Supporting Statement

Collections of Information

Employing Statistical Means

Part B

July 24, 2009

TABLE OF CONTENTS

Tables

Exhibits

Supporting Statement B - Collections of Information Employing Statistical Methods

1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection methods to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.

The 2009 National Survey of Veterans, Active Duty Service Members, Activated National Guard and Reserve Members, Family Members, and Survivors Design Study (NSV) targets six distinct populations. The largest target population for this survey is the population of Veterans living in households in the U.S. The survey population excludes Veterans in institutional or non-institutional group quarters. Based on the 2006 results from the American Community Survey, we expect this population to be about 23.4 million Veterans living in about 22.5 million households.

The survey will also collect information from the following other populations:

- Spouses of Veterans
- Surviving Spouses of deceased Veterans
- Active Duty Service Members (includes currently activated National Guard and Reserve members)
- Demobilized National Guard and Reserve members
- Spouses of Active Military (including spouses of currently activated National Guard and Reserve members).

The survey will exclude any active duty member currently in theater or in combat.

Six different questionnaires have been designed for these populations, including one for Veterans. Each of the populations will be surveyed with its own questionnaire, with the exception of active duty military and currently activated National Guard and Reserves, who will use the same questionnaire. Table 1 summarizes the target populations, their associated questionnaires, and their estimated population sizes.

The Department of Defense maintains essentially complete lists for three of these populations, namely, active duty military (which includes currently activated National

Guard and Reserves), demobilized National Guard and Reserves, and spouses of active military. (The primary data source for these lists is the DoD Defense Enrollment Eligibility Reporting System [DEERS] file.) Many, but not all, of the DEERS variables are present on copies of the DEERS file maintained by the VA. The VA files are of suitable quality to form sampling frames for these populations, and the NSV will use these frames to draw a list sample of individuals.

Table 1. Target Populations for the 2009 NSV, Associated Questionnaires, and Estimated Population Sizes

Initial Population of Interest	Final Target Subpopulation	Questionnaire	Sample Frame	Approximate population size (in millions)
Veterans	(1)Veterans	Veteran Survey	Address-Based	23.4 M
Active military	(2) Currently active duty service members and currently activated NG/Reserves*	Active Duty Service Member Survey	List**	1.4 M
Demobilized NG/RC	(3) Demobilized NG/RC	Demobilized National Guard/Reserve Survey	List	0.6 M
Family members (spouses) and survivors	(4) Active duty spouse	Active Duty Spouse Survey	List	0.6 M
	(5) Veteran spouse	Veteran Spouse Survey	Address-Based	13.5 M
	(6) Surviving spouse	Surviving Spouse Survey	Address-Based	2.0 M

*The survey will exclude active military and currently activated National Guard and Reserve members in combat/theater.
** List frame for active duty military service members include currently activated National Guard and Reserve members.

Table 2 provides the designated sample size for each of the list frames. For each of the three frames, individuals will be selected with equal probability. The sample will be geographically stratified. The survey will exclude those in institutions or in areas of active conflict. For purposes of design, the response rate has been estimated at 50%, so the designated sample size has been set at 3,158 to yield an expected 1,579 completed interviews.

Table 2. Designated Sample Size and Expected Final Response, for Target Subpopulations Sampled Through List Frames

Final Target Subpopulation	Designated Sample Size	Expected Final Response
Active military (including currently activated NG/RC)	1,504	752
Demobilized NG/RC	1,504	752
Active duty spouse	150	75
Total	3,158	1,579

The remaining populations will be sampled through an address frame. Although VA administrative files, such as the Health Care Enrollment File, contain lists of Veterans and their addresses, less than half of all Veterans can be identified this way. Instead, the sampling strategy for Veterans rests on drawing a stratified, unequal probability sample of addresses. The survey administration will be in two phases: (1) a screening phase to identify Veterans and members of other target populations; and (2) an extended questionnaire phase to collect the characteristics of interest.

For the screening phase, a random sample of U.S. addresses will be purchased from a vendor. The sample households will be contacted first by mail, and a household respondent will be asked to complete the short screening paper questionnaire.The screener will ascertain whether the household includes one or more members who are Veterans, one or more members who are spouses of Veterans, or one or more members who are surviving spouses of Veterans. Because of their greater personal interest in the survey, we expect that households including one or more Veterans, Veteran spouses, or surviving spouses will respond at a higher rate than other households.

