Appendix B

National Survey of Veterans
Detailed Description of
Weighting Procedures

This document describes the methods used to prepare the sampling weights. Section 1 discusses the weighting method for the list sample surveys—the Active Duty Service Member (ADSM) survey, the ADSM Spouse survey, and the Demobilized National Guards and Reserves (DM) survey. Section 2 presents the weighting method for the address-based sample surveys—the National Survey of Veterans (NSV), the Veteran Spouse Survey, and the Veteran Surviving Spouse Survey. Each section provides a brief description of the sampling frame, the samples for each survey, and the process for developing the sampling weights. Section 3 describes the variance estimation method and the replicate weights for the replication method of estimation.

For each survey, sampling weights were developed to account for the selection probability of each sample unit and to compensate for unit nonresponse. Ratio adjustments were applied such that the sums of the sampling weights match known population targets. The weighting methods followed the approaches agreed to in the Final Sampling Plan (Deliverable 16, Chapter 6, page 49-56, May 21, 2009).

1.   List Sample Weights

Sampling Frames

The list samples for the ADSM and DM surveys were selected from a sampling frame created from a file prepared by VA (see Functional Specification Documentation, July 9. 2009 V1.6), which was transmitted to Westat on August 17, 2009. There were 2,261,139 observations on this file. The sampling frame included all service members whose mailing address was in one of the 50 states and the District of Columbia. Service members with overseas addresses or APO addresses were excluded. By branch of service, the sampling frame included all service members in the Army, Air Force, Navy, Marine Corps, and Coast Guard. Service members in the commissioned corps of the Public Health Service, commissioned corps of the National Oceanic and Atmospheric Administration, and Merchant Marines were excluded from the sampling frame.

 The sampling frame for the list samples included 1,970,814 service members—1,494,035 ADSM and 470,171 DM members in the National Guard and Reserves. VA created the classification variable for identifying the DM members, which included members in the National Guard and Reserves who had one or more activations since 2004 and were deactivated from active duty at the time of VA's data file preparation.

Survey Samples

Two-phase sampling was used to select the sample for the ADSM survey. The first-phase sample was a sample of 6,000 members selected from the ADSM sampling frame. This sample was sent to DMDC to check activation and deployment status (as of August 31, 2009). Those members whose status was demobilized, deployed, deceased, or unknown were excluded from the survey by being assigned disposition codes of ineligible. The ADSM second-phase sample of 2,507 members was then selected from the eligible members of the first-phase sample.

The sample for the ADSM Spouse Survey was also selected using a two-phase sampling approach. The person-type and marital-status codes were used to identify all married sponsors in the ADSM sampling frame. From this list of 341,204 married sponsors, a first-phase sample of 1,000 service members was selected. This sample was then sent to DMDC to append the spouse's name and mailing address. A second-phase sample of 250 spouses was selected from the spouses of ADSMs with known mailing address in the US.

The sample for the DM survey was selected directly from the DM sampling frame consisting of 470,171 members. A systematic sample was selected from this list after implicit stratification resulting from sorting by organization code, age, sex, marital status, person type, education, race and ethnicity. The DM sample contained 3,008 members—1,473 reserves and 1,535 national guards.

Sampling Weights

The sampling weights for each survey were developed in multiple steps. First, a base weight was computed as the inverse of the selection probability from the finite population list. Second, the base weight was adjusted for unknown eligibility and for nonresponse. For the sampled members in the ADSM and the ADSM Spouse surveys, adjustments accounted for the eligibility status at each sampling stage, the probability of selection in the second stage, and survey nonresponse. For the sampled members in the DM survey, adjustment compensated for eligibility status and for nonresponse. Sampling weight adjustments were conducted within adjustments cells defined by age (born since 1979/before 1979), sex (male/female), and service branch (Army, Air Force, and other). Small adjustment cells were collapsed so that there were at least 30 members in most adjustment cells. The final step was to apply ratio adjustment to rake the adjusted weights to frame totals. The raking dimensions were defined to match frame totals by age, sex, and branch.

Tables B-1 through B-3 show the frame totals, and the sum of final weights for ineligible members and respondents in the survey. For the ADSM survey, ineligible members included deactivation, deployment, deceased, or not found by DMDC. For the DM survey, ineligible members included members back on active duty, never been activated from the Reserves or National Guard, never served in at least one of the five service branches included in this survey, and deceased. For the ADSM spouse survey, those with unknown spouse contact information from DMDC were ineligible for the survey.

Table B-1.   Active Duty Service Members Survey: Frame Total and Sum of Final Weights

Characteristics	Frame Total	Sum of weights	
		Ineligibles*	Respondents
Total ADSM	1,494,035	450,674	1,043,361
Born since 1979	904,191	251,685	652,506
Born before 1979	589,844	198,989	390,855
Male	1,277,298	376,963	900,335
Female	216,737	73,711	143,026
Army	667,402	277,581	389,821
Air force	312,772	80,175	232,597
Navy, Marine, Coast Guard	513,861	92,918	420,943

*Ineligible ADSM members included deactivation, deployment, deceased, or not found by DMDC

Table B-2. Demobilized National Guards and Reserves Survey: Frame Total and Sum of Weights

