



Information and Technology for Better Decision Making

July 2002 Status of Forces Survey of Active Duty Members

Statistical Methodology Report



Additional copies of this report may be obtained from:

Defense Technical Information Center

ATTN: DTIC-BRR

8725 John J. Kingman Rd., Suite #0944

Ft. Belvoir, VA 22060-6218

Or from:

<http://www.dtic.mil/dtic/order.html>

Ask for report by ADA 566 077

**JULY 2002 STATUS OF FORCES SURVEY OF
ACTIVE DUTY MEMBERS:
STATISTICAL METHODOLOGY REPORT**

**Defense Manpower Data Center
Human Resources Strategic Assessment Program
1600 Wilson Boulevard, Suite 400, Arlington, VA 22209-2593**

Acknowledgments

Master file data processing for the *July 2002 Status of Forces Survey of Active Duty Members* was performed at the Defense Manpower Data Center by J. Robert Hamilton, Carole Massey, and Susan Reinhold.

JULY 2002 STATUS OF FORCES SURVEY OF ACTIVE DUTY MEMBERS: STATISTICAL METHODOLOGY REPORT

Executive Summary

This report describes the sampling design, sample selection, estimation procedures, and the weighting procedures used for the *July 2002 Status of Forces Survey (SOFS) of Active Duty Members (July 2002 SOFS-A)*. The first section of this report presents a general overview of the survey and the sampling design. Subsequent sections provide information on the statistical methods used in weighting and variance estimation. Several types of response rates were calculated and are described in the last section of the report.

Target Population

To be eligible for the *July 2002 SOFS-A* respondents must have met all four the following criteria:

1. Active duty members of the Army, Navy, Marine Corps, and Air Force.
2. At least six months service at the time the Web questionnaire was first fielded.
3. Up to and including paygrade O6 personnel when the sample was drawn.
4. Not a National Guard or Reserve member in an active duty program.

Sample Design

Single-stage, nonproportional stratified random sampling procedures were employed to ensure adequate sample sizes for the reporting categories. In stratified random sampling, all members of a population are categorized into mutually exclusive homogeneous groups. Sample members are then chosen using simple random sampling techniques (SRS) within each group. The advantage of using stratified random sampling is that small population subgroups can be oversampled in comparison to their proportion in the population, which ensures that there will be a sufficient number of completed surveys within these subgroups for analysis. The stratifying variables used in the *July 2002 SOFS-A* to form these mutually exclusive groups were:

1. Service Branch (Army, Navy, Marine Corps, Air Force)
2. Paygrade Group (E1-E3, E4, E5-E6, E7-E9, W1-W5, O1-O3, O4-O6)
3. Sex (Male, Female)
4. Minority-status (Minority, Non-minority)
5. Duty Location (US, Asia/Pacific Islands, Europe)

6. Family-status (Single w/ Children versus Single w/o Children, Married w/ Children or Married w/o Children)
7. Dual Service Spouse (Yes/No)

Weighting

Weighting of the survey involved several stages that took into account the sample design and the response rates that were achieved in the survey. These steps were:

- Calculation of base weights
- Adjustments for unknown eligibility
- Adjustments for nonresponse among eligible sample persons
- Poststratification to counts of persons at the beginning of the data collection period.

The survey was a stratified simple random sample of persons and the weights were generated to ensure that estimates from the survey represent the target population. The first step in weighting was to compute a base weight, which was the inverse of the selection probability for each initially sampled person. Since the eligibility of some persons could not be determined due to nonresponse, an adjustment was made to apportion the weights of the unknowns among the other persons in the sample. The third step above adjusted the weights of eligible respondents to account for the eligibles who did not respond. The final step in weighting was to poststratify weights to frame counts made for the beginning of the data collection period (July 2002). The poststratification step compensates for some changes in the population that occur between the time of sample selection and data collection.

Response rates for the *July 2002 SOFS-A* were computed in accordance with the standards defined by the Council of American Survey Research Organizations (CASRO). The response rates for the full sample and for subgroups and how they were computed are described in the last section of this report.

Table of Contents

	<u>Page</u>
Introduction.....	1
Sampling Design.....	2
Overview of the Sampling Design.....	2
Inferential Requirements.....	3
Population Definition.....	3
Key Reporting Domains.....	4
Precision Requirements.....	7
Sampling Frame Construction and Stratification.....	7
Sample Size and Allocation.....	9
Weighting.....	10
Assigning Eligibility and Disposition Codes.....	10
Frame Eligibility.....	11
Survey Control Disposition Codes.....	12
Self-Reported Eligibility.....	13
Completed Questionnaire.....	14
Weighting Procedures.....	15
Calculation of Base Weights.....	15
Weighting Adjustments.....	15
Unit Nonresponse Adjustments.....	16
Poststratification Adjustment.....	16
Computation of Variance for Estimates.....	19
Software for Computing Taylor Series Variances.....	19
Calculation of Response Rates.....	21
Exhibit 1:.....	29
Executive Summary of Weighting Procedures.....	29
Weighting Variables.....	33
References.....	37

Appendixes

A. Sample Allocation.....	39
---------------------------	----

Table of Contents (Continued)

Page

List of Tables

1.	Factors Defining Key Reporting Domains	5
2.	Creation of the Frame Eligibility Variable (ELIGFLGW)	12
3.	Description of the Survey Control System Disposition Codes (FLAG_FIN)	13
4.	Self-Reported Eligibility (SSRINEL)	14
5.	Complete Questionnaires (COMPFLAG).....	14
6.	Definitions of 52 Poststrata.....	18
7.	Features of Three Software Packages for the Analysis of Survey Data	21
8.	Disposition Codes for CASRO Response Rates (CAS_ELIG)	23
9.	Observed Sample Disposition Counts, Response Rates, and Design Effects.....	26
ES 1.	Sample Disposition Codes (SAMP_DC).....	29
ES 2.	Eligibility Status Response Flag (ESTATRSP).....	29
ES 3.	Complete Eligible Response Flag (COMPERSP)	30
ES 4.	Postratification Flag (POSTSTRT).....	31
ES 5.	Eligibility Indicator (ELIGFLGW).....	31
ES 6.	Cases Assigned Weights in Each Step of the Weighting Process by Type of Disposition	32

JULY 2002 STATUS OF FORCES SURVEY OF ACTIVE DUTY MEMBERS: STATISTICAL METHODOLOGY REPORT

Introduction

The *July 2002 Status of Forces Survey of Active Duty Members (July 2002 SOFS-A)* is the first active duty survey under the Defense Manpower Data Center's (DMDC) new Web-based *Status of Forces Survey (SOFS)* program. When fully operational in 2004, SOFS will include nine short surveys a year, with three cross-sectional samples of each population: active duty members, Reserve component members, and DoD civilian employees.

The *July 2002 SOFS-A* continues a long series of Department of Defense (DoD) research initiatives to provide policymakers with information about the military life of active duty personnel. The first small-scale survey was conducted in 1969 and was administered every two years until 1978, when it was expanded to become representative of the entire population of directly involved with active duty military life (Doering, Grissmer, Hawes and Hutzler, 1981). The DoD also conducted large-scale active duty surveys in 1985 (Hunt et al., 1986), 1992 (Westat, 1993, 1994a, 1994b) and 1999 (Helba, Lee, Perry, Wright and Williams, 2001).

The *July 2002 SOFS-A* is sponsored by the Office of the Assistant Secretary of Defense for Force Management Policy (OASD[FMP]) with particular interest in analysis by the Offices of the Deputy Assistant Secretaries of Defense for Military Community and Family Policy (ODASD[MCFP]) and for Military Personnel Policy (ODASD[MPP]). The *July 2002 SOFS-A* was administered to respondents using one of two methods: a Web-based method and a paper-and-pencil method. This report documents the sample design, weighting, and variance estimation for the Web-based method only.

In formulating policy, the DoD relies on both administrative data and survey data. The administrative data contain personnel-related information collected from individuals, or maintained about them. These data are largely automated and readily available for policy research and formulation purposes (e.g., to determine amounts of military compensation, eligibility for various forms of health and program benefits, and performance assessments) (LaVange et al., 1986).

Survey data can be used to supplement administrative data, as well as to address issues that cannot be studied from the administrative data. Especially when collected periodically, these data can serve as a basis for assessing the response of military personnel to policy changes and for identifying areas for future policy action.

DMDC has performed military personnel surveys of active duty personnel approximately every seven years since 1978. In 1985, it began fielding a spouse questionnaire in addition to the member form. These earlier surveys allowed policy makers to view trends in high-interest areas. Information from previous surveys illustrates the wide variety of uses found for active duty survey data. For example, previous surveys have been used to study: the effects of Operation Desert Shield/Desert Storm on the family, how attitudes on the military way of life change over

time, the effect of separation and deployment on the family, and how military couples deal with military life. Information from the earlier surveys was used in congressional reports (on topics such as military members qualifying for food stamps) and data have been used extensively by the Quadrennial Reviews of Military Compensation.

Sampling Design

This section of the report describes:

- the inferential requirements for the survey including the population definition, key reporting domains or subpopulations defined within the overall population, and the precision requirements imposed on sample estimates of parameters describing the key domains;
- the construction and stratification of the sampling frame;
- the procedure followed to determine the sample size and allocation; and
- selection of the sample.

A distinction is made between *sample size* and *number of observations*. Sample size refers to the number of persons selected into the sample. Sample sizes are determined to provide a specified number of observations given the anticipated eligibility and response rates for the survey. The sample is the group of persons to whom a questionnaire is to be administered. Number of observations, on the other hand, refers to the number of persons eligible to participate in the survey who returned a questionnaire with key items completed.

A distinction is also made between *strata and domains*. Stratification is a feature of the sampling design, used to control the distribution of the sample. Strata partition the inferential population in the mathematical sense. That is, each individual in the population is classified into only one stratum, and the set of all strata includes the entire population. By contrast, a single individual can simultaneously belong to one or more domains. The set of domains, as a consequence, does not partition the population and is itself arbitrary, depending largely on the interests of the investigators analyzing the data. *Key domains* are identified in advance of the survey to provide the basis for determining the sample size and allocation.

Overview of the Sampling Design

A stratified random sampling design was used. Source information for constructing the sampling frame and identifying key domains consisted of a computer accessible file totaling 1,357,526 records. The sample frame was constructed from the DMDC's December 2001 Active Duty Master Edit File.

Within each stratum, persons were sampled with equal conditional probabilities, and without replacement. Stratum level sample sizes were determined by variance constraints imposed on key parameter estimates of the proportion of persons belonging to specified reporting domains. The sample drawn from the sampling frame initially consisted of 37,918

individuals. Some sample members (n = 1,499) were then flagged as ineligible if they were ineligible for benefits in the March 2002 Defense Enrollment Eligibility Reporting System (DEERS) Medical Point-in-Time Extract (PITE). Still other sample members (n = 117) also became ineligible if they indicated in the survey, by other means (such as telephone calls or e-mails to the data collection contractor), or through a proxy (such as a spouse) that they were not in active duty service as of the first day of the Web-based survey, July 8, 2002.

Inferential Requirements

The inferential requirements for a survey are described in terms of

- a fully operational definition of the population of inferential interest (i.e., the target population),
- key parameters used in developing the design, and
- the precision requirements for the survey, stated in terms of the maximum values of the variances to be associated with the sample estimates of the key parameters.

The population definition identifies all individuals for whom conclusions are to be reached or about whom inferences are to be made based on the survey data. The definition generally includes a spatial and a temporal component.