For the extended questionnaire phase, responding households that do not include anyone in the NSV target populations will not be asked to participate further. Those households with one or more individuals in the target populations become eligible for selection to receive an extended questionnaire.

In selecting the sample for the screening phase, the sample of addresses purchased from the vendor will be stratified into three sampling strata. The three sampling strata are further described in response to Question 2 (pp. 6–9). Based on an analysis of data from the pilot study for the 2009 NSV, we currently estimate that the stratum sizes will be approximately 7.8 million, 0.3 million, and 104 million households.

Table 3 provides our planned allocation of the sample to the sampling strata. Based on the results of the pilot survey, we expect that the overall response rate in the screening phase will be 35% and the overall response rate to the extended questionnaire will be 66%. Under our assumptions that 80% of the households in strata 1 and 2 will contain Veterans, Table 3 gives the total number of households expected to respond to the screener, the number of Veteran households expected to respond to the screener, and the expected number of Veteran respondents to the extended questionnaire.

In addition to the design goals for Veterans as a whole, the 2009 NSV has goals for subpopulations of Veterans defined by period of service, age, sex, race, and Hispanic origin. Table 4 identifies the target subpopulations, their precision goals, and their associated expected samples during the screening and extended questionnaire phases. All Veterans identified during the screening phase will be included in the extended phase.

Table 3. Expected Allocation and Sample Sizes for Veterans

Sampling Strata	Sampled Addresses (For mailing)	Expected Number of Deliverable Addresses	Expected Screening Response (Number of Households)	Expected Veteran HH Screening Response (Number of Households)	Expected Veteran Extended Response (Number of Veterans)
Stratum 1	8,359	8,075	5,277	4,683	3,189
Stratum 2	3,071	2,807	759	408	222
Stratum 3	122,471	109,367	36,401	11,897	8,071
Total	**130,207**	**120,249**	**42,437**	**16,988**	**11,482**

Table 4. Estimated Sizes of the NSV Veteran Population and Subpopulations

Subpopulation	Precision goal*	Veterans (millions)	Expected Screening Yield	Expected Extended Yield
WW II	3.0%	2.27	4,664	1,643
Korean Conflict	3.0%	2.62	5,321	1,835
Vietnam	3.0%	7.65	14,185	4,010
Persian Gulf	3.0%	5.51	9,524	1,926
Black	4.0%	2.60	2,256	511
Hispanic	4.0%	1.33	4,193	752
Young (≤ 30 yrs)	4.0%	1.26	2,340	401
Female	4.0%	1.82	3,633	1,372
All Veterans		**23.44**	**16,988**	11,482

*Half width of the 90% confidence interval for the estimate of a percentage of 50%.

The two remaining populations, spouses of Veterans and surviving spouses, will also be identified during the screening phase. (For purposes of the 2009 NSV, spouses of Veterans who are Veterans themselves will be surveyed as Veterans, not as spouses of Veterans.) In the majority of cases, the screening questionnaire will identify Veteran spouses in households also containing Veterans. In other words, the marital status of most Veteran spouses can be classified as “married, spouse present,” although some will be

"married, spouse absent." Thus, most Veteran spouses will be among the 16,988 households identified in the next-to-last column of Table 3. For the most part, surviving spouses will not be currently living with Veterans, so they will primarily be included among the 25,449 non-Veteran households that are counted among the expected 42,437 screening respondents in Table 3 but primarily excluded from the 16,988 expected responses from Veteran households.

The sample for spouses of Veterans and surviving spouses will be drawn from the screening returns. Table 5 shows the sample to be drawn from respondents to the screening questionnaire and the expected response rates to the extended questionnaires.

Table 5. Designated Sample Size, and Expected Final Response, for Remaining Target Subpopulations Sampled Through the Address Frame

Final Target Subpopulation	Sample for Extended Questionnaire (To Mail)	Expected Final Extended Response (Complete)
Veteran Spouses	860	602
Surviving Spouses	107	75

In Table 1, the estimate of 13.5 million married, spouse-present Veteran spouses was based on a public-use file from the American Community Survey. The response to the screening survey will provide an ample number of Veteran spouses from which to sample. Our plan is to sample married, spouse-present Veteran spouses by subsampling the Veteran sample. For example, if the Veteran is a woman married to and living with her non-Veteran husband, we will plan to sample him as a Veteran spouse conditional on her selection in the sample. Veteran spouses living apart from their spouses, that is, "married, spouse absent," will be identified as a separate stratum and sampled separately in proportion to the relative sizes of these two groups of Veteran spouses. A memo with the pilot's final results will be forthcoming.