Characteristics	Frame Total	Sum of weights	
		Ineligibles*	Respondents
Total Demobilized	470,171	152,627	317,544
Born since 1979	243,685	112,108	131,578
Born before 1979	226,486	40,519	185,966
Male	385,558	124,098	261,460
Female	84,613	28,529	56,084
Army	298,339	107,222	191,117
Air force	110,814	30,084	80,730
Navy, Marine, Coast Guard	61,018	15,321	45,697

*Ineligible DM members included members back on active duty, never been activated from the Reserves or National Guard, never served in at least one of the five service branches included in this survey and deceased

Table B-3. ADSM Spouse Survey: Frame Total and Sum of Final Weights

Characteristics	Frame Total	Sum of weights	
		Ineligibles*	Respondents
Total Married Sponsor	341,204	82,049	259,155
Sponsor in Army	168,869	52,838	116,031
Sponsor in Other Branches	172,335	29,211	143,124

*Ineligibles include selected members with no spouse contact information available from DMDC.

2. Address Sample Weights

Sampling Frame

The sampling frame for the address samples was the U.S. Postal Service's (USPS's) Computerized Delivery Sequence (CDS) database of delivery point validated-addresses, including emergency service 911- converted addresses. This data source includes all postal addresses receiving mail deliveries. Two-phase sampling was used to select the address samples.

Phase 1 Sample

The first-phase sample for the NSV address-based surveys was purchased from a commercial vendor, Marketing Systems Group (MSG), licensed for CDS products. Westat purchased from MSG a sample of residential addresses in the 50 states and the District of Columbia. Addresses eligible for the survey included all residential addresses, Post Office (PO) Boxes, PO Box throwback addresses, and addresses labeled as seasonal or vacant. The PO Box throwback addresses are street addresses for which residents have requested that mail be delivered instead to their PO Box address. Addresses labeled as drop points were excluded. A drop point is a single delivery point or receptacle that services multiple residences.

The phase 1 address sample was selected from the July 2009 CDS file from MSG. This source included 134, 029, 967 addresses in the target survey population. Westat purchased a total of 1,855,000 addresses in two parts—an initial purchase of 1,100,000 addresses and a second purchase of an additional 755,000 addresses. These samples were selected independently. The approach involved implicit stratification resulting from sorting addresses by ZIP code (5-digit and 4-digit extension), Carrier Route, and Walking Sequence. The sorted list was divided into equally-sized contiguous intervals; and in each interval, one address was selected at random. The length of the intervals was determined by the target sample size.

The phase 1 address sample was assigned into one of three sampling strata based on the results of address matching to VA administrative record files. VA provided four administrative record files for this purpose—the Health Care Enrollment File (8,052,585 observations), the Compensation and Pension File (3,323,611 observations), a retired service member file (3,151,673 observations), and a recent discharge file for Veterans discharged between 2004 and 2009 (896,199 observations).

The data preparation for address matching involved removing administrative records with overseas addresses, incomplete addresses, addresses of deceased members, and overlapping addresses across data sources. The address data fields on each record were standardized into standard address components comprising street number, street pre-direction, street name, street suffix, street post-direction, secondary components (units, apartment, etc.), secondary component number, city name, state, and postal ZIP code. The components were reformatted. For example, street names were capitalized, punctuations removed, and standard codes were applied for common components (e.g., ST for Street, APT for apartment). The software AutoStan (MatchWare, 1996a) was used for this process.

Address matching used the record-linkage software AutoMatch (MatchWare, 1996b). AutoMatch uses a match weight to measure the likelihood of a correct match and a decision rule (a cutoff threshold) to select matched pairs. Address matching was used for stratification purpose. The match outcome can be imperfect and still support a valid sample design if all strata are represented in the selected sample and appropriately weighted. Therefore, a single threshold was used to select matched pairs, and there was limited review of the resulting matches.

The sampling strata were as follows:
- Stratum 1. Addresses of VA benefit users (addresses matched to at least one administrative record in the VA files),
- Stratum 2. Addresses of recent discharge Veterans (addresses matched only to administrative records in the recent discharged Veterans file), and
- Stratum 3. All other addresses not matched to any VA administrative records.

Screener Sample and Sampling Weights

The phase 2 address sample of 137,247 addresses was selected for the screener survey. The allocation of the second-phase sample to the three strata minimized the total sample size subject to constraints that specified target margins of error be were to be achieved by domain-level estimates calculated from the completed person-level questionnaires.

Screener sampling weights were developed for each of the three surveys—the survey of Veterans, the survey of Veteran spouses, and the survey of surviving spouses. For each survey, the screener sampling weights accounted for the inverse of the selection probability in each sampling stage, the subsampling for the spouse and surviving spouse surveys, and for screener nonresponse.

Nonresponse adjustment cells were defined by stratum (and by substrata young/other in stratum 1), Fedex mailing option for nonresponse followup, address type (PO Box, secondary unit type address, and single dwelling addresses), and region (4 Census regions). Adjustment cells were collapsed to ensure a minimum of at least 30 observations per cell.

Veterans Survey Sample and Sampling Weights

The sampling weights for each Veteran were the products of the following components—the final screener household weights (after nonresponse adjustment), a multiple-address adjustment factor, and a ratio adjustment factor to external population control total. All Veterans in the same household were eligible for receiving the Veteran survey questionnaire. For those Veterans who indicated more than one mailing address, we used an approximate compensation factor dividing their sampling weight by 2.