Key parameters used as the basis for the design may be defined in terms of characteristics of the overall population, characteristics of subpopulations of special interest (key domains), tests of hypotheses (including standardized comparisons), and the relations that exist at population levels among specified observation variables. For this survey, the key parameters were prevalence rates, defined as the proportion of persons belonging to specified domains who would report having the various attitudes and experiences measured on the survey.

The precision requirements were defined in terms of the maximum *confidence interval half-widths* to be associated with a priori estimates of 50% prevalence rates. For this survey, the maximum values of the variances to be associated with the sample estimates of the prevalence rates were specified in the form of confidence interval half-widths. Both the cost implications and the objectives of the survey were considered in specifying these values. On the one hand, the intervals had to be small enough to provide an informative study. On the other hand, they could not be so restrictive as to be unaffordable. Table A-1 lists the half-width confidence intervals together with the domain definitions, domain sizes, and prevalence rates.

Population Definition

The population of interest for the *July 2002 SOFS-A* consisted of all active duty personnel in the Army, Navy, Marine Corps and Air Force up to and including paygrade O6, with at least six months of service at the time of the survey mailings, and not a National Guard or Reserve member in an active duty program. The sample frame included only members who were active duty on July 8, 2002. The sample for the July 2002 SOFS consisted of 37,918 individuals, of whom 36,419 were determined to be eligible members of the target population,

with eligibility conditional on them being eligible for benefits in the March 2002 Defense Enrollment Eligibility Reporting System (DEERS) Medical Point-in-Time Extract (PITE).

Key Reporting Domains

The factors used to define the key reporting domains are listed in Table 1. An initial set of candidate domains was generated by considering various combinations of, and crosses among, the factors listed in the table. Because the domain sizes interact with the precision requirements imposed on the domain prevalence estimates to determine the overall sample size and allocation, several iterations were required to develop domain definitions consistent with the objectives of the survey and the resources available to carry out the survey.

Table 1.
Factors Defining Key Reporting Domains

Variable	Categories
Service branch*	<ul style="list-style-type: none"> • Army • Navy • Marine Corps • Air Force
Paygrade Group 1*	<ul style="list-style-type: none"> • E1-E3 • E4 • E5-E6 • E7-E9 • W1-W5 • O1-O3 • O4-O6
Sex*	<ul style="list-style-type: none"> • Male • Female
Race/Ethnic Category 2*	<ul style="list-style-type: none"> • Non-minority (Non-Hispanic White) • Minority
Family-status 1*	<ul style="list-style-type: none"> • Unmarried with children • Without children or married with children
Duty Location*	<ul style="list-style-type: none"> • U.S. • Asia/Pacific Islands • Europe
Dual Service Spouse*	<ul style="list-style-type: none"> • Unmarried or married to a civilian • Married to a military spouse
Race Code	<ul style="list-style-type: none"> • White • Asian or Pacific Islander • Black • American Indian or Alaskan native • Other
Race/Ethnic Code	<ul style="list-style-type: none"> • American Indian or Alaskan Native • Asian or Pacific Islander • Black (not Hispanic) • White (not Hispanic) • Hispanic • Other

*Stratification variables

Table 1. (continued)

Variable	Categories
Regions 1	<ul style="list-style-type: none"> • Unknown • US & US territories • Europe • Other • Asia & Pacific Islands
Regions 2	<ul style="list-style-type: none"> • US & US territories, Other, Unknown • Europe • Asia & Pacific Islands
Regions 3	<ul style="list-style-type: none"> • US & US territories, Other, Unknown • Europe • Asia & Pacific Islands • All Regions
CONUS	<ul style="list-style-type: none"> • CONUS (all 48 contiguous states and D.C.) • OCONUS (non contiguous states, territories and countries)
CINCs	<ul style="list-style-type: none"> • America • Europe • Pacific • Central • South
6 Ranges of Months Away for Duty Occupations	<ul style="list-style-type: none"> • .321 – 1.06 Months • 1.07 – 1.82 Months • 1.83 – 2.58 Months • 2.59 – 3.34 Months • 3.35 – 4.10 Months • 4.11 – 4.86 Months
3 Ranges of Months Away for Duty Occupations	<ul style="list-style-type: none"> • .321 – 1.82 Months • 1.83 – 3.34 Months • 3.35 – 4.86 Months
2 Ranges of Months Away for Duty Occupations	<ul style="list-style-type: none"> • .321 – 2.58 Months • 2.59 – 4.86 Months
Paygrade Group 2	<ul style="list-style-type: none"> • E1-E9 • W1-W5 • O1-O6
Paygrade Group 3	<ul style="list-style-type: none"> • E1-E3 • E4-E5 • E6-E9 • W1-W5 • O1-O3 • O4-O6
Paygrade Group 4	<ul style="list-style-type: none"> • E1-E3 • E4 • E5-E6 • E7-E9 • All Officers

Table 1. (continued)

Variable	Categories
Paygrade Group 5	<ul style="list-style-type: none">• E1-E4• E5-E9• W1-W5• O1-O3• O4-O6
Paygrade Group 6	<ul style="list-style-type: none">• Enlisted• Officer
Basic Allowance for Quarters (BAQ)	<ul style="list-style-type: none">• BAQ with Dependents• BAQ without Dependents• Partial BAQ• BAQ with Dependents, inadequate quarters• NO BAQ, adequate quarters
Off Base/Basic Allowance for Housing (BAH) status	<ul style="list-style-type: none">• Off Base/Receiving BAH• On Base/No BAH
Marital Status	<ul style="list-style-type: none">• Married• Not Married
Family Status 2	<ul style="list-style-type: none">• Single with Children• Other

Precision Requirements

Domains and their associated precision constraints were defined to allow in-depth analysis for the overall active duty personnel population and some depth of analysis for other smaller domains. More specifically, the survey precision requirements were set for domains that would facilitate analyses. Special attention was given to allow for Service branch and Paygrade group-level analyses.

In general, precision requirements are specified as the maximum values of the sampling variances to be associated with parameters estimates for key domains. Both the values of the parameters and the values of the variances are needed to complete the specification. The sampling variances are functions of the sample size, the distribution of the sample, population variances, and design constants. The parameter values used for the design are the prevalences¹ and are available upon request.

Sampling Frame Construction and Stratification

A distinction is made between *dimensions of stratification* and *levels of stratification*. The dimensions are the variables used to stratify the sample/population whereas the levels are the values present within a dimension.

¹ As defined above, prevalence rates are the proportion of persons belonging to specified domains who would report having the various attitudes and experiences measured on the survey.

Variables were selected to define the strata because they could be used to define the domains of most analytic issues. A sample can be optimally designed for domains that can be exactly defined as sets of one or more strata. As discussed below there are limits on how small strata can be, so only a limited set of strata can be defined. The variables used to define strata for the member sample were:

- Service of the member: Army, Navy, Marine Corps, and Air Force.
- Paygrade of the member: enlisted E1-E3, E4, E5-E6, E7-E9, warrant officers W1-W5, and commissioned officers O1-O3, and O4-O6.
- Sex: male and female.
- Duty Location: inside the continental U.S., Asia/Pacific Island, Europe.
- Race/Ethnicity: minority versus non-minority (non-Hispanic White).
- Family-status: unmarried with children versus unmarried without children, married with children, or married without children.
- Dual Service Spouse: Joint Service married (i.e., both the member and spouse were in the military) versus married non-joint (i.e., the member was married to a non-military spouse or unmarried).
- Unknown stratum: all individuals for whom one or more variables of the above stratum variables were missing.

As a starting point, a candidate set of strata was constructed by crossing all of the levels of the stratification variables and one for unknowns². If unbiased variances for linear statistics are to be a design requirement, then a minimum of two observations is needed in any stratum. However, if a stratum is too small, then insisting on at least two observations from that stratum introduces an unequal weighting effect that acts to increase variances for no reason other than the stratum is simply too small. Even if only a few strata are too small, the cumulative unequal weighting effects can compromise any variance advantage associated with having stratified in the first place.

This consideration led to defining “too small” in terms of a proportional allocation of the total sample. A proportional allocation of the sample cannot, by definition, introduce unequal weighting effects. Given a proportional allocation and a minimum requirement of two observations per stratum, the minimum stratum size was computed as,

$$\min\{N_h\} = \frac{2N}{n},$$

² Note that certain combinations do not exist because there are no warrant officers in the Air Force.

where,

N_h = the size of the h - th stratum,

N = the size of the population, and,

n = the total size of the sample.

For example, for $N = 1,325,037$ and $n = 60,000$, a minimum stratum size of $\min\{N_h\} = 44$ (rounded up to 45) would be indicated.

The decisions about which strata to collapse were based on identifying the candidate stratification dimensions with consistent patterns of deficient strata and on a consideration of the relative importance of specific candidate stratification dimensions to the surveys. Some specific levels that were collapsed were:

- Family-status and Dual Service Spouse-status categories were collapsed within the duty location levels: *Europe* and *Asia/Pacific Island*.
- Duty Locations – *Europe* and *Asia/Pacific Island* – were generally collapsed into one category (*Overseas*) across Females.
- O1-O3s were collapsed into O4-O6s across females and across Air Force minority males.
- Warrant officers were collapsed into O1 – O3s across males and into O1 – O6s across females.
- E5-E6s were collapsed into E7-E9s across Navy, Marine Corps and Air Force females and across Army non-minority females who are overseas.
- E1-E3s were collapsed into E4s across Navy non-minority females who are overseas.

The final strata definitions are listed in A, Table A-1. A total of 201 strata were constructed. The “unknown” stratum (stratum 201) contains persons for whom one or more of the stratum dimensions was missing from the source information.

Sample Size and Allocation

After the strata and domains were constructed, the total sample size and its allocation to the sampling strata were determined. The DMDC Sampling Tool (Kavee & Mason, 1997) was used to allocate the sample so that the precision requirements were met for the different reporting domains. This software is designed to produce optimal sample designs for stratified, equal probability samples for a specified cost model. The cost model used is the same as described by Wheelless, Mason, and Kavee (1997). Response rates were estimated from data collected in the 1992 Reserve components survey (Perry, Hintze, Mackin, & Weltin, 1997). Within each

stratum, units on the frame were sorted in a random order and the first n_h were selected for the sample where n_h was the sample size allocated to the stratum.

In determining the solution for the domains from the final sample design, the Lagrange Ratios identified those variance constraints that were driving the solution, and thus the size (cost) of the survey. Ratios closest to 100 have the greatest impact; the smaller the ratio the smaller the impact on the final design. Precision constraints with no impact on the solution have a zero ratio—these are domains for which the expected precision will meet the precision requirement as a result of meeting other more difficult to achieve constraints. The precision expected to be achieved from the design if the expected response rates were achieved can be made available upon request³.

The design effects are the ratios of the true variance expected from the design compared to the variances that would be achieved by a simple random sample. An overall design effect of 2.2 resulted from the final sample plan. This indicates that if the only estimates that had to be made were for the overall population, those estimates would have as small an error variance as a simple random sample 45%⁴ the size of the final sample.

Weighting

This section describes the weighting and estimation procedures for the *July 2002 SOFS-A*. The first step in weighting was to compute a base weight, which is the inverse of the selection probability for each initially sampled person. Since the eligibility of some persons cannot be determined due to nonresponse, the second step was an adjustment made to apportion the weights of the members with unknown eligibility among the known eligible and ineligible in the sample. The third step adjusted the weights of eligible respondents to account for the eligible members who did not complete the questionnaire. This final step in weighting was to poststratify weights to frame counts made at the beginning of the data collection period. The final step, poststratification, compensates for changes in the population that occur between the time of sample selection and data collection.