Whereas the American Community Survey provides a reliable estimate of the number of married, spouse-present Veteran spouses, we lack a comparable source for an estimate of the number of surviving spouses of Veterans. If there are 2 million or more such surviving spouses, then the screening questionnaire could provide approximately 674 households with surviving spouses, a more than adequate sample size for purposes of subsampling.

A previous cycle of the NSV was conducted in 2001. The initial sampling was based on overlapping RDD and list frames. The RDD component had a response rate of 67.6% to the screener, which identified eligible or potentially eligible Veterans. For Veterans determined to be eligible in the RDD component, the response rate to the extended interview was 76.4%. For the list sample, 54.0% were eligible Veterans who

completed the interview and 8.8% were determined to be ineligible, but eligibility could not be determined for 34.0%.

2. Describe the procedures for the collection of information including:

- **Statistical methodology for stratification and sample selection,**
- **Estimation procedure,**
- **Degree of accuracy needed for the purpose described in the justification,**
- **Unusual problems requiring specialized sampling procedures, and**
- **Any use of periodic (less frequent than annual) data collection cycles to reduce burden.**

As noted in the response to the previous question, the 2009 NSV will collect information from distinct target populations using different questionnaires and sampling strategies. The previous section divided the discussion of the overall sampling design into (1) a list sample for four target populations to be surveyed with three questionnaires and (2) an address sample for the remaining three target populations and their corresponding questionnaires. This division will be continued here.

For the three target populations to be sampled from list samples, only national estimates are sought. Consequently, an equal probability sample from the list frame is appropriate. The sample will be geographically stratified by sorting on the ZIP Code of the available address and forming the same number of strata as the target sample size and selecting one from each.

Westat will not survey the active military and activated National Guard and Reserves on assignment in a conflict zone, primarily Iraq and Afghanistan. As described in documentation from the Department of Defense, deployment codes on DEERS will enable us to identify these individuals and select our sample excluding them. These codes are not present on the VA copy of DEERS, however. Consequently,, a two-phase sampling method will be used. Using an estimate of the excluded population, we will compute a larger initial sample size so that we can obtain from DoD/DMDC the deployment codes for the initial sample and then reject sampled individuals in the excluded population and, in expectation, obtain our target sample size with the remaining sampled individuals.

For simplicity, we will independently sample active military and their spouses. The DoD files maintain separate records for each, including different contact information if appropriate.

With respect to estimation, the list frames will provide accurate totals for use in ratio estimation. All three target populations can be divided by branch of the military (Army, Navy, Air Force, and Marines) and region (Northeast, Midwest, South, West) determined by U.S. address, and these 16 cells will be used to form simple ratio estimates. On average, the cells will contain about 67 respondents each, an appropriate

number for ratio estimation. Cell collapsing will be implemented, however, for cells with fewer than 20 individuals.

Because the design is initially self-weighting and unclustered, we expect the reliability of the targeted respondents for each population to correspond to a simple random sample of the same size. For populations with a target sample size of 752, a 95% confidence interval for a population proportion estimated at 50% will be +/- 3%. Particularly because this is the first time these populations have been surveyed with the questionnaires developed for the 2009 NSV, this precision should prove sufficient for reasonably anticipated analytic purposes. Analysis of the results by region or branch of the military is not anticipated.

Sampling for the remaining three target populations, including Veterans, will be based on address. The primary frame for the sample will be based on the Delivery Sequence File (DSF) of the U.S. Postal Service, a file listing all addresses in the U.S. eligible to receive mail. We will obtain an equal probability sample of addresses from the DSF from a commercial vendor. Before drawing the sample, the vendor will stratify the addresses geographically using ZIP Code and other delivery information available on the file.

For purposes of stratification, we will also assemble a separate list of addresses based on administrative records from VA and DoD. The administrative records pertain to Veterans and identify likely addresses where Veterans may live. VA will help with the initial processing of the VA and DoD/DMDC files by encrypting the Social Security Number across the files into a unique ID. VA will then provide the VA and DoD files containing the unique ID to Westat for additional processing. Based on the experience from the Pilot Study for this survey, we will use files from VA extracted from the Health Care Enrollment File and the Compensation and Pension File and a file from DoD extracted from the Prior Military Service Address File. Westat will merge the DoD and VA files on the basis of the unique ID to create a file of unique individuals, setting flags to indicate which files contained records for each individual. Finally, we will convert the merged file of individuals into a file of unique addresses where Veterans may live.