The final Veteran sampling weights were adjusted by ratio adjustment factors such that the sums of weights for the respondents matched external population control totals. Ratio adjustment used a raking procedure to control the marginal totals by age group, sex, race/ethnicity, and service era (a total of 7 dimensions). Table B-4 shows the Veteran population totals from VetPop (http://www1.va.gov/VETDATA/Demographics/Demographics.asp) and the adjusted population control totals used for this survey. The VetPop totals were adjusted for two coverage differences. First, we excluded Veterans in Puerto Rico and the territories. Second, we used the American Community Survey to derive the proportions of Veterans living in group quarters and used these to exclude such Veterans from the control totals.

Table B-4. Veteran Survey: Population Control Totals

Characteristics	VetPop Total 12/15/2009	Veteran Survey Population Control Total*
Total Veterans	22,764,393	22,172,806
Male	20,953,553	20,415,995
Female	1,810,840	1,756,811
18-30 year olds	1,248,220	1,180,127
31-54 year olds	6,975,021	6,792,895
55-74 year olds	9,725,722	9,600,078
75+ year olds	4,815,531	4,599,706
Hispanic	1,213,345	1,175,939
Black Only	2,583,361	2,458,054
Other	18,967,687	18,538,813
Gulf war	5,505,533	5,412,779
Vietnam Era	7,557,404	7,429,396
Korean War	2,552,953	2,487,774
World War II	2,193,476	2,060,745

* Totals excluding Veterans in group quarters and institutions and Veterans in Puerto Rico and territories. The proportion of the Veteran population in group quarters and institutions were derived from ACS 2008

The denominators of the ratio adjustments were calculated from the respondent data in the demographic section of the questionnaire. Item nonresponse in demographic data was relatively minor (2 percent for age and service era variables; 10 percent for gender, and 9 percent for race/ethnicity). Screener data were used to impute for missing Veteran demographic data in single Veteran households. The remaining missing data (1 percent or less) were imputed using hotdeck imputation. Hotdeck imputation was conducted sequentially, one variable at a time. The imputation cells were defined by stratum, presence of spouse/survivor in household, gender, age group, race/ethnicity, and service era. Donors were limited to fewer than three donations.

Veteran Spouse Survey Sampling Weights

The sampling weights for the Veteran spouse survey were developed in the same manner as the Veteran survey. The screener address weights were the products of the inverse of the selection probability in each sample stage and the subsampling rate of addresses for the spouse survey. Nonresponse adjustment cells were formed for the subsample of addresses selected for the spouse survey. The cells were defined again using stratum (and by substrata young/other in stratum 1), Fedex mailing option for nonresponse followup, address type (PO Box, secondary unit type address, and single dwelling addresses), and region (4 Census regions). The adjustment cells were defined to ensure at least 30 observations per cell.

The sampling weight for each Veteran spouse was the product of the sampling weight for the screener household, a sample person nonresponse adjustment factor, and a ratio adjustment factor to external population totals. Spouse nonresponse adjustments used adjustment cells by stratum and by the presence of any Veterans in the household. For those respondents who indicated more than one mailing address, an approximate compensation was applied by dividing their sampling weight by 2.

The ratio adjustment factors were derived such that the sums of weights for the respondents matched external control totals, which were population estimates of married Veterans calculated from the 2008 American Community Survey and adjusted to the target Veteran population in 2009. The ratio adjustments used a raking procedure to control the marginal totals by age group, sex, and race/ethnicity (a total of 3 dimensions). Table B-5 shows the Veteran Spouse population totals. Each control totals was calculated from 2008 American Community Survey data and then multiplied the ratio of the number of Veterans in the 2009 target population to the number in the 2008 target population.

Table B-5. Veteran Spouse Survey: Population Control Totals

Characteristics	Population Control Total*
Total Veteran Spouses	15,027,660
Male	721,326
Female	14,306,334
54 year olds and under	5,558,414
55-74 year olds	7,300,305
75+ year olds	2,168,941
Hispanic/Black only	1,778,204
Other	13,249,456

Source: American Community Survey 2008 population totals adjusted to the population of Veterans in 2009

The denominators of the ratio adjustments were calculated from the respondent data in the demographic section of the questionnaire. Item nonresponse on demographic data was relatively minor (ranging from 3 to 9 percent). We used hotdeck imputation for item nonresponse. Hotdeck imputation was conducted one variable at a time using imputation cells by stratum and controlling for screener data on the characteristics of the Veterans in the household.

Veteran Surviving Spouse Survey Sampling Weights

The sampling weights for the surviving spouse survey were derived using the same method as the Veteran spouse survey. Screener sampling weights were developed for the subset of addresses selected for the surviving spouse survey. Nonresponse adjustments were defined using adjustment cells defined for these addresses. The same criteria that minimum requirement was at least 30 observations in each adjustment cell.

The sampling weight for each surviving spouse was developed following the same approach as the weighing for Veteran spouse. Ratio adjustment used the same factors derived for the Veteran spouse survey.

3. Sampling Weight Adjustments for Nonresponse

This section documents the weighting adjustment cells and weighting adjustment factors to compensate for nonresponse in each survey.

List Sample Adjustment Factors

Table B-6 shows the adjustment cells and the adjustment factors for the list sample surveys. Nonresponse adjustments were applied to compensate for nonresponse with unknown eligibility status and for nonresponse with known eligible status. The composite adjustment factor is a product of these two components.