Response rates for the *July 2002 SOFS-A* have also been computed in accordance with the standards defined by the Council of American Survey Research Organizations (CASRO, 1982). The response rates for the full sample and for subgroups and how they are computed are described in the last section of this report.

Assigning Eligibility and Disposition Codes

Each person in the *July 2002 SOFS-A* survey was assigned a sample disposition code indicating whether the person was an eligible respondent (*ER*), an eligible nonrespondent (*ENR*), an ineligible (*IN*), or a person whose eligibility status was unknown (*UNK*). These codes were a

³ Precisions can only be given in expectation for those domains that do not exactly match a set of strata. It is only by chance that persons meeting the domain definition would be randomly selected in a strata not fully comprised of persons in that domain.

⁴ Forty-five percent is the inverse of 2.2, the overall design effect.

key input in weighting and in computation of response rates, discussed in later sections. In this section we describe in detail the codes that were assigned.

The method used to assign disposition was a sequential process that uses the values of the variables that determine the final disposition code. The variables are:

- ELIGFLGW—Defense Enrollment Eligibility Reporting System (DEERS) Medical Point-in-Time Extract (PITE) eligibility as of March 2002
- COMPFLAG—Completed questionnaire indicator
- FLAG_FIN—Survey Control System Disposition code
- SSRINEL—Self-reported eligibility from question 1 (SRSVC1)

In general, for each sampled member, we first determined if the member's eligibility was known or unknown. For members whose eligibility status was known, they were classified as either eligibles or ineligibles. For eligible members we then determined whether the questionnaire was complete or incomplete.

Frame Eligibility

In order to create the variable for frame eligibility an updated frame file, created from the March 2002 DEERS/PITE, was obtained from DMDC as close to the beginning of the field period as possible. This frame was constructed in the same way as the December 2001 frame from which the sample was selected. To be eligible for the survey, an active duty member had to have been eligible based on both the December 2001 and the updated March 2002 frames. A member who was eligible in December 2001 might have become ineligible by the beginning of the field period for any of several reasons. For example, the member may have no longer been in the active duty service as a result of a separation, retirement, death, or permanent illness. Any sample person not on the updated March 2002 frame was considered ineligible.

The December 2001 sample was matched against the updated frame file using social security number to determine eligibility. The frame eligibility was based on three criteria:

- The member had to be in both the December 2001 frame and the updated March 2002 frame.
- The member had more than five months of service at the end of March 2002.⁵
- The members who were in paygrades O6 or below in the December 2001 frame.⁶

⁵ This criterion may not have been met for all sample members. Sample members with TAFMS=0 on the March frame were deemed eligible despite its meaning either “unknown” or “less than one month of service.”

⁶ Included were members with the O6 paygrade on the December 2001 frame who had the O7 paygrade on the updated March 2002 frame.

The variable ELIGFLGW was created by merging the December 2001 and updated March 2002 frames. Table 2 shows the frame eligibility counts for active duty members based on matching the December 2001 and the updated March 2002 frames.

Table 2.
Creation of the Frame Eligibility Variable (ELIGFLGW)

ELIGFLGW	In Dec. 01 Frame?	In Mar. 02 Frame?	Total Members	Percentage
1 - Eligible	Yes	Yes	36,419	96%
2 -Ineligible	Yes	No	1,499	4%
Total			37,918	100.0%

Survey Control Disposition Codes

The Survey Control System created a variable with the disposition code (FLAG_FIN) of each mailed survey as determined during data collection. During data collection, returns received codes based on whether they were determined to be eligible respondents, eligible nonrespondents, ineligible, or with unknown eligibility. Table 3 shows the counts and descriptions of values of FLAG_FIN found in the sample.

Table 3.**Description of the Survey Control System Disposition Codes (FLAG_FIN)**

FLAG_FIN	Descriptions	Sample cases	% Sample cases
1	Returned survey - a non-blank survey was returned with no additional information.	12,241	32.28%
2	Return (deceased) – a non-blank survey was returned with information that the sample member was deceased	1	0.00%
6	Returned (separated or retired from Reserves) – a non-blank survey was returned with information that the sampled member had separated or retired from the Reserves.	4	0.01%
7	Return (deployed) – a non-blank survey was returned with additional information that the sample member was deployed.	6	0.02%
8	Returned (other) – a non-blank survey was returned with a reason other than that the sampled member was deceased, incarcerated, deployed, or separated/retired from the Reserves.	10	0.03%
17	Returned Blank (no reason) – a blank survey was returned and no reason was given by sample member	8	0.02%
18	Not returned (deceased) – survey was not returned, sample member deceased.	2	0.01%
19	No Return (incarcerated) – survey was not returned, sample member was incarcerated	1	0.00%
22	Not returned (separated or retired) – a survey was not returned, sample member had separated or retired from the Reserves.	24	0.06%
23	Not returned (active refusal) – survey was not returned, sample member refused to take part in the survey.	14	0.04%
24	No Return (deployed) – survey was not returned, sample member unreachable at this address	1	0.00%
25	Not returned (other) – survey was not returned, sample member was not an active refuser, gave a reason for nonresponse other than being deceased, incarcerated, deployed, unreachable or separated/retired from Reserve.	46	0.12%
26	Not returned (no reason) – survey was not returned, no reason was given by sample member.	20,767	54.77%
27	PND (no address remaining) – all addresses were attempted and returned as postal non-deliveries.	951	2.51%
28	PND (active address) – at the close of field the last address used was found invalid, next available was not attempted.	2,643	6.97%
29	Originally non-locatable (no address at start of mailing) – substantially incomplete or blank address field before the start of administration of the survey, no mailings attempted.	18	0.05%
99	Original ineligible as identified by DMDC	1,181	3.11%
Grand Total		37,918	100.00%

Self-Reported Eligibility

We used the answer to question 1 (SCSVC1) from the questionnaire to determine self-reported eligibility (variable *SSRINEL*) for members. Question 1 and its response categories are:

Q1. In what Service were you on active duty on July 8, 2002?

1. Army
2. Navy
3. Marine Corps
4. Air Force
5. None, I was separated or retired

The responses to this question and their use in assigning eligibility are summarized in Table 4.

Table 4.
Self-Reported Eligibility (SSRINEL)

Self reported eligibility (SSRINEL)	Values of Question 1 (SRSVC1)	Description	Sample Cases	Percent
1. Eligible	1 2 3 4	Yes, person is an active duty member of the Army, Navy, Marine Corps, or Air Force.	12,100	31.9%
2. Ineligible	5	No, a person is not an active duty member of the Army, Navy, Marine Corps or Air Force.	141	0.4%
3. Unknown	. A B	No Response/Multiple Response/No Survey Returned	25,677	67.7%
Grand Total			37,918	100.00%

Completed Questionnaire

The variable that indicates whether a questionnaire was completed (*COMPFLAG*) is shown in Table 5 along with the corresponding percentages.

Table 5.
Complete Questionnaires (COMPFLAG)

COMPFLAG	SAMPLE CASES	% SAMPLE CASES
0 – Incomplete	1,178	9.6%
1 – Complete	11,122	90.4%
Grand Total		100.00%

Weighting Procedures

The analysis of survey data from complex sample designs requires the use of weights to (1) compensate for variable probabilities of selection; (2) adjust for differential response rates; and (3) improve the precision of the survey-based estimates (Skinner *et al.*, 1989). To develop the weights for the *July 2002 SOFS-A*, the following steps were taken. First, base weights equal to the reciprocal of the probability of selection were assigned to each member selected for the sample. Next, the base weights were adjusted for nonresponse using weighting classes defined by relevant variables available on the December 2001 frame file. Finally, the nonresponse-adjusted weights were ratio-adjusted to population counts from the December 2001 frame file with March 2002 DEERS excluded (the beginning of the data collection period). This ratio or poststratification adjustment compensated for some changes in the eligible population between the times of sample selection and March 2002. Details of this weighting methodology are described below.

Calculation of Base Weights

The *July 2002 SOFS-A* sample was randomly selected without replacement from a stratified frame. The overall probabilities of selection vary by design strata in order to satisfy the precision goals specified by the study. Let U be the frame of the N units in the population (i.e., active duty members at the time of sampling). Note that the frame size N includes some units who were ineligible at the time the survey was conducted because, for example, they had left the Service. The frame U was partitioned into H non-overlapping strata U_1, \dots, U_H consisting of N_h units in each stratum h so that

$$N = \sum_{h=1}^H N_h.$$

A simple random sample of size n_h was selected without replacement within each stratum U_h . Given this design, the base weight for the i -th sampled member in stratum h will be calculated as:

$$w_{hi} = \frac{N_h}{n_h} \quad i = 1, \dots, n_h.$$

For each individual classified in stratum h , the base weight is the ratio of the total number of individuals in the stratum to the stratum-level sample size. The base weight w_{hi} is equal to the reciprocal of the probability of selection and is attached to each sample unit in the data file. Note that n_h is the number of persons initially sampled in stratum h without regard to whether or not the member ultimately participated in the survey.

Weighting Adjustments

In an ideal survey, all the units in the inference population are eligible to be selected into the sample and all those that are selected participate in the survey. In practice, neither of these conditions occurs. Some of the sampled units do not respond (unit nonresponse); some sample

units are discovered to be ineligible; and the eligibility status of some units cannot be determined. If these problems are not addressed, the estimates of the survey will be biased. We used nonresponse weight adjustments to deal with unknown eligibility and unit nonresponse. Poststratification was used to account for additional variables of interest (occupation & education level). The following sections describe these methodologies in detail.

Unit Nonresponse Adjustments

Unit nonresponse (i.e., whole questionnaire nonresponse) occurs when a sampled member fails to respond for any reason. For example, nonresponse could result from failure to locate the member because of mobility or invalid/incorrect addresses in the frame, or from the unwillingness of some members to participate in the survey. Because the (unweighted) response rate (defined in a later section) in the *July 2002 SOFS-A* will be substantially less than 100 percent, adjusting for unit nonresponse is an important step in attempting to avoid bias. To compensate for losses due to nonresponse, we adjusted weights in three stages:

- The first non-response adjustment uses logistic regression to adjust for the propensity of sample members to report their eligibility status either via a survey response or self/proxy report of ineligibility. This adjustment is applied to the selection weights.
- The second non-response adjustment uses logistic regression to adjust for survey completion among eligible responses. This adjustment is applied to the eligibility adjusted selection weights from the prior weighting step.
- Finally, a poststratification adjustment is applied to the eligibility and completion adjusted weights from the prior weighting step.

The drawback to nonresponse adjustment is that it increases the variability of the weights and, thus, tends to increase the sampling variance of some estimates (Kish, 1992). Ideally, the reduction in bias from using a nonresponse adjustment more than compensates for the increase in variance.

Note that after the two stages of nonresponse adjustment, the persons with non-zero weight are those in *ER*, *IN_SR*, and *IN_FR*. The members with unknown eligibility (*UNK*) and eligible nonrespondents (*ENR*) have zero weight.

Poststratification Adjustment

The nonresponse-adjusted weights were poststratified to force certain sample estimates of numbers of persons to equal known population totals. In the *July 2002 SOFS-A* the primary functions of poststratification were bias reduction.