By matching the merged VA/DoD information with the sample from the DSF, Westat will construct three sampling strata:

1. DSF addresses that can be matched to addresses identified on the VA file,
2. DSF addresses that can be matched to addresses on the DoD file, but not the VA file, and
3. DSF addresses that cannot be matched to addresses on the DoD and VA files.

The first two sampling strata are of interest for the main study because (1) stratum 1 represents a high-yield stratum of households likely to include Veterans who are users of VA services, and (2) stratum 2 should include a number of households with Veterans, particularly young Veterans, who do not yet use VA services. The pilot study indicates

that the second stratum is small but still of interest because its yield for young Veterans is higher than it is in stratum 1 or stratum 3.

Westat will then subsample the resulting address file at varying rates by stratum to yield the designated sample sizes in Table 1. Additionally, in strata 1 and 2 Westat will oversample addresses appearing to contain young (age 30 or younger), old (age 75 or older), or female Veterans, based on information from the merged VA/DoD files.

Technically, the design is a two-phase sample, with the selection of a sample from the DSF by the vendor constituting the first phase and the merging of stratification information from the VA/DoD and subsequent subsampling representing the second phase.

The screening questionnaire will provide further information to classify Veterans by age, sex, race, and Hispanic origin. All responding Veterans will be sent an extended questionnaire. Only spouses and surviving spouses identified at the screener will be subsampled. Rather than a full three-phase design, which would stratify and subsample the screening responses after effectively all screening returns have been received or declared nonresponse, selection into the extended sample will be pre-designated based on the characteristics reported in the screener. This approach will allow the extended questionnaire sample to be selected on a flow basis, accelerating data collection.

Westat will be able to compute initial weights based on the inverse probability of selection, including the selection probability from the DSF, the probability from the second phase design, and the subsampling probability for the extended questionnaire. Comparison of estimates of the Veteran population based on these weights to independently derived estimates of the Veteran population will inform us on the effective coverage of the survey, including the combined impact of non-response and coverage issues from the frame.

The initial weights will then be revised through ratio estimation to conform to known totals for the Veteran population. Veteran age is potentially correlated with several characteristics to be measured through the extended questionnaire, so at a minimum the survey weights will be controlled through ratio estimation to independent estimates of the Veteran population by age produced by VA. Raking-ratio estimation to a more detailed set of controls by sex, race, and Hispanic origin in addition to age will be attempted, again based on estimates prepared by VA.

Because differential sampling is employed at different stages of selecting the Veteran sample, the survey weights vary and increase the sampling variance relative to a simple random sample of the same size. The Veteran sample is virtually unclustered, however, so weight-variation is the principal factor that must be taken into account in designing the sample. The sample was designed to achieve 90% confidence intervals of +/- 3% for a population proportion of 50% for each of the subpopulations in Table 4. Some groups, particularly Persian Gulf Veterans, are assigned larger sample sizes than

required to meet this minimum standard. The 90% confidence interval for Veterans as a whole will be about +/–1%.

The sample of married, spouse-present Veteran spouses can be selected to overlap with the Veteran sample. A separate stratum of married, spouse-absent Veteran spouses will be formed and sampled at approximately the same rate as married, spouse-present Veteran spouses. The results from the pilot study will furnish an improved estimate of the size of this population, which will be used to set appropriate sampling rates for this group.

The pilot sample will also provide an indication of the size of the surviving spouse population, currently crudely estimated in Table 1 at 2 million. If the number of surviving spouses is in fact 2 million or more, then the screening sample should yield a sufficient number of them to achieve the goals stated in Table 5. If the true number of surviving spouses is instead lower, then it may be necessary to reduce the stated goals for precision for surviving spouses.

The American Community Survey provides controls for married, spouse-present Veteran spouses suitable for ratio estimation. Raking by broad age group, sex, race, and ethnicity will be possible. But similar controls will be lacking for married, spouse-absent Veteran spouses. We propose to compute ratio factors by age for married, spouse-present Veteran spouses and to apply these factors to married, spouse-absent Veteran spouses. We view this as an underadjustment for non-response, because the Veteran survey may prove less engaging to married, spouse-absent Veteran spouses than married, spouse-present ones.