Table B-6.   List Samples: Response Adjustment Cells and Adjustment Factors

Adjustment Cells	Sampled Members Eligibility Known	Nonresponse Adjustment Factors		
		Eligibility Unknown	Eligible	Composite
ADSM Survey				
Young, Male, Army	47	9.77	1.05	10.25
Young, Male, Air Force	39	6.54	1.03	6.73
Young, Male, Navy	37	7.03	1.05	7.38
Young, Male, Marine and Coast Guard	48	7.00	1.02	7.14
Young , Female, All branches	25	9.44	1.00	9.44
Born before 1979, Male, Army	98	3.40	1.00	3.40
Born before 1979, Male, Air Force	69	2.74	1.00	2.74
Born before 1979, Male, Navy	59	3.49	1.00	3.49
Born before 1979, Male, Marine and Coast Guard	46	2.26	1.00	2.26
Born before 1979, Female, All branches	33	3.91	1.00	3.91
Demobilized National Guard and Reserve				
Young , Male, All branches	123	10.07	1.00	10.07
Young , Female, All branches	40	7.83	1.00	7.83
Born before 1979, Male, All branches	366	3.35	1.00	3.35
Born before 1979, Female, All branches	53	4.34	1.00	4.34
Spouse of ADSM Survey				
All Married ADSM	71	3.52	1.0	3.53

The adjustment cells were defined by age (young-born since 1979/born before 1979, sex (male/female), and service branch (army, air force, navy, marine corps, and coast guards). For the ADSM survey, the nonresponse adjustment factors ranged between 6.73 to 10.25 for young ADSMS born since 1979; and these factors ranged between 2.26 to 3.4 for the older ADSM (born before

1979). The size of the adjustment factors was similar for the DM survey. Small adjustment cells were combined across branches. Service branch was not used in the DM survey because of small cell sizes. The goal was to combine cells with comparable response rates and to keep cell size at least 30.

Address Screener Sample - Adjustment Factors for the Veteran Survey

Table B-7 shows the adjustment cells and adjustment factors for the screener address sample. The five factors used to form the nonresponse adjustment cells were Stratum (1=Addresses of VA benefit users, 2=recent discharge veterans in the PSMAF, and 3=all other addresses), domain in stratum 1 (young 30 and under/other), FedEx (experiment/control group), address type (PO Box addresses, SUD-secondary unit addresses, and single house); and Census region (Northeast, Midwest, South, West).

Table B-7.   Address Screener Samples: Response Adjustment Cells and Adjustment Factors (Full Sample)

Stratum	Domain	FedEx	Address	Region	Number Response	Adjustment Factor
1 UA Users	Young (30& under)	No	PO Box/SUD	All	87	3.89
1 UA Users	Young (30& under)	No	Single House	NorthEast	60	2.46
1 UA Users	Young (30& under)	No	Single House	MidWest	125	2.15
1 UA Users	Young (30& under)	No	Single House	South	195	2.93
1 UA Users	Young (30& under)	No	Single House	West	79	3.34
1 UA Users	Other (31+)	No	PO Box	NorthEast	36	1.76
1 UA Users	Other (31+)	No	PO Box	MidWest	51	1.77
1 UA Users	Other (31+)	No	PO Box	South	116	1.74
1 UA Users	Other (31+)	No	PO Box	West	71	1.71
1 UA Users	Other (31+)	No	Secondary Unit	NorthEast	33	1.85
1 UA Users	Other (31+)	No	Secondary Unit	MidWest	37	1.93
1 UA Users	Other (31+)	No	Secondary Unit	South	86	1.70
1 UA Users	Other (31+)	No	Secondary Unit	West	67	1.99
1 UA Users	Other (31+)	No	Single House	NorthEast	404	1.42
1 UA Users	Other (31+)	No	Single House	MidWest	610	1.42
1 UA Users	Other (31+)	No	Single House	South	1,237	1.55
1 UA Users	Other (31+)	No	Single House	West	532	1.52
2 PSMAF Recent Discharge	NA	No	PO Box/SUD	All	59	5.15
2 PSMAF Recent Discharge	NA	No	Single House	NorthEast	43	3.33
2 PSMAF Recent Discharge	NA	No	Single House	MidWest	73	3.07
2 PSMAF Recent Discharge	NA	No	Single House	South	154	3.34
2 PSMAF Recent Discharge	NA	No	Single House	West	66	3.20
2 PSMAF Recent Discharge	NA	Yes	PO Box/SUD	All	81	3.65

Table B-7. Address Screener Samples: Response Adjustment Cells and Adjustment Factors (Full Sample) (Continued)