The poststratification population totals or controls were produced using the updated December 2001 frame. The December 2001 sample frame was matched against the March 2002 DEERS file, and only members who were eligible on both files were retained.

Given the definitions of the poststrata (see Table 6), the mechanics of the poststratification weight adjustment proceeded as follows. The population was partitioned into

groups (or poststrata) denoted by U_1, \dots, U_G . The groups were by definition mutually exclusive and cover the entire population. Let N_g be the size of U_g , so that $N = \sum_{g=1}^G N_g$. The eligible

respondents in the sample can be also partitioned in groups s_1, \dots, s_G . The expression for the poststratification weighting adjustment factor for all the units classified in cell g is

$$f_g^p = \frac{N_g}{\sum_{i \in s_g} w_i^{A2}}.$$

the poststratified final weight w_i^p , for the i -th sample person classified in poststratum g was then computed as

$$w_i^p = f_g^p w_i^{A2}, \quad i \in s_g.$$

Some sample members who were eligible on the December 2001 frame self-reported (or through a proxy) being ineligible. These persons received a separate ineligibility code (*IN_SR*) as noted earlier. Existence of such persons was evidence that the March 2002 DEERS file also contained some ineligible cases. Sample persons coded as eligible respondents (*ER*) and ineligibles (*IN_SR*) were, consequently, both included in poststratification.

After poststratification the cases with non-zero weight were those in *ER* and *IN_SR*. Cases coded as *ENR*, *IN_FR*, and *UNK* had zero weight.

Table 6.
Definitions of 52 Poststrata

Cell No.	Poststratum Description
001	Army_No College_.321-1.06 Months
002	Army_No College_1.07-1.82 Months
003	Army_No College_1.83-2.58 Months
004	Army_No College_2.59-3.34 Months
005	Army_No College_3.35-4.10 Months
006	Army_No College_4.11-4.86 Months
007	Army_Some College_.321-1.82 Months
008	Army_Some College_1.83-2.58 Months
009	Army_Some College_2.59-3.34 Months
010	Army_Some College_3.35-4.86 Months
011	Army_4-year Degree or Higher_.321-1.06 Months
012	Army_4-year Degree or Higher_1.07-1.82 Months
013	Army_4-year Degree or Higher_1.83-2.58 Months
014	Army_4-year Degree or Higher_2.59-3.34 Months
015	Army_4-year Degree or Higher_3.35-4.86 Months
016	Navy_No College_.321-1.06 Months
017	Navy_No College_1.07-1.82 Months
018	Navy_No College_1.83-2.58 Months
019	Navy_No College_2.59-3.34 Months
020	Navy_No College_3.35-4.10 Months
021	Navy_No College_4.11-4.86 Months
022	Navy_Some College_.321-1.82 Months
023	Navy_Some College_1.83-3.34 Months
024	Navy_Some College_3.35-4.86 Months
025	Navy_4-year Degree or Higher_.321-1.06 Months
026	Navy_4-year Degree or Higher_1.07-1.82 Months
027	Navy_4-year Degree or Higher_1.83-2.58 Months
028	Navy_4-year Degree or Higher_2.59-3.34 Months
029	Navy_4-year Degree or Higher_3.35-4.86 Months
030	USMC_No College_.321-1.06 Months
031	USMC_No College_1.07-1.82 Months
032	USMC_No College_1.83-2.58 Months
033	USMC_No College_2.59-3.34 Months
034	USMC_No College_3.35-4.86 Months
035	USMC_Some College_All Months
036	USMC_4-year Degree or Higher_.321-1.06 Months
037	USMC_4-year Degree or Higher_1.07-1.82 Months
038	USMC_4-year Degree or Higher_1.83-2.58 Months
039	USMC_4-year Degree or Higher_2.59-3.34 Months
040	USMC_4-year Degree or Higher_3.35-4.86 Months
041	USAF_No College_.321-1.06 Months

Table 6. (continued)

Cell No.	Poststratum Description
042	USAF_No College_1.07-1.82 Months
043	USAF_No College_1.83-2.58 Months
044	USAF_No College_2.59-3.34 Months
045	USAF_No College_3.35-4.86 Months
046	USAF_Some College_.321-2.58 Months
047	USAF_Some College_2.59-4.86 Months
048	USAF_4-year Degree or Higher_.321-1.06 Months
049	USAF_4-year Degree or Higher_1.07-1.82 Months
050	USAF_4-year Degree or Higher_1.83-2.58 Months
051	USAF_4-year Degree or Higher_2.59-3.34 Months
052	USAF_4-year Degree or Higher_3.35-4.86 Months

Computation of Variance for Estimates

Variance estimation procedures have been developed to account for the sample design and estimators employed in a complex survey. Using these procedures, factors such as the selection of sample in multiple stages and the use of differential sampling rates to oversample a targeted subpopulation can be appropriately reflected in estimates of sampling error. One method for estimating variances from a complex survey is known as Taylor series variance estimation and replication. Wolter (1985) is a useful reference on the theory and application of this method. Shao (1996) is a more recent review paper that compares Taylor series variance estimation with another method called replication. The next section describes how Taylor series variance estimation can be implemented to compute variances of the estimates for the *July 2002 SOFS-A*.

Software for Computing Taylor Series Variances

The Taylor series method is a linear approximation to a statistic (i.e., the variance of a mean estimate) that is then substituted into the formula for calculating the variance of a linear estimate appropriate for the sample design. The Taylor series method relies on the simplicity associated with estimating the variance for a linear statistic even with a complex sample design and is valid in large samples. In this formulation, the variance strata and primary sampling units (PSUs) must be defined.

SUDAAN[®] (Software for the Statistical Analysis of Correlated Data) (Research Triangle Institute, 1997) is a software package designed to produce variance estimates for complex surveys using the Taylor series method. SUDAAN computes standard errors of the estimates taking into account most features of complex sample designs and estimators. SUDAAN is also capable of reflecting stratum-by-stratum finite population correction (*fpc*) factors in the computation of variances. This is particularly important for the *July 2002 SOFS-A*, where some strata are sampled at high rates.

For descriptive statistics, SUDAAN offers three procedures: PROC CROSSTAB for categorical variables, PROC DESCRIPT for continuous variables, and PROC RATIO for ratios of totals. These procedures can be used to compute statistics of interest, such as estimated totals, means, and percentages along with their corresponding standard errors, design effects, and confidence intervals. SUDAAN can be used to reflect the facts that:

1. the updated March frame contains members who self-reported or proxy-reported as ineligible, or would have been found ineligible if they had been surveyed,
2. the *fpc* is important in some strata, and
3. the weights were poststratified.

SUDAAN can account for the effect of poststratifying weights through the use of POSTVAR and POSTWGT statements. The estimates of standard errors will reflect the effect of poststratification. The option is valid only in PROC DESCRIPT and PROC RATIO and design effects are not computed with this option.

Differences of table cell estimates can also be computed in PROC DESCRIPT and PROC RATIO. The statements that control these calculations are CONTRAST, DIFFVAR, and PAIRWISE.

To reflect the effect of the design in variance estimation, SUDAAN requires variables that indicate the sampled PSUs and the variance estimation strata. The sampled PSU corresponds to the individual sampled person. In most cases the variance estimation strata (TVSTR) are the sampling design strata. However, in some strata, the initial sample was small (less than 30) and was even further reduced due to nonresponse. In these cases the weighting classes, created by collapsing design strata, were used.

The variance strata and PSU indicator variables are part of the dataset delivered to DMDC so estimates and their standard errors can be computed using SUDAAN. The statistical methods report for the *July 2002 SOFS-A* includes an with several examples of SUDAAN programs to illustrate how points (1), (2), (3) above are handled along with examples of how to calculate differences of table cell estimates.

SAS version 8 has two procedures for analyzing survey data: PROC SURVEYMEANS and PROC SURVEYREG. Both use the Taylor series linearization approach to estimate standard errors and can handle sample designs with or without replacement. SURVEYMEANS produces estimates of means, proportions, and totals, while SURVEYREG fits linear regression models (logistic regression is not yet available). No design effects are estimated with either PROC. Estimates of differences or other linear combinations are not available in SURVEYMEANS.

These procedures are new in SAS and do not contain as many features as some other packages. Finite population correction factors can be included in variance estimates for *July 2002 SOFS-A*, but the effect of nonresponse adjustments and poststratification cannot. Accounting for the poststratification frame containing some ineligible units can be done by using

a DOMAIN statement to treat the eligibles as a subpopulation. Table 7 summarizes the survey data analysis capabilities and limitations of the three software packages discussed above.

Table 7.
Features of Three Software Packages for the Analysis of Survey Data

Feature	SUDAAN	SAS	WesVar
Estimation features reflected in variance estimates			
Stratification	x	x	x
Ineligible cases in poststratification frame	x	x	x
Differential weights among cases	x	x	x
Nonresponse adjustments (unknown eligibility, eligible nonrespondents)	x*	NA	x
Poststratification	x	NA	x
Finite population correction factors	x	x	x**
Tables			
Totals/standard errors	x	x	x
Means/standard errors	x	x	x
Proportions/standard errors	x	x	x
Multi-way tables	x	x	x
Differences of cell estimates/standard errors	x	NA	x
Ratios of cell estimates	x	NA	x
Linear regression			
Parameter estimates/standard errors	x	x	x
Confidence intervals for parameters	x	x	x
Logistic regression			
Parameter estimates/standard errors	x	NA	x
Confidence intervals for parameters	x	NA	x
Odds ratios/confidence intervals	x	NA	x
Multinomial logistic regression (unordered categories)			
Parameter estimates/standard errors	x	NA	x
Odds ratios/confidence intervals	x	NA	x
Multinomial logistic regression (ordered categories)			
Parameter estimates/standard errors	x	NA	NA
Odds ratios/confidence intervals	x	NA	NA

Note. NA = not available.

* Available in SUDAAN when estimates based on replication methods are computed.

** Common fpc's at the replicate level.

Calculation of Response Rates

Several rates for the *July 2002 SOFS-A* were computed in accordance with the standards defined by the Council of American Survey Research Organizations (CASRO, 1982). The rates are referred to as:

- Location rate (LR),
- Completion rate (CR), and
- Response rate (RR).

These quantities were computed in such a way that $RR = LR * CR$. The rates are adjusted, as described below, to account for the fact that the eligibility of some units is unknown.

The location rate used for July 2002 SOFS is

$$LR = \frac{\text{adjusted located sample}}{\text{adjusted eligible sample}} = \frac{N_L}{N_E}.$$

The *completion rate* is defined as

$$CR = \frac{\text{usable responses}}{\text{adjusted located sample}} = \frac{N_R}{N_L}.$$

The *response rate* is defined as

$$RR = \frac{\text{usable responses}}{\text{adjusted eligible sample}} = \frac{N_R}{N_E}.$$

where

N_L = Adjusted located sample

N_E = Adjusted eligible sample

N_R = Usable responses.

The adjustments account for the fact that the eligibility status of some persons is unknown so that the proportion of eligibles among the unknowns must be estimated. An assumption in these calculations is that the only ineligibles among the persons with unknown disposition ($ELIG = UNK$) are ones who would have self-reported themselves as ineligible if they had returned a survey form. That is, the updated frame file is assumed to properly identify all other ineligibles. To facilitate computation of the CASRO rates, a separate code (CAS_ELIG) was created that identifies cases that contribute to the components of LR , CR , and RR , as defined in Table 8.