We are unaware of a credible, independent estimate of surviving spouses. The survey responses weighted by the inverse probability of selection may provide the most readily available option. Here again, use of a ratio factor computed for a related population, such as married, spouse-present Veteran spouses, can be justified as a minimal correction for nonresponse.

In terms of expected reliability, the survey estimates for Veteran spouses and surviving spouses will also meet a standard of producing 90-percent confidence intervals of +/- 3% for population proportions estimated at 50%.

3. Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.

The purpose of the data collection is to estimate the awareness for VA benefits and services for six different groups, including: 1) Veterans, 2) spouses of Veterans, 3) surviving spouses of Veterans, 4) active duty military, 5) spouses of active duty military

and 6) demobilized National Guard and Reserves. As described above, to collect these data, two basic methodologies will be followed. For the Veterans, spouses of Veterans, and surviving spouses of Veterans, a two-step procedure will be used. This uses a sample frame of addresses (hereafter referred to the "address frame"). For the active duty service members, spouses of active duty, and demobilized National Guard/Reserves, the survey will be completed in one step using a list of names taken from different administrative files (hereafter referred to as the "list frame").

Methods for the Address Frame

As noted above, the procedures for the address frame will follow a two-stage design. The first stage will send out a short (screening) survey to each sampled address. This survey will ask a household respondent to report on whether there are any Veterans, spouses of Veterans, or surviving spouses of Veterans living in the household (Attachment 1). It will also include a question on the mode preference for the extended interview. After the screening survey is returned, a second (extended) survey will be sent to those households where an eligible member of the population resides (i.e., a Veteran, spouse of a Veteran or a surviving spouse of a Veteran). One extended questionnaire will be sent for each Veteran identified at the screener, but spouses and survivors will be subsampled. These questionnaires are provided in Attachments 2, 3, and 4, one for each of the three different groups.

The study will be using a number of methods to maximize the response rate and data quality. First, the survey will make multiple contacts with the household, following the Total Design Method recommended by Dillman, et al. (2008). Both the first and second stages will have a total of four contact attempts. These letters will be sent on VA stationary, with outer envelopes containing the VA logo. For the screener, the first mailing will be a pre-notification letter that alerts members of the household of the NSV (Attachment 5). Approximately 5 days later, the survey instrument will be mailed, with a cover letter providing the details about the study as well as an insert to encourage response (Attachment 6). If the household does not respond after another week, a postcard will be sent that reminds them to return the questionnaire (Attachment 7). If the questionnaire is not returned in another 2 weeks, a follow-up letter (Attachment 8) and survey will be sent to the address. Veteran that indicated on the screener questionnaire that they were 30 years old or younger or indicated they were Hispanic and do not respond to the first or second mailing of the extended questionnaire will receive a third mailing of the questionnaire that will be sent by express delivery.

When a screener is returned, it will be scanned and entered into a database. All identified Veterans will be sent an extended interview. If a spouse of a Veteran, and/or a surviving spouse of a Veteran is identified, a sampling algorithm will be applied to decide whether to send an extended interview. If the address is selected to receive an extended interview, the mode preference of the household will be used to decide on the type of request to send to the sampled individual. As noted above, the screener contains a question which asks respondents whether they would prefer to respond by the web or by

mail. If they choose the web, they are asked to provide an e-mail address. Respondents that choose postal mail will be mailed a paper copy of the appropriate survey.

Web preference. For those eligible sample members choosing the web for the extended survey, an e-mail and a letter will be sent to them. Attachment 9 provides this e-mail and letter for each of the three populations. The letter will also contain an insert that encourages their use of the Internet to complete the survey (also in Attachment 9). One week after these materials are sent, a reminder postcard will be sent to the respondent (Attachment 10). Two weeks after the postcard, a second letter will be sent that requests the respondent to fill out the survey (Attachment 11). This request will include a paper survey, while also giving the respondent the option to answer on the web. Two weeks after this request, a third request will be sent asking respondents to fill out the survey (Attachment 12). This will also include a paper survey and the option to complete the survey by telephone.

Mail preference. The sequence of mailings for this group will be very similar to the web preference group. The only difference is that this group will not be given an option to complete the survey on the Internet. They will be sent a paper questionnaire at each of the mailings, and the last request will include the option to complete the survey by telephone. Attachments 13–16 provide the letters for this group of respondents. Exhibit 1 illustrates the data collection procedure for the address sample.