Stratum	Domain	FedEx	Address	Region	Number Response	Adjustment Factor
			Nonresponse Cells			
2 PSMAF Recent Discharge	NA	Yes	Single House	NorthEast	66	2.11
2 PSMAF Recent Discharge	NA	Yes	Single House	MidWest	96	2.35
2 PSMAF Recent Discharge	NA	Yes	Single House	South	199	2.66
2 PSMAF Recent Discharge	NA	Yes	Single House	West	70	2.84
3 All Other Address	NA	No	PO Box	NorthEast	324	4.29
3 All Other Address	NA	No	PO Box	MidWest	449	3.38
3 All Other Address	NA	No	PO Box	South	812	4.62
3 All Other Address	NA	No	PO Box	West	657	3.98
3 All Other Address	NA	No	Secondary Unit	NorthEast	1,226	4.57
3 All Other Address	NA	No	Secondary Unit	MidWest	1,012	4.17
3 All Other Address	NA	No	Secondary Unit	South	1,345	5.39
3 All Other Address	NA	No	Secondary Unit	West	1,263	4.60
3 All Other Address	NA	No	Single House	NorthEast	3,945	2.90
3 All Other Address	NA	No	Single House	MidWest	6,480	2.71
3 All Other Address	NA	No	Single House	South	8,320	3.38
3 All Other Address	NA	No	Single House	West	4,768	3.14
3 All Other Address	NA	Yes	PO Box	All	182	3.57
3 All Other Address	NA	Yes	Secondary Unit	NorthEast	109	3.83
3 All Other Address	NA	Yes	Secondary Unit	MidWest	97	3.14
3 All Other Address	NA	Yes	Secondary Unit	South	142	3.96
3 All Other Address	NA	Yes	Secondary Unit	West	125	3.52
3 All Other Address	NA	Yes	Single House	NorthEast	434	2.05
3 All Other Address	NA	Yes	Single House	MidWest	598	2.12
3 All Other Address	NA	Yes	Single House	South	815	2.52
3 All Other Address	NA	Yes	Single House	West	473	2.23

Adjustment used 48 cells—17 cells for addresses in stratum 1, 10 cells for addresses in stratum 2, and 21 cells for addresses in stratum 3. By domains in stratum 1, the adjustment factors ranged from 2.15 to 3.89 for cells in the young veteran domain; and they ranged from 1.42 to 1.93 for cells in the older veteran domain. In stratum 2, the adjustment factors ranged by FedEx experiment and address type. For example, the adjustment factor was 2.11 for single house type address in Northeast, FedEx followup); it was 5.15 for PO Box and addresses for secondary units.

Address Subsamples - Adjustment Factors for the Surviving Spouse and the Veteran Spouse Surveys

Nonresponse adjustment cells were defined for the subsamples of addresses for the surviving spouse survey and for the veteran spouse surveys. Table B-8 shows the adjustment cells and adjustment factors for the Surviving Spouse Survey and Table B-9 shows those for the Veteran Spouse Survey.

Table B-8. Address Screener Samples: Response Adjustment Cells and Adjustment Factors (Surviving Spouse Subsample)

Stratum	Nonresponse Cells				Number Response	Adjustment Factor
	Domain	FedEx	Address	Region		
1, 2	All	No	PO Box	NE/MW	73	1.75
1, 2	All	No	PO Box	West	97	1.88
1, 2	All	No	PO Box	South	58	1.84
1, 2	All	No	Secondary Unit	NorthEast	30	1.95
1, 2	All	No	Secondary Unit	MidWest	36	1.88
1, 2	All	No	Secondary Unit	South	68	1.90
1, 2	All	No	Secondary Unit	West	64	2.04
1, 2	Young (30& under)	No	Single House	NE/MW	38	2.21
1, 2	Young (30& under)	No	Single House	South/West	46	3.22
1 UA Users	Other (31+)	No	Single House	NorthEast	348	1.41
1 UA Users	Other (31+)	No	Single House	MidWest	515	1.42
1 UA Users	Other (31+)	No	Single House	South	1,032	1.56
1 UA Users	Other (31+)	No	Single House	West	445	1.52
3 All Other Address	NA	No	PO Box	NorthEast	214	4.30
3 All Other Address	NA	No	PO Box	MidWest	290	3.32
3 All Other Address	NA	No	PO Box	South	518	4.59
3 All Other Address	NA	No	PO Box	West	409	4.00
3 All Other Address	NA	No	Secondary Unit	NorthEast	778	4.58
3 All Other Address	NA	No	Secondary Unit	MidWest	662	4.10
3 All Other Address	NA	No	Secondary Unit	South	860	5.52
3 All Other Address	NA	No	Secondary Unit	West	800	4.63
3 All Other Address	NA	No	Single House	NorthEast	2,524	2.89
3 All Other Address	NA	No	Single House	MidWest	4,198	2.69
3 All Other Address	NA	No	Single House	South	5,257	3.40
3 All Other Address	NA	No	Single House	West	3,085	3.13
2,3	NA	Yes	PO Box	NE/MW	50	2.48
2,3	NA	Yes	PO Box	South	40	4.35
2,3	NA	Yes	PO Box	West	41	3.12
2,3	NA	Yes	Secondary Unit	NorthEast	66	3.95
2,3	NA	Yes	Secondary Unit	MidWest	59	3.34
2,3	NA	Yes	Secondary Unit	South	95	3.61
2,3	NA	Yes	Secondary Unit	West	83	3.58
2,3	NA	Yes	Single House	NorthEast	293	2.00
2,3	NA	Yes	Single House	MidWest	386	2.09
2,3	NA	Yes	Single House	South	513	2.58
2,3	NA	Yes	Single House	West	310	2.21

Table B-9. Address Screener Samples: Response Adjustment Cells and Adjustment Factors (Veteran Spouse Subsample)