Table 8.
Disposition Codes for CASRO Response Rates (CAS_ELIG)

Eligibility Code for CASRO Response Rates (CAS_ELIG)	FLAG_FIN Values	Weighting Eligibility Code (ELIG)	Sample Cases	Sum of Base Weights	Description
<i>ER</i>	1	<i>ER</i>			Eligible respondent (usable)
<i>ENR_NOQCOMP</i>	1,8	<i>ENR</i>			Eligible nonrespondent (questionnaire not completed)
<i>ENR_BLANK</i>	25	<i>ENR</i>			Eligible nonrespondent (returned blank questionnaire)
<i>ENR_ACTIVE</i>	23,	<i>ENR</i>			Eligible nonrespondent (active refusal)
<i>IN_SR</i>	1,6,18,22	<i>IN_SR</i>			Self-reported or proxy-reported ineligible
<i>UNK_NOLOC</i>	27,28,29	<i>UNK</i>			Unknown eligibility (nonlocatable member)
<i>UNK_NORET</i>	26	<i>UNK</i>			Unknown eligibility (questionnaire not returned)
<i>IN_FR</i>	N/A	<i>IN_FR</i>			Ineligible member in updated frame file

The expressions for the numbers of located persons, eligible persons, and usable responses in terms of CAS_ELIG are given below. As a notational shorthand, CAS_ELIG codes are used to stand for counts of persons in the formulas. For example, *ER* denotes the count of eligible respondents.

Adjusted located sample:

$$\begin{aligned}
 N_L &= (\text{Eligible respondents}) + (\text{Eligible nonrespondents}) + (\text{Estimate of eligibles among unknowns who were located but did not return a questionnaire}) \\
 &= ER + ENR + UNK_NORET \cdot \left(\frac{ER + ENR}{ER + ENR + IN_SR} \right) \\
 &= ER + ENR + UNK_NORET \cdot P_E
 \end{aligned}$$

$$P_E = \frac{ER + ENR}{ER + ENR + IN_SR}$$

where

and $ENR = ENR_NOQCOMP + ENR_BLANK + ENR_ACTIVE$.

Adjusted eligible sample:

$$\begin{aligned}
 N_E &= (\text{Eligible respondents}) + (\text{Estimate of eligibles among all unknowns}) \\
 &= ER + ENR + (UNK_NORET + UNK_NOLOC) \cdot \left(\frac{ER + ENR}{ER + ENR + IN_SR} \right) \\
 &= ER + ENR + UNK \cdot P_E
 \end{aligned}$$

where $UNK = UNK_NORET + UNK_NOLOC$.

Usable responses:

$$N_R = E_R.$$

The adjusted located count, N_L , and the adjusted eligible count, N_E , can also be expressed by subtracting various counts from the total sample as shown below. DMDC has used this on earlier surveys.

$$\begin{aligned}
 N_E &= (\text{Total sample}) \\
 &\quad - (\text{Known ineligible}) \\
 &\quad - (\text{Estimate of self-reported ineligible among non-located unknowns}) \\
 &\quad - (\text{Estimate of self-reported ineligible among other unknowns}) \\
 &= \frac{TOTAL - (IN_FR + IN_SR) - (UNK_NOLOC + UNK_NORET) \cdot \frac{IN_SR}{ER + ENR + IN_SR}}{ER + ENR + UNK \cdot P_E}
 \end{aligned}$$

using the facts that $TOTAL = ER + ENR + IN_FR + IN_SR + UNK_NOLOC + UNK_NORET$ and $IN_SR / (ER + ENR + IN_SR) = 1 - P_E$.

$$\begin{aligned}
 N_L &= (\text{Total sample}) \\
 &\quad - (\text{Known ineligible}) \\
 &\quad - (\text{Non-located unknowns}) \\
 &\quad - (\text{Estimate of self-reported ineligible among other unknowns}) \\
 &= \frac{TOTAL - (IN_FR + IN_SR) - UNK_NOLOC - UNK_NORET \cdot \left(\frac{IN_SR}{ER + ENR + IN_SR} \right)}{ER + ENR + UNK \cdot P_E}
 \end{aligned}$$

$$= ER + ENR + UNK_NORET \cdot P_E$$

The weighted response rates and design effects were calculated for the key domains used in the sample design and are shown in Table 9. In all cases, base weights were used in computing the weighted response rates.

Table 9.
Observed Sample Disposition Counts, Response Rates, and Design Effects

Domain	Observed Sample Disposition Counts					Response Rates		Deff
	All	Resp	Inel	N-Resp	RecInel	EligResp	Complete	
All	37,918	11,060	117	25,242	1,499	31%	91%	2.32
Service								
Army	14,773	3,880	48	10,120	725	28%	90%	2.25
Navy	7,792	2,077	27	5,360	328	28%	91%	2.41
Marine Corps	8,448	1,777	31	6,292	348	22%	88%	2.19
Air Force	6,905	3,326	11	3,470	98	49%	95%	1.98
Paygrade Group 3								
Enlisted	27,575	6,097	63	20,316	1,099	23%	90%	2.01
Officer	10,340	4,963	54	4,924	399	50%	94%	1.68
Paygrade Group 2								
E1-E4	17,021	2,187	34	13,971	829	14%	86%	1.86
E5-E9	10,554	3,910	29	6,345	270	38%	92%	2.05
W1-W5	1,614	787	5	769	53	51%	93%	1.68
O1-O3	4,842	1,994	18	2,572	258	44%	93%	1.65
O4-O6	3,884	2,182	31	1,583	88	58%	94%	1.43
Paygrade Group 1								
E1-E3	9,411	872	14	8,104	421	10%	86%	1.75
E4	7,610	1,315	20	5,867	408	19%	87%	1.94
E5-E6	7,714	2,643	8	4,868	195	35%	92%	2.06
E7-E9	2,840	1,267	21	1,477	75	47%	92%	1.81
W1-W5	1,614	787	5	769	53	51%	93%	1.68
O1-O3	4,842	1,994	18	2,572	258	44%	93%	1.65
O4-O6	3,884	2,182	31	1,583	88	58%	94%	1.43
Imputed duty location - 2 level								
US (Incl. Territories)	31,533	9,045	105	21,070	1,313	30%	91%	2.35
Overseas	6,385	2,015	12	4,172	186	33%	93%	2.15
Imputed duty location								
US (Incl. Territories)	31,533	9,045	105	21,070	1,313	30%	91%	2.35
Europe	3,006	1,045	3	1,865	93	36%	94%	2.00
Asia and Pacific	2,816	884	6	1,861	65	32%	93%	2.09
Other/Unknown	563	86	3	446	28	17%	85%	2.63

Table 9. (continued)

Domain	Observed Sample Disposition Counts					Response Rates		Deff
	All	Resp	Inel	N-Resp	RecInel	EligResp	Complete	
Imputed Receiving BAH								
No BAH	16,731	3,616	27	12,489	599	23%	90%	2.12
Rec BAH	19,996	7,436	84	11,657	819	39%	92%	2.42
Education - 4 Levels								
No College	20,687	1,338	37	18,317	995	7%	86%	1.94
Some College	6,552	4,776	21	1,682	73	74%	91%	2.09
4-year Degree	5,857	2,591	30	3,003	233	47%	93%	2.20
Grad/Prof Degree	4,032	2,354	26	1,534	118	61%	94%	1.68
Race/Ethnicity - 2 level								
White	17,751	6,326	80	10,628	717	38%	92%	2.48
Total minority	19,491	4,732	36	13,988	735	25%	90%	1.94
Race/Ethnicity - 7 level								
White	17,751	6,326	80	10,628	717	38%	92%	2.48
Black	11,309	2,358	21	8,493	437	22%	91%	1.81
Hispanic	5,442	1,366	8	3,865	203	26%	89%	1.94
American Indian/Alaskan Native	135	118	2	14	1	90%	89%	2.15
Asian/Pacific Islander	676	600	3	71	2	89%	89%	1.91
More than one race marked	1,929	290	2	1,545	92	16%	91%	2.22
Family Status								
Single w/ Child(ren)	846	764	2	78	2	91%	94%	2.18
Single w/o Child(ren)	17,215	2,521	35	13,827	832	16%	90%	2.16
Married w/ Child(ren)	6,403	5,857	24	497	25	92%	92%	2.32
Married w/o Child(ren)	13,454	1,918	56	10,840	640	15%	89%	2.33
Spouse Full Unemployment Rate								
Not applicable not in labor force	6,363	5,772	27	535	29	92%	92%	2.30
Working Spouse	5,181	4,729	19	407	26	92%	92%	2.33
Unemployed	615	543	2	66	4	89%	89%	2.13
Dual Service Spouse								
Not Dual Service Spouse	33,914	9,113	108	23,285	1,408	28%	91%	2.30
Dual Service Spouse	4,004	1,947	9	1,957	91	50%	91%	2.29
Gender								
Male	29,831	8,609	92	20,017	1,113	30%	91%	2.30
Female	8,069	2,451	25	5,209	384	32%	92%	2.15

Table 9. (continued)

Domain	Observed Sample Disposition Counts					Response Rates		Deff
	All	Resp	Inel	N-Resp	RecInel	EligResp	Complete	
Gender - Paygrade concatenation								
Enlisted Male	21,544	4,569	44	16,095	836	22%	89%	1.94
Enlisted Female	6,013	1,528	19	4,205	261	27%	91%	1.98
Officers Male	8,285	4,040	48	3,921	276	51%	93%	1.69
Officers Female	2,055	923	6	1,003	123	48%	95%	1.51
Service - Paygrade								
Army Enlisted	11,034	2,078	35	8,378	543	20%	87%	1.90
Army Officers	3,739	1,802	13	1,742	182	51%	93%	1.61
Navy Enlisted	5,546	1,084	10	4,248	204	20%	90%	2.00
Navy Officers	2,244	993	17	1,111	123	48%	93%	1.60
Marines Enlisted	6,196	873	15	5,019	289	15%	85%	1.64
Marines Officers	2,252	904	16	1,273	59	42%	91%	2.26
Air Force Enlisted	4,799	2,062	3	2,671	63	44%	94%	1.87
Air Force Officers	2,105	1,264	8	798	35	61%	96%	1.46
Service - Paygrade								
Army E1 - E4	7,030	740	22	5,849	419	12%	84%	1.77
Army E5 - E9	4,004	1,338	13	2,529	124	35%	89%	1.89
Army W1 - W5	1,302	640	3	613	46	51%	92%	1.43
Army O1 - O3	1,341	552	5	670	114	45%	93%	1.42
Army O4 - O6	1,096	610	5	459	22	57%	95%	1.22
Navy E1 - E4	3,175	330	3	2,693	149	11%	85%	1.82
Navy E5 - E9	2,371	754	7	1,555	55	33%	92%	2.04
Navy W1 - W5	103	49	2	51	1	50%	96%	1.59
Navy O1 - O3	1,204	417	6	686	95	38%	92%	1.61
Navy O4 - O6	937	527	9	374	27	59%	94%	1.25
Marines E1 - E4	4,401	418	7	3,761	215	10%	82%	1.50
Marines E5 - E9	1,795	455	8	1,258	74	27%	88%	1.47
Marines W1 - W5	209	98	.	105	6	48%	92%	1.58
Marines O1 - O3	1,167	382	6	747	32	34%	91%	2.37
Marines O4 - O6	876	424	10	421	21	51%	91%	1.17
Air Force E1 - E4	2,415	699	2	1,668	46	30%	92%	1.68
Air Force E5 - E9	2,384	1,363	1	1,003	17	58%	96%	1.98
Air Force O1 - O3	1,130	643	1	469	17	58%	95%	1.52
Air Force O4 - O6	975	621	7	329	18	66%	96%	1.31

Exhibit 1: Executive Summary of Weighting Procedures

1. First, the sample member’s response and eligibility status were encoded in a sample disposition code (SAMP_DC) whose frequencies are shown in Table ES 1.

