883 & 4884)	
Air	4881	6
Rail	4882	3
Truck	4885	5
Couriers & messengers	492	4
Warehousing & storage	493	5
Communications	515 (except Internet), 517	7
Utilities	221	8
Sanitation services		
Waste Collection	5621	10
Waste treatment & disposal	5622	10
Wholesale trade		
Durables	423	9
Nondurables	424	9

Size class groupings: 50 – 99; 100 – 249; 250 – 499; 500 – 999; 1,000 – 2,499; 2,500 and more.

Non-Appropriated Fund CT NAICS Groupings

Industry Groupings for CT Establishments		
Industry	NAICS Code	Grouping Number
All Wholesale Trade	423 – 424	1
Gas Stations and Tire Dealers	441320, 447110, 447190	2
Electronic Stores	443111, 443112, 453210	3
Home Centers	444110	4
Drug and Variety Stores	446110, 452990	5
Clothing and Department Stores	448140, 452111, 452112	6
Vending Machine Operators	454210	7
Golf Courses	713910	8
Bowling Centers	713950	9
Hotels	721110	10
Eating and Drinking Places	722110, 722211, 722212, 722213, 722410	11

Size class groupings:

For all above NAICS except 441320, 447110, 447190, 454210, 713910, and 713950: 20 – 49; 50 – 99; 100 – 249; 250 – 499; 500 – 999; 1,000 – 2,499; 2,500 and more.

For NAICS 441320, 447110, 447190, 454210, 713910, and 713950: Same size class groupings as above plus 8 – 19 grouping.

Non-Appropriated Fund AS/PS NAICS Groupings

Industry Groupings for AS/PS Establishments		
Industry	NAICS Code	Grouping Number
Banking and Credit	522110, 522120, 522130, 522210	12
Insurance	524113, 524126, 524114	13

Size class groupings:

For all above NAICS: 20 49; 50 – 99; 100 – 249; 250 – 499; 500 – 999; 1,000 – 2,499; 2,500 and more.

APPENDIX B

Allocation of Establishments to Strata

The procedure used to allocate the number of desired sample units over the strata that are to be sampled is described below. 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 :

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

6. If n_r is greater than or equal to $N_r - 2$, then set $n_r = N_r$, i.e., all units in stratum r are in the 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 :

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

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, i.e., 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, i.e., 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.