Stratum	Domain	FedEx	Address	Region	Number Response	Adjustment Factor
		Nonresponse Cells			**Number**	**Adjustment**
Stratum	**Domain**	**FedEx**	**Address**	**Region**	**Response**	**Factor**
1 UA Users	All	No	PO Box	All	66	1.79
1,2	All	No	Secondary Unit	All	72	1.90
1, 2	All	No	Single House	NorthEast	112	1.53
1, 2	All	No	Single House	MidWest	177	1.47
1, 2	All	No	Single House	South	336	1.61
1, 2	All	No	Single House	West	141	1.66
3 All Other Address	NA	No	PO Box	NorthEast	56	4.09
3 All Other Address	NA	No	PO Box	MidWest	80	3.29
3 All Other Address	NA	No	PO Box	South	148	4.57
3 All Other Address	NA	No	PO Box	West	113	3.86
3 All Other Address	NA	No	Secondary Unit	NorthEast	244	4.28
3 All Other Address	NA	No	Secondary Unit	MidWest	184	4.10
3 All Other Address	NA	No	Secondary Unit	South	251	5.18
3 All Other Address	NA	No	Secondary Unit	West	252	4.23
3 All Other Address	NA	No	Single House	NorthEast	710	2.89
3 All Other Address	NA	No	Single House	MidWest	1,207	2.63
3 All Other Address	NA	No	Single House	South	1,501	3.38
3 All Other Address	NA	No	Single House	West	898	3.04
3 All Other Address	NA	Yes	PO Box	All	35	3.29
2,3	NA	Yes	Secondary Unit	All	70	3.96
2,3	NA	Yes	Single House	NorthEast	87	1.98
2,3	NA	Yes	Single House	MidWest	107	2.14
2,3	NA	Yes	Single House	South	156	2.39
2,3	NA	Yes	Single House	West	70	2.69

The nonresponse cells for the subsamples used the same constraints to limit cell size and to collapse cells with comparable response rates. There were relatively few responses in stratum 2. Adjustment cells were defined combining stratum 1 and stratum 2 cells. All adjustment cells under FedEx experiment were combined with similar experiment cells in Stratum 3. Nonresponse compensation used 36 adjustment cells for the Surviving Spouse Survey and 24 cells for the Veteran Spouse Survey.

Person Nonresponse - Adjustment Factors for the Surviving Spouse and the Veteran Spouse Surveys

For the Veteran survey, sampling weights for the respondents were directly post-stratified to external population control totals. For the Spouse and Surviving Spouse Surveys, a person-level nonresponse was applied before population ratio adjustment. This was necessary because there were

no external population control totals of the surviving spouse population. Hence, the adjustments were done in two stages. Each sample was adjusted for nonresponse, then the population ratio adjustment factors for the spouse survey were applied to the surviving spouse survey.

Table B-10 and B-11 show the nonresponse adjustment cells and the adjustment factors for the Veteran Spouse survey and the Surviving Spouse survey. For the Veteran Spouse Survey, the nonresponse adjustment cells were defined by stratum, domain, and the screener response on age group of veteran in single veteran household. (Age group was imputed if there were veterans in multiple age groups.) The adjustment factor was 1.62 for spouse of young veterans sampled in stratum 1 and 2. It was 2.34 for spouse of veterans where the screener age group was 75+. For the Surviving Spouse survey, the adjustment cells were defined by stratum and presence of a veteran in the screener response.

Table B-10.   Veteran Spouse Survey: Nonresponse Adjustment Cells and Adjustment Factors

	Nonresponse Cell		Number of	Adjustment
Stratum	Domain	Screener Age Group of Veteran	Response	Factor
1, 2	75 and under	All	93	1.62
1	75+	All	65	1.64
3	All	Veteran Under 55 years old	124	1.77
3	All	Veteran 55-74 years old	188	1.78
3	All	Veteran 75+ years old	60	2.34

Table B-11.   Veteran Surviving Spouse Survey: Nonresponse Adjustment Cells and Adjustment Factors

	Nonresponse Cell	Number of	Adjustment
Stratum	Present of a Veteran in screener	Response	Factor
1, 2	Any	50	1.92
3	No veteran	569	1.41
3	At least one veteran	139	2.02

4. Ratio Adjustments

This section documents the raking dimensions and the ratio adjustment factors applied to the sampling weights in each survey. Each table shows the raking dimension, the control totals, the mean and the coefficient of variation (% relative error) of the adjustment factors.

Ratio adjustments for the ADSM, the DM samples used age, sex, and service branch to form raking dimensions and adjustment cells. Control totals were derived from the sampling frame. The data used were: YR1979 1= born since 1979 2= born after 1979; SEX_N 1= Male 2=Female BRANCH_N 1=Army 2=Air Force 3=Navy 4=Marine Corps 5= Coast Guards. For the spouse of ADSM survey, the sum of weights were post-stratified to

Ratio adjustment for the veteran surveys used age, sex, race/ethnicity, and service era. The data used were YR1979 1= born since 1979 2= born after 1979; SEX 1= Male 2=Female, Race/Ethnicity 1= Hispanic, 2= non-Hispanic Black only and 3=non Hispanic other race. The service eras were Gulf war era, Vietnam War era, Korean War era, and World War II era (yes/no). The adjustment factors for the Veteran Spouse survey were applied to the Surviving Spouse survey.