Table ES 1.
Sample Disposition Codes (SAMP_DC)

SAMP_DC	Value	Frequency	Percent
Record ineligible based on ELIG	1	1,499	3.95
Self/Proxy-report ineligible	2	21	0.06
Survey Self Report Ineligible	3	96	0.25
Complete Eligible Response	4	11,060	29.17
Incomplete Eligible Response	5	1,051	2.77
Refused/deployed/other	8	50	0.13
Blank	9	8	0.02
PND	10	3,573	9.42
Non-respondents	11	20,560	54.22

2. The next step adjusted for eligibility status determination with the weights for ‘eligibility non-responses’ being distributed to the ‘eligibility responses.’ A variable called *Eligibility Status Response Flag* (ESTATRSP) was constructed to identify these two groups (Table ES 2). Record ineligible cases were excluded from the non-response and poststratification adjustments. A logistic model of ESTATRSP resulted in a response propensity estimate that was inverted and applied to the selection weights.

Table ES 2.
Eligibility Status Response Flag (ESTATRSP)

ESTATRSP	SAMP_DC	Frequency	Percent
	Record ineligible based on ELIG	1,499	3.95
0 Eligibility non-response	Refused/deployed/other	50	0.13
0 Eligibility non-response	Blank	8	0.02
0 Eligibility non-response	PND	3,573	9.42
0 Eligibility non-response	Non-respondents	20,560	54.22
1 Eligibility response	Self/Proxy-report ineligible	21	0.06
1 Eligibility response	Survey Self Report Ineligible	96	0.25
1 Eligibility response	Complete Eligible Response	11,060	29.17
1 Eligibility response	Incomplete Eligible Response	1,051	2.77

- The second non-response adjustment adjusts for completion status where the weights for the ‘incomplete eligible responses’ are distributed to the ‘complete eligible responses.’ A variable called Complete Eligible Response Flag (COMPERSP) was constructed to identify these two groups as shown in Table ES 3. Taking the weight from the previous step, a ratio adjustment equal to the weighted mean of $1 \div \text{COMPERSP}$ within weighting classes was applied to the weight from the previous step.

Table ES 3.
Complete Eligible Response Flag (COMPERSP)

COMPERSP	SAMP_DC	Frequency	Percent
.	Record ineligible based on ELIG from	1,499	3.95
.	Self/Proxy-report ineligible--decease	21	0.06
.	Survey Self Report Ineligible	96	0.25
.	Refused/deployed/other	50	0.13
.	Blank	8	0.02
.	PND	3,573	9.42
.	Non-respondents	20,560	54.22
0 Incomplete Eligible response	Incomplete Eligible Response	1,051	2.77
1 Complete Eligible response	Complete Eligible Response	11,060	29.17

- Next, a poststratification adjustment was applied. The sums of weights from the previous step within the poststrata are ratio adjusted to equal known population totals. The poststrata were selected to reduce bias in the survey estimates by introducing additional variables to the stratification (Service * Education * PersTempo of Occupation). A Poststratification Flag (POSTSTRT) was constructed to identify the cases used for poststratification (Table ES 4). The population file based on the December 2001 active duty master file contains 1,368,423 records; however, 49,014 were found to be ineligible based on a match to the March 2002 Defense Enrollment Eligibility Reporting System resulting in a poststratification total of 1,319,409.

Table ES 4.
Poststratification Flag (POSTSTRT)

POSTSTRT	SAMP_DC	Frequency	Percent
.	Record ineligible based on ELIG from	1,499	3.95
.	Incomplete Eligible Response	1,051	2.77
.	Refused/deployed/other	50	0.13
.	Blank	8	0.02
.	PND	3,573	9.42
.	Non-respondents	20,560	54.22
Poststratification cases	Self/Proxy-report ineligible--decease	21	0.06
Poststratification cases	Survey Self Report Ineligible	96	0.25
Poststratification cases	Complete Eligible Response	11,060	29.17

- Finally, a weighted eligible/ineligible indicator (ELIGFLGW) was constructed to identify those records that are needed to estimate point estimates and variances (Table ES 5). A value of '1' indicates those records needed to obtain point estimates (e.g., means, proportions, and percents). Values of '1' and '2' are needed to produce Taylor series estimates of variances when using SAS PROC SURVEYMEANS.

Table ES 5.
Eligibility Indicator (ELIGFLGW)

ELIGFLGW	SAMP_DC	Frequency	Percent
1 Eligible weighted	Complete Eligible Response	11,060	29.17
2 Ineligible weighted	Self/Proxy-report ineligible	21	0.06
2 Ineligible weighted	Survey Self Report Ineligible	96	0.25
3 Unweighted Non-response	Incomplete Eligible Response	1,051	2.77
3 Unweighted Non-response	Refused/deployed/other	50	0.13
3 Unweighted Non-response	Blank	8	0.02
3 Unweighted Non-response	PND	3,573	9.42
3 Unweighted Non-response	Non-respondents	20,560	54.22
4 Unweighted Record Ineligible	Record ineligible based on ELIG	1,499	3.95

Table ES 6 summarizes which cases were included in each step of the weighting process. The last column shows the general form of the final weight applied to persons in the various disposition categories. Only eligible respondents (*ER*) and self-reported or proxy-reported ineligible (*IN_SR*) received a non-zero final weight.

Table ES 6.
Cases Assigned Weights in Each Step of the Weighting Process by Type of Disposition

Disposition	Nonresponse Adjustment Factor, Step 1	Nonresponse Adjustment Factor, Step 2	Nonresponse Adjusted Weight	Post-Stratification Factor	Final Weight
Eligible Respondents (ER)	f_c^{A1}	f_c^{A2}	$f_c^{A1} f_c^{A2} w_i$	f_g^P	$f_c^{A1} f_c^{A2} f_g^P w_i$
Eligible Non-respondents (ENR)	f_c^{A1}	0	0	0	0
Self- or Proxy-reported ineligible (IN_SR)	f_c^{A1}	1	$f_c^{A1} w_i$	f_g^P	$f_c^{A1} f_g^P w_i$
Ineligible in updated frame (IN_FR)	1	1	w_i	0	0
Unknown (UNK)	0	0	0	0	0

Exhibit 2: Weighting Variables

Four groups of variables were needed for weighting: Completion criteria, Sample disposition, Constructed record variables, and Weighting adjustments. The narrative below provides a brief description of each variable. For most of the variables below, crosstabulations of the variable and its sources are needed to evaluate the validity of the SAS code prior to interpretation of the variable's values.

Completion Criteria. Four completion criteria variables are used to identify complete returns. These variables should be on the dataset prior to weighting because they are also used for deduping the dataset:

QCOMPN	<i>Questions Completed Count.</i> This variable is list of the number of questions answered by each respondent.
QCOMP	<i>Questions Completed Proportion.</i> This variable uses the unique count of the everyone to answer items as the denominator of the proportion.
CRITFLAG	<i>Critical Questions Complete Flag.</i> This variable codes the completeness of the critical items that are unique to the survey.
COMPFLAG	<i>Questionnaire Complete Flag.</i>

Sample Disposition. Four sample disposition variables are used for weighting:

REC_INEL	<i>Record Ineligible Flag.</i> This variable captures all information from the administrative record regarding ineligibility. For SF0207A there is only one record ineligibility variable (ELIG).
SSRINEL	<i>Survey Self-Report Ineligible.</i> This variable codes the eligibility status of a respondent based on survey items and depends on the availability of eligibility information on the population file. If the eligibility information on the population file were incomplete then missing information on the self-report items cause eligibility to be unknown. If eligibility is unknown a complete response is treated like a non-response and is tabulated as a unique form of non-response in SAMP_DC. When SSRINEL is correctly coded this distinction is properly implemented in the rest of the weighting variables. DMDC samples usually have complete eligibility information on the administrative record, therefore, complete returns with missing self-report eligibility questions are usually treated as eligible returns.
SAMP_DC	<i>Sample Disposition Code.</i> This variable codes the information from several variables above (i.e., REC_INEL, SSRINEL, COMPFLAG, QCOMP) and several variables on the Survey Control System (i.e., SCSINEL, BLKREAS, FLAG_FIN).

ELIGFLGW *Eligibility Flag.* This variable is constructed from SAMP_DC and is used to indicate which cases are used during analysis.

Constructed Sampling Variables. One or more constructed variables may be needed for weighting when important information was not known during the development of the sample file. For *July 2002 SOFS* one additional variable was needed to construct the post-strata:

CEDUC Education level.

Weighting Adjustments. Fourteen variables were used for the weighting of the SF0207A survey. Crosstabulations of the variable with its sources were needed to evaluate the validity of the SAS code prior to interpretation of the variable's values. Relatively large weight adjustments or coefficients of variation indicate that collapsing of strata or trimming of adjustments may be needed. For example, if the overall response rate were 0.5 then we would expect to have an overall adjustment factor around 2. In this case, cells adjustment factors greater than 3 would be considered for collapsing. In addition the maximum weights within a stratum were investigated and considered for trimming if they exceeded the population average plus three standard deviations.

SAMP_WGT *Sampling weight (inverse of probability of selection).* Samp_Wgt is considered self-explanatory. See SAS code for details.

WC_STRAT *Weighting class strata (25 or more usable responses).* WC_STRAT may simply be set equal to the sampling stratum when there are 25 or more usable responses in each sampling stratum. For SF0207A the non-response evaluation consisted of looking at the response rates within sampling strata prior to collapsing small strata. WC_Strat will also serve as the variance estimation strata.

ESTATRSP *Eligibility status response flag.* This variable is standardized with respect to Samp_DC. The mean of this variable equals the eligibility status response propensity. It is important to note that record ineligibles are excluded from the weighting because they will also be excluded from the population poststratification counts.

ELIGS_RP *Eligibility status response propensity.* This variable equals the response propensity resulting from a logistic regression model of ESTATRSP on the weighting class strata plus interactions of Service with the remaining key reporting domains (e.g., paygrade, duty location, basic allowance for housing status, education level, race-ethnicity, family status, spouse employment status, dual Service spouse status, gender, and gender x paygrade).

ELIGS_A *Eligibility Status Adjustment.* This variable equals the inverse of the response propensity, ELIGS_RP.

ELIGSA_W *Eligibility Adjusted Weight.* This variable equals SAMP_WGT multiplied by ELIGS_A.

COMPERSP	<i>Complete Eligible response flag.</i> This variable is standardized with respect to SAMP_DC. The mean of this variable equals the completion rate of eligible respondents.
COMPER_A	<i>Complete Eligible Response Adjustment.</i> This variable is the inverse of COMPERSP within WC_STRAT.
COMPER_W	<i>Complete Eligible Response Adjusted Weight.</i> This variable equals ELIGSA_W multiplied by COMPER_A, except for self-report/proxy-report ineligible where it equals ELIGSA_W.
PSTRATA	<i>Poststratification Strata.</i> This is a unique variable to a particular population for the Status of Forces surveys. Different PSTRATA are expected for active duty members, spouses of active duty members, Active Guard and Reserves, and DoD civilians. Standardized PSTRATA are used because non-response analyses used to design the sampling strata had also led to a predetermined poststratification scheme.
NPSTRAT	Poststratification population counts based on the PSTRATA and the particular population file (e.g., month and year) used to draw the sample.
POSTSTRT	Poststratification Flag. This variable is standardized with respect to SAMP_DC. It is important to note that record ineligible are excluded from the weighting because they will also be excluded from the population poststratification counts.
POSTST_A	Poststratification Adjustment. This variable is equal to known population totals for the poststrata divided by the sums of weights within poststrata.
FINALWGT	Final Poststratified Weight. This variable is equal to COMPER_W multiplied by POSTST_A

References

- 1999 Survey of Active Duty Personnel* [CD-ROM]. (2001). Arlington, VA: Defense Manpower Data Center [Producer and Distributor].
- Council of American Survey Research Organizations (1982). *On the definition of response rates* (special report of the CASRO task force on completion rates, Lester R. Frankel, Chair). Port Jefferson, NY: Author.
- Doering, Z. D., Grissmer, D. W., Hawes, J. A., & Hutzler, W. P. (1981). *1978 DoD Survey of Officers and Enlisted Personnel: User's manual and codebook* (Rand Note N-1604-MRAL). Santa Monica, CA: Rand.
- Folsom, R. E. (1991). Exponential and logistic weight adjustments for sampling and nonresponse error reduction. *1991 Proceedings of the Section on Survey Research Methods* (pp. 197-202). Alexandria, VA: American Statistical Association.
- Helba, C., Lee, K., Perry, S., Wright, L.C. & Williams, K.H. *Overview of the 1999 Survey of Active Duty Personnel*. Report No. 2000-008. Arlington VA: Defense Manpower Data Center.
- Kalton, G. and Kasprzyk, D. (1986). The Treatment of missing survey data. *Survey Methodology* 12, 1–16.
- Kass, G. 1980. An exploratory technique for investigating large quantities of categorical data. *Applied Statistics*, 29, 119–127.
- Kavee, J. D., and Mason, R. E. (1997) *DMDC sample planning tool: User's manual (Version 2.1)* (Report No. 97-028) Arlington VA: Defense Manpower Data Center.
- Kish, L. (1992). Weighting for unequal Pi. *Journal of Official Statistics*, 8, 183–200.
- LaVange, L. M., McCalla, M. E., Gabel, T. J., Rakoff, S. H., Doering, Z. D., & Mahoney, B. S. (1986a, 1986b, 1986c). *Descriptions of officers and enlisted personnel in the U.S. Armed Forces: 1985—Supplementary tabulations from the 1985 DoD Survey of Officer and Enlisted Personnel, Vols. 1-3*. Arlington, VA: Defense Manpower Data Center.
- Rizzo, L., & Nixon, M.N. (1996). Nonresponse analysis report for the 1992 DOD Reserve components surveys of officers and enlisted personnel and their spouses. Contract no. MDA 903-92-C-0226. Rockville MD: Westat.
- Rust, K. (1986). Efficient replicated variance estimation. *1986 Proceedings of the Section on Survey Research Methods* (pp. 81-87). Alexandria, VA: American Statistical Association.
- Rust, K. F. and J. N. K. Rao (1996). Variance estimation for complex surveys using replication techniques. *Statistical Methods in Medical Research*, 5: 282–310.

- Shao, J. (1996). Resampling methods in sample surveys, (with discussion). *Statistics*, 27, 203-254.
- Skinner, C., Holt, D., and Smith, T., eds. (1989). *Analysis of complex surveys*. New York: J. Wiley & Sons.
- SPSS (1998), *WesVar[®] Complex samples[™] 3.0, User's Guide*, Chicago: SPSS.
- STATA (1998), *STATA[®] User's guide*, College Station: STATA.
- SUDAAN (1997), *SUDAAN[®] User's manual, release 7.5*. Research Triangle Park: Research Triangle Institute.
- Westat. (2000). *WesVar[™] 4.0 user's guide*, Rockville, MD: Westat.
- Wheless, S.C., Mason, R. E., Kavee, J. D. (1997). *Armed Forces 1996 Equal Opportunity Survey: Statistical methodology report* (Report No. 97-025). Arlington, VA: Defense Manpower Data Center.
- Wright, L. C., George, B. J., Flores-Cervantes, I., Valliant, R., & Elig, T.W. (Eds.). (2000). *1999 Survey of Spouses of Active Duty Personnel: Statistical methodology report* (Report No. 2000-021). Arlington, VA: Defense Manpower Data Center.
- Wright, L. C., Williams, K. H., & Willis, E. J. (2000). *1999 Survey of Active Duty Personnel: Administration, datasets, and codebook* (Report No. 2000-005). Arlington, VA: Defense Manpower Data Center.
- Wright, L. C., Williams, K. H., & Willis, E. J. (In preparation). *1999 Survey of Spouses of Active Duty Personnel: Administration, datasets, and codebook* (Report No. 2000-011). Arlington, VA: Defense Manpower Data Center.
- Wolter, K. (1985). *Introduction to variance estimation*. New York: Springer-Verlag.
- Woodruff, R. (1952). Confidence intervals for medians and other positional measures. *Journal of the American Statistical Association*, 47, 635-646.

**Appendix A.
Sample Allocation**

Table A1.
Sample Allocation

Stratum Number	Frame Size	Sample Allocation	Sample Size	Pct Sampled	Description
1	51,916	73	905	1.74	A_E1-E3_M_Non_US_***_****
2	8,980	17	157	1.75	A_E1-E3_M_Non_OS_***_****
3	1,532	9	96	6.27	A_E1-E3_M_Min_US_SWK_0000
4	28,062	107	1,213	4.32	A_E1-E3_M_Min_US_000_****
5	3,288	17	124	3.77	A_E1-E3_M_Min_Eu_***_****
6	2,520	12	101	4.01	A_E1-E3_M_Min_AP_***_****
7	13,693	37	193	1.41	A_E1~E4_F_Non_**_***_0000
8	2,118	24	104	4.91	A_E1~E4_F_Non_**_000_DUAL
9	9,894	46	462	4.67	A_E1-E3_F_Min_US_***_****
10	1,853	10	69	3.72	A_E1-E3_F_Min_OS_***_****
11	2,459	13	125	5.08	A_E4___M_Non_US_SWK_0000
12	34,938	62	610	1.75	A_E4___M_Non_US_000_****
13	8,256	23	192	2.33	A_E4___M_Non_Eu_***_****
14	3,301	11	72	2.18	A_E4___M_Non_AP_***_****
15	2,369	11	212	8.95	A_E4___M_Min_US_SWK_0000
16	22,187	69	1,341	6.04	A_E4___M_Min_US_000_****
17	5,626	25	305	5.42	A_E4___M_Min_Eu_***_****
18	2,441	10	142	5.82	A_E4___M_Min_AP_***_****
19	7,859	36	607	7.72	A_E4___F_Min_US_***_****
20	2,339	13	137	5.86	A_E4___F_Min_OS_***_****
21	3,491	28	112	3.21	A_E5-E6_M_Non_US_SWK_0000
22	44,055	53	207	0.47	A_E5-E6_M_Non_US_000_0000
23	1,542	20	70	4.54	A_E5-E6_M_Non_US_000_DUAL
24	8,194	45	122	1.49	A_E5-E6_M_Non_Eu_***_****
25	4,334	27	64	1.48	A_E5-E6_M_Non_AP_***_****
26	3,491	26	155	4.44	A_E5-E6_M_Min_US_SWK_0000
27	31,815	137	795	2.50	A_E5-E6_M_Min_US_000_0000
28	2,708	30	153	5.65	A_E5-E6_M_Min_US_000_DUAL
29	7,510	64	193	2.57	A_E5-E6_M_Min_Eu_***_****
30	3,996	33	109	2.73	A_E5-E6_M_Min_AP_***_****
31	3,154	18	53	1.68	A_E5-E6_F_Non_US_***_0000
32	1,173	17	45	3.84	A_E5-E6_F_Non_US_000_DUAL
33	1,400	18	37	2.64	A_E5~E9_F_Non_OS_***_****
34	2,509	16	114	4.54	A_E5-E6_F_Min_US_SWK_0000
35	4,994	15	109	2.18	A_E5-E6_F_Min_US_000_0000
36	2,496	25	151	6.05	A_E5-E6_F_Min_US_000_DUAL
37	2,748	20	76	2.77	A_E5~E9_F_Min_OS_***_0000
38	1,006	14	48	4.77	A_E5~E9_F_Min_OS_000_DUAL

Table A1. (continued)

Stratum Number	Frame Size	Sample Allocation	Sample Size	Pct Sampled	Description
39	18,822	73	287	1.52	A_E7-E9_M_Non_US_***_****
40	3,710	19	48	1.29	A_E7-E9_M_Non_OS_***_****
41	17,401	100	577	3.32	A_E7-E9_M_Min_US_***_****
42	2,582	21	64	2.48	A_E7-E9_M_Min_Eu_***_****
43	1,525	13	43	2.82	A_E7-E9_M_Min_AP_***_****
44	1,209	12	35	2.89	A_E7-E9_F_Non_US_***_****
45	3,232	23	165	5.11	A_E7-E9_F_Min_US_***_****
46	6,241	285	728	11.66	A_W1-W5_M_Non_US_***_****
47	963	46	111	11.53	A_W1-W5_M_Non_Eu_***_****
48	597	30	64	10.72	A_W1-W5_M_Non_AP_***_****
49	2,053	87	266	12.96	A_W1-W5_M_Min_US_***_****
50	599	29	72	12.02	A_W1-W5_M_Min_OS_***_****
51	4,421	84	239	5.41	A_W1~O6_F_Non_US_***_0000
52	1,357	33	87	6.41	A_W1~O6_F_Non_US_000_DUAL
53	1,197	27	70	5.85	A_W1~O6_F_Non_OS_***_****
54	2,606	41	253	9.71	A_W1~O6_F_Min_US_***_0000
55	629	14	72	11.45	A_W1~O6_F_Min_US_000_DUAL
56	703	16	73	10.38	A_W1~O6_F_Min_OS_***_****
57	19,398	173	476	2.45	A_O1-O3_M_Non_US_***_0000
58	645	12	30	4.65	A_O1-O3_M_Non_US_000_DUAL
59	2,824	28	74	2.62	A_O1-O3_M_Non_Eu_***_****
60	1,234	14	32	2.59	A_O1-O3_M_Non_AP_***_****
61	5,611	61	204	3.64	A_O1-O3_M_Min_US_***_****
62	1,233	15	41	3.33	A_O1-O3_M_Min_OS_***_****
63	15,587	229	542	3.48	A_O4-O6_M_Non_US_***_0000
64	561	12	27	4.81	A_O4-O6_M_Non_US_000_DUAL
65	2,035	32	73	3.59	A_O4-O6_M_Non_Eu_***_****
66	744	13	26	3.49	A_O4-O6_M_Non_AP_***_****
67	3,368	52	147	4.36	A_O4-O6_M_Min_US_***_****
68	564	10	24	4.26	A_O4-O6_M_Min_OS_***_****
69	3,007	22	105	3.49	N_E1~E4_M_Non_US_SWK_0000
70	64,738	89	439	0.68	N_E1~E4_M_Non_US_000_****
71	5,477	14	73	1.33	N_E1~E4_M_Non_OS_***_****
72	1,683	14	83	4.93	N_E1-E3_M_Min_US_SWK_0000
73	29,758	145	926	3.11	N_E1-E3_M_Min_US_000_****
74	2,529	13	87	3.44	N_E1-E3_M_Min_OS_***_****
75	13,396	62	183	1.37	N_E1~E4_F_Non_**_***_****
76	8,537	49	281	3.29	N_E1-E3_F_Min_US_***_****
77	1,535	9	54	3.52	N_E1~E4_F_Min_OS_***_****
78	2,519	20	126	5.00	N_E4___M_Min_US_SWK_0000
79	18,150	89	576	3.17	N_E4___M_Min_US_000_****