List Samples

Table B-12.   Raking dimensions, Frame Control Total, Mean and CV of Adjustment Factors, ADSM Survey

DIM1	YR1979	SEX_N	BRANCH_N	FRMTOTAL	Mean Adjustment	CV of Adjustment
111	1	1	1	331,646	1.003	0.00
112	1	1	2	133,908	1.010	0.00
113	1	1	3	137,322	1.020	0.00
114	1	1	4,5	163,786	0.996	0.00
121	1	2	1,2,3,4,5	137,529	1.146	29.76
211	2	1	1	242,114	0.991	0.00
212	2	1	2	104,142	0.988	0.00
213	2	1	3	112,998	0.989	0.00
214	2	1	4,5	51,382	0.988	0.00
221	2	2	1,2,3,4,5	79,208	1.003	27.30

Table B-12.   Raking dimensions, Frame Control Total, Mean and CV of Adjustment Factors, ADSM Survey (Continued)

DIM2	BRANCH_N	FRMTOTAL	Mean Adjustment	CV of Adjustment
1	1	667,402	1.014	5.48
2	2, 4, 5	513,861	0.956	11.58
3	3	312,772	1.097	18.44

Table B-13.   Raking dimensions, Frame Control Total, Mean and CV of Adjustment Factors, DM Survey

DIM1	YEAR1979	SEX	BRANCH	FRMTOTAL	Mean Adjustment	CV (%) Adjustment
111	1	1	1	137,013	1.115	0.00
112	1	1	2,3,4,5	56,965	0.803	8.39
121	1	2	All	49,707	1.016	5.24
211	2	1	1	107,363	0.982	0.00
212	2	1	2,3,4,5	84,217	1.022	7.87
221	2	2	All	34,906	0.970	6.83

Table B-13. Raking dimensions, Frame Control Total, Mean and CV of Adjustment Factors, DM Survey (Continued)

DIM2	BRANCH	FRMTOTAL	Mean Adjustment	CV (%) Adjustment
1	1	298,339	1.020	5.40
2	2,4,5	61,018	1.070	10.30
3	3	110,814	0.923	8.86

Table B-14.   Raking dimensions, Frame Control Total, Mean and CV of Adjustment Factors, Spouse of ADSM Survey

DIM1	BRANCH (Sponsor)	FRMTOTAL	Mean Adjustment	CV (%) Adjustment
1	1	168,869	0.981	0.00
2	2,3,4,5	172,335	1.019	0.00

Veteran Samples

Table B-15.   Raking dimensions, Frame Control Total, Mean and CV of Adjustment Factors, Veterans Survey

DIM1	SEX	MSTOTAL	Label	Mean Adjustment	CV (%) Adjustments
1	1	20,415,995	Male	0.98	48.55
2	2	1,756,811	Female	1.51	48.10
DIM2	AGEGRP	MSTOTAL			
1	1	1,180,127	Under 30	2.69	13.73
2	2	6,792,895	31-54	1.62	20.26
3	3	9,600,078	55-74	0.82	16.67
4	4	4,599,706	75+	0.70	17.50
DIM3	RACETHN3	MSTOTAL			
1	1	1,175,939	Hispanic	1.29	50.97
2	2	2,458,054	Blackonly	1.73	43.49
3	3	18,538,813	Other	0.95	45.27
DIM4	GULFWAR	MSTOTAL			
1	1	5,412,779	Yes	1.84	33.72
2	2	16,760,027	No	0.85	30.96
DIM5	VIETNAM	MSTOTAL			
1	1	7,429,396	Yes	0.79	18.70
2	2	14,743,410	No	1.19	51.78
DIM6	KOREAN	MSTOTAL			
1	1	2,487,774	Yes	0.68	15.66
2	2	19,685,032	No	1.07	49.62
DIM7	WWII	MSTOTAL			
1	1	2,060,745	Yes	0.63	15.92
2	2	20,112,061	No	1.07	49.14

Table B-16. Raking dimensions, Population Control Total, Mean and CV of Adjustment Factors, Veteran Spouse Survey

DIM1	Gender	SPTotal	Label	Mean Adjustment	CV (%) Adjustment
1	1	721,326	Male	0.45	23.56
2	2	14,306,334	Female	1.22	26.03
DIM2	AGEGRP	Sptotal			
1	1,2	5,558,414	Under 54 years old	1.69	18.14
3	3	7,300,305	55-74 years old	1.03	21.57
4	4	2,168,941	75+ years old	0.82	25.32
DIM3	RACETHN3	SPTotal			
1	1,2,	1,778,204	Hispanic, Black only	1.38	38.34
3	3	13,249,456	Non Hispanic ,other	1.10	33.19

Gender 1=Male 2=Female Age Group 1= 30 or under, 2= 31 to 54, 3=55 to 74 and 4= 75+
RaceEthn3 1=Hispanic 2=Black Only 3=Other

5.   Replicate Weights for Variance Estimation

Replicate weights were constructed for variance estimation using the grouped jackknife estimation method (JK1 in WesVar). Replicate weighting is a convenient and often used technique to compute estimates of sampling variance. Once replicate weights are constructed, little additional effort is required to estimate the variance of most estimates. A number of widely available statistical software packages now support the use of replicate weights, including Wesvar, SUDAAN, the *survey* package in R, and Stata.

To create the replicate weights, the same basic strategy was followed for each component survey. In each case, the first-phase sampling units were assigned into 100 replicate groups following the sample selection order. Then for each replicate group, sampling weights were computed following the same sequence of adjustments as the full sample weight.