Table A1. (continued)

Stratum Number	Frame Size	Sample Allocation	Sample Size	Pct Sampled	Description
80	2,518	13	88	3.49	N_E4___M_Min_OS_***_****
81	1,127	9	51	4.53	N_E4___F_Min_US_SWK_0000
82	4,482	25	147	3.28	N_E4___F_Min_US_000_****
83	5,040	48	132	2.62	N_E5-E6_M_Non_US_SWK_0000
84	58,545	89	241	0.41	N_E5-E6_M_Non_US_000_0000
85	927	14	35	3.78	N_E5-E6_M_Non_US_000_DUAL
86	3,591	17	43	1.20	N_E5~E9_M_Non_Eu_***_****
87	3,639	20	44	1.21	N_E5~E9_M_Non_AP_***_****
88	3,935	38	146	3.71	N_E5-E6_M_Min_US_SWK_0000
89	32,968	206	772	2.34	N_E5-E6_M_Min_US_000_****
90	1,813	15	43	2.37	N_E5~E9_M_Min_Eu_***_****
91	3,883	26	79	2.03	N_E5~E9_M_Min_AP_***_****
92	1,156	13	28	2.42	N_E5~E9_F_Non_**_SWK_0000
93	5,885	13	28	0.48	N_E5~E9_F_Non_**_000_0000
94	1,213	21	41	3.38	N_E5~E9_F_Non_**_000_DUAL
95	1,785	16	69	3.87	N_E5~E9_F_Min_**_SWK_0000
96	4,957	24	110	2.22	N_E5~E9_F_Min_**_000_0000
97	1,054	14	55	5.22	N_E5~E9_F_Min_**_000_DUAL
98	1,400	13	36	2.57	N_E7-E9_M_Non_US_SWK_0000
99	19,908	64	172	0.86	N_E7-E9_M_Non_US_000_****
100	7,152	46	174	2.43	N_E7-E9_M_Min_US_***_****
101	2,033	67	148	7.28	N_W1~O3_M_Non_US_SWK_0000
102	18,573	246	549	2.96	N_W1~O3_M_Non_US_000_****
103	663	11	29	4.37	N_W1~O3_M_Non_Eu_***_****
104	878	15	35	3.99	N_W1~O3_M_Non_AP_***_****
105	728	21	55	7.55	N_W1~O3_M_Min_US_SWK_0000
106	4,165	66	179	4.30	N_W1~O3_M_Min_US_000_****
107	453	10	29	6.40	N_W1~O3_M_Min_OS_***_****
108	5,490	100	234	4.26	N_W1~O6_F_Non_US_***_****
109	501	12	31	6.19	N_W1~O6_F_Non_OS_***_****
110	1,684	24	109	6.47	N_W1~O6_F_Min_US_***_****
111	166	8	39	23.49	N_W1~O6_F_Min_OS_***_****
112	443	11	22	4.97	N_O4-O6_M_Non_US_SWK_0000
113	13,951	301	600	4.30	N_O4-O6_M_Non_US_000_****
114	710	15	35	4.93	N_O4-O6_M_Non_Eu_***_****
115	571	13	27	4.73	N_O4-O6_M_Non_AP_***_****
116	1,892	41	96	5.07	N_O4-O6_M_Min_**_***_****
117	43,745	112	1,376	3.15	M_E1-E3_M_Non_US_***_****
118	4,208	13	119	2.83	M_E1-E3_M_Non_OS_***_****
119	19,723	99	1,117	5.66	M_E1-E3_M_Min_US_***_****
120	3,234	20	159	4.92	M_E1~E4_M_Min_OS_***_****

Table A1. (continued)

Stratum Number	Frame Size	Sample Allocation	Sample Size	Pct Sampled	Description
121	2,628	48	261	9.93	M_E1-E3_F_Non_**_***_****
122	1,880	41	390	20.74	M_E1-E3_F_Min_**_***_****
123	16,734	83	346	2.07	M_E4___M_Non_US_***_****
124	1,817	10	34	1.87	M_E4___M_Non_OS_***_****
125	7,684	50	378	4.92	M_E4___M_Min_US_***_****
126	1,043	27	64	6.14	M_E4___F_Non_**_***_****
127	947	24	162	17.11	M_E4___F_Min_**_***_****
128	19,886	133	477	2.40	M_E5-E6_M_Non_US_***_****
129	2,889	24	57	1.97	M_E5~E9_M_Non_OS_***_****
130	11,062	91	475	4.29	M_E5-E6_M_Min_US_***_****
131	2,149	22	71	3.30	M_E5~E9_M_Min_OS_***_****
132	1,390	41	103	7.41	M_E5~E9_F_Non_**_***_****
133	1,466	41	239	16.30	M_E5~E9_F_Min_**_***_****
134	7,132	48	170	2.38	M_E7-E9_M_Non_US_***_****
135	4,304	32	167	3.88	M_E7-E9_M_Min_US_***_****
136	8,597	358	910	10.59	M_W1~O3_M_Non_US_***_****
137	678	33	75	11.06	M_W1~O3_M_Non_OS_***_****
138	2,127	89	265	12.46	M_W1~O3_M_Min_**_***_****
139	970	49	153	15.77	M_W1~O6_F_***_**_***_****
140	4,515	305	680	15.06	M_O4-O6_M_Non_US_***_****
141	478	34	70	14.64	M_O4-O6_M_Non_OS_***_****
142	549	35	93	16.94	M_O4-O6_M_Min_**_***_****
143	37,371	114	408	1.09	F_E1-E3_M_Non_US_***_****
144	4,630	19	64	1.38	F_E1-E3_M_Non_OS_***_****
145	13,481	94	316	2.34	F_E1-E3_M_Min_US_***_****
146	1,667	14	41	2.46	F_E1-E3_M_Min_OS_***_****
147	8,882	23	51	0.57	F_E1-E3_F_Non_US_***_0000
148	1,276	21	43	3.37	F_E1-E3_F_Non_US_000_DUAL
149	1,361	14	30	2.20	F_E1-E3_F_Non_OS_***_****
150	6,457	53	163	2.52	F_E1-E3_F_Min_US_***_****
151	1,910	18	72	3.77	F_E1~E4_F_Min_OS_***_****
152	21,143	47	204	0.96	F_E4___M_Non_US_***_0000
153	1,637	20	77	4.70	F_E4___M_Non_US_000_DUAL
154	3,125	16	70	2.24	F_E4___M_Non_Eu_***_****
155	2,430	13	47	1.93	F_E4___M_Non_AP_***_****
156	8,529	50	386	4.53	F_E4___M_Min_US_***_****
157	2,315	15	99	4.28	F_E4___M_Min_OS_***_****
158	4,072	18	45	1.11	F_E4___F_Non_US_***_0000
159	1,949	30	69	3.54	F_E4___F_Non_US_000_DUAL
160	1,453	18	43	2.96	F_E4___F_Non_OS_***_****
161	4,102	32	214	5.22	F_E4___F_Min_US_***_****

Table A1. (continued)

Stratum Number	Frame Size	Sample Allocation	Sample Size	Pct Sampled	Description
162	3,727	35	97	2.60	F_E5-E6_M_Non_US_SWK_0000
163	49,830	77	207	0.42	F_E5-E6_M_Non_US_000_0000
164	3,787	56	140	3.70	F_E5-E6_M_Non_US_000_DUAL
165	8,521	57	128	1.50	F_E5-E6_M_Non_Eu_***_****
166	5,255	39	78	1.48	F_E5-E6_M_Non_AP_***_****
167	1,449	14	53	3.66	F_E5-E6_M_Min_US_SWK_0000
168	13,543	74	274	2.02	F_E5-E6_M_Min_US_000_0000
169	1,704	23	78	4.58	F_E5-E6_M_Min_US_000_DUAL
170	2,719	27	67	2.46	F_E5-E6_M_Min_Eu_***_****
171	2,561	24	64	2.50	F_E5-E6_M_Min_AP_***_****
172	1,332	13	37	2.78	F_E7-E9_M_Non_US_SWK_0000
173	18,235	26	71	0.39	F_E7-E9_M_Non_US_000_0000
174	1,049	16	40	3.81	F_E7-E9_M_Non_US_000_DUAL
175	4,165	26	56	1.34	F_E7-E9_M_Non_OS_***_****
176	6,968	52	196	2.81	F_E7-E9_M_Min_US_***_****
177	1,603	15	38	2.37	F_E7-E9_M_Min_OS_***_****
178	1,467	16	34	2.32	F_E5~E9_F_Non_US_SWK_0000
179	5,983	11	24	0.40	F_E5~E9_F_Non_US_000_0000
180	4,244	70	139	3.28	F_E5~E9_F_Non_US_000_DUAL
181	2,495	38	68	2.73	F_E5~E9_F_Non_OS_***_****
182	1,438	12	53	3.69	F_E5~E9_F_Min_US_SWK_0000
183	3,407	14	62	1.82	F_E5~E9_F_Min_US_000_0000
184	2,072	26	102	4.92	F_E5~E9_F_Min_US_000_DUAL
185	1,823	24	68	3.73	F_E5~E9_F_Min_OS_***_****
186	22,200	203	419	1.89	F_O1-O3_M_Non_US_***_0000
187	1,065	21	40	3.76	F_O1-O3_M_Non_US_000_DUAL
188	1,721	18	39	2.27	F_O1-O3_M_Non_Eu_***_****
189	1,053	12	23	2.18	F_O1-O3_M_Non_AP_***_****
190	5,405	86	195	3.61	F_O1~O6_M_Min_US_***_****
191	657	11	24	3.65	F_O1~O6_M_Min_OS_***_****
192	6,253	85	184	2.94	F_O1~O6_F_Non_US_***_0000
193	1,706	36	72	4.22	F_O1~O6_F_Non_US_000_DUAL
194	1,092	18	39	3.57	F_O1~O6_F_Non_OS_***_****
195	2,325	93	367	15.78	F_O1~O6_F_Min_**_***_****
196	501	10	19	3.79	F_O4-O6_M_Non_US_SWK_0000
197	18,674	269	504	2.70	F_O4-O6_M_Non_US_000_0000
198	595	14	25	4.20	F_O4-O6_M_Non_US_000_DUAL
199	1,790	27	53	2.96	F_O4-O6_M_Non_Eu_***_****
200	702	13	19	2.71	F_O4-O6_M_Non_AP_***_****
201	12,577	123	455	3.62	Unknown on one or more strata
Total	1,368,424	9,381	37,918	2.77	

This page is reserved for insertion of Standard Form 298, page 1 -- this is best accomplished by replacing this page after the document has been converted to PDF

This page is reserved for insertion of Standard Form 298, page 2 -- this is best accomplished by replacing this page after the document has been converted to PDF