With the exception of the DM survey, the sample designs for each of the components of the NSV followed a two-phase or three-phase strategy. In each case, the first-phase sample followed a relatively simple, unclustered sampling design. In fact, the variance from the first-phase sample can be approximated in each instance by the variance of a simple random sample. Each of the surveys sampled only a small fraction of the total universe of interest, so it is not important to include finite population corrections in any of the variance calculations.

- During the first phase of the ADSM Survey, an unclustered sample of 6,000 members was sent to DMDC to identify ineligible cases based on activation and deployment status. A second-phase sample of 2,507 members was then selected from those classified as eligible on the basis of the information from DMDC.

  The application of the general variance estimation approach to the ADSM Survey began by assigning the first-phase sample of 6,000 members to 100 groups of 60 each. On average, the groups each contain about 25 members of the second-phase sample, although the method of assignment does not prevent random variation between groups in the numbers selected. The initial jackknife weights for each of the 100 replicates were formed by deleting (in turn) one of the groups from the base-weight calculations and reweighting the remaining groups by the factor 100/99. The replicate weights were then adjusted by repeating the same weighting steps as for the full sample, including both the nonresponse adjustment to the second-phase

sample and the final raking, where all of the full sample information is replaced by the weighted results for the replicate.

- Similarly, the sample for the ADSM Spouse Survey was selected on the basis of a first-phase unclustered sample of 1,000 service members who were married sponsors. Where possible, DMDC provided contact information for the spouses of the sponsors, and a second-phase sample of 250 spouses was drawn from those with known mailing addresses in the U.S.

  The variance design for the ADSM Spouse Survey followed principles similar to the ADSM Survey. The sample of 1,000 married sponsors in the first phase was assigned to 100 groups of 10 each. After DMDC identified those having a spouse with a mailable address in the U.S., a second-phase sample of 250 spouses was selected. Again, each step of estimation was replicated, including the adjustments for nonresponse and raking to the frame totals, and incorporated into the replicate weights.

- The DM survey used a single-phase design, drawing an unclustered sample of 3,008 members from the list frame. The sample included 1,473 reserve members and 1,535 national guard members.

  The DM sample was divided into 100 groups of about 30 each for variance estimation. Because of the single-phase design, this represents a standard application of the grouped jackknife.

- The first-phase sample for the Veterans Survey was a systematic sample of addresses, purchased as two independent samples of 1,100,000 and 755,000 from MSG. Although the systematic sample would have provided a degree of beneficial geographic stratification, the variance from the first-phase design can be approximated by the variance of a simple random sample without replacement. The addresses were stratified on the basis of matching to administrative files and a second-phase sample of addresses was selected for the screening sample, with sampling rates varying by second-phase strata. All Veterans identified by the screening questionnaire became eligible for the detailed questionnaire.

  For the Veterans Survey, the approach assigned the 1,855,000 addresses in the first-phase sample to 100 groups. Address matching to VA administrative data was used to stratify the

first-phase sample, from which a second-phase sample of 137,247 addresses was selected. The 100 groups would have each contained an average of 1,347 addresses in the second-phase sample, but the actual number would have varied to some degree.

- The spouses and surviving spouses of Veterans were sampled using the same first- and second-phase sample as the Veterans. Responses to the screening questionnaire were used to for a further phase of sample, technically creating a three-phase design in both cases.

  The assignments of first-phase addresses to groups for Veterans were also used for the Veteran Spouse Survey and the Veteran Surviving Spouse Survey.

In the general case, variance estimation for two-phase sampling is an active area of research, especially in the last ten years. (A new paper (Farrell and Singh, 2010) just appeared in the most recent issue of *Survey Methodology,* for example.) A general approach to the problem is to decompose the variance into the separate contributions from the first and second phases of sampling. This decomposition is described at length, for example, by Särndal, Swensson, and Wretman (1992, Ch. 9) and by Fuller (2009, Sec. 3.3). Finite population corrections are readily incorporated into this approach, and they can be particularly important in the expression for the variance from the second phase sample.

In some situations, simpler approaches are also appropriate. The application of the grouped jackknife to NSV is supported by its close ties to work of Rao and Sitter (1995), who applied the jackknife to a two-phase sample design with simple random sampling at both phases, and examined the variance of the ratio estimator. Because (1) a large number of replicates are used (100) for NSV, (2) the first-phase samples closely approximate simple random samples, and (3) the application does not require incorporation of finite population corrections, Rao and Sitter's findings for the jackknife extend to the use of the grouped jackknife in each of the NSV applications. The approach is far simpler to implement and provides a method of constructing replicate weights supported by several software packages.

References

Farrell, P.J, and Singh, S. (2010). Some contributions to jackknifing two-phase sampling estimators. *Survey Methodology*, 36, 57-68.

Fuller, W.A. (2009). *Sampling Statistics*. Hoboken, NJ: John Wiley & Sons.

Matchware Technologies, Inc. (1996a). *Autostan: Generalized Standardization System User's Manual*. Silver Spring, MD: Matchware Technologies, Inc.

Matchware Technologies, Inc. (1996b). *Automatch: Generalized Record Linkage System User's Manual*. Silver Spring, MD: Matchware Technologies, Inc.

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

Rao, J.N.K., and Sitter, R.R. (1995). Variance estimation under two-phase sampling with application to imputation for missing data. *Biometrika*, 82, 453-460.