Use of Express Delivery. In addition to following the above methodology, we will send an express package for the second mailing of the screener for a sample of 10,000 non-respondents. The purpose is to test whether this procedure is effective at reaching groups that respond at lower rates to the screener. For example, the results of the Pilot study indicated that younger veterans, female veterans and Hispanic veterans responded to the screener at lower rates. Express packages have been shown to increase response, especially for groups with lower response rates. The 10,000 sample will be drawn from stratums 2 and 3, as these are the groups that have the lowest response rates. This procedure could not only improve the overall response rate, but it will provide an opportunity to conduct non-response bias analyses.

Nonresponse Bias Analyses for Address Frame

The survey design for the 2009 NSV incorporates several features to maximize response rates in an increasingly difficult survey environment. In spite of these efforts, past experience with the 2001 NSV and current experience with similar mail surveys form the evidence to extrapolate a response rate for the 2009 NSV well below 80 percent. The design in fact assumes an overall 35 percent response by Veterans to the extended questionnaire. Following OMB guidelines, we propose a plan for a nonresponse bias study. The study would evaluate the potential impact of nonresponse on the ability to generalize from the 2009 NSV to the population of Veterans, Spouses of Veterans and Surviving Spouses.

We will use two approaches to this analysis. The first will follow the general principle of using the best available *auxiliary information* in the analysis, where auxiliary information data are data obtained from sources other than the survey's data collection efforts. Example of sources of auxiliary data are data contained in the sampling frame—such as, data from DEERS—or aggregate data from other surveys—such as, the American Community Survey. The second approach will compare estimates for the respondents sent the express package to comparable groups that were not sent the package.

Use of Auxiliary Information. The Veteran component of the 2009 NSV is by far the largest. It is the one component of the 2009 NSV where specific subpopulation targets, such as female and young Veterans, have been specified. The large sample size increases the chances that practically significant findings will also be statistically significant. There are also two possible approaches to studying nonresponse with auxiliary data, given the two-phase design of the data collection through screening and extended questionnaires. Because essentially all of the substantively important data on Veterans come from the extended questionnaire, one approach is to analyze the effect of nonresponse on the data from the extended questionnaire, without specific reference to the data from the screening questionnaire. The second approach is to analyze nonresponse as a two-step process, namely the separate effects of nonresponse to the screening questionnaire and nonresponse to the extended questionnaire among the respondents to the screening questionnaire. This section discusses the two approaches in this order.

Although the 2009 NSV should follow the standard practice of computing nonresponse rates according to current professional standards, these rates by themselves will not be a direct reflection on the quality of the NSV data. The design has incorporated the assumption, which will be tested during the pilot study, that Veteran households (including those with Veteran spouses or surviving spouses) will respond to the screening questionnaire at substantially higher rates than other households. Response to the screening questionnaire by non-Veteran households is desirable from a cost perspective, but the response rate could be arbitrarily lower than the assumed 30 percent without affecting the ability of the 2009 NSV to represent the target populations of interest.

An important source of auxiliary data to inform the nonresponse analysis for Veterans is the VetPop model. The marginal totals from the model will be used to adjust the final weights for Veterans through raking-ratio estimation. However, parts of the nonresponse bias study can be informed by an alternative set of estimates using the base weights. The comparison of the NSV respondent data weighted by the base weights to the VetPop estimates represents the combined effect of nonresponse and incomplete coverage by the address frame.

Following method 1, *weighted response rates by subgroups* the first step in the nonresponse bias study could compare two sets of rates:

- Standard nonresponse rates overall and by sampling stratum. Because stratum 3 is expected to contain a lower proportion of Veterans than the other two strata, its response rate is expected to be lowest of the three strata.

- Effective coverage rates for each of the 2009 NSV subpopulations of interest. The *effective coverage rate* will be the ratio of the estimated total based on the extended questionnaires for a Veteran subpopulation, weighted by the base weights, to the corresponding estimated total from the VetPop model.

Although raking-ratio estimation will achieve consistency between the VetPop model and the 2009 NSV data for the specific margins used in the estimation, the weighted NSV data will not necessarily agree with the estimates from the VetPop model for other possible subdomains. An exploratory analysis could compare estimates from the VetPop and the NSV for other subdomains of interest, using statistical methods to differentiate between random variation likely due to sampling error and differences suggestive of systematic bias.

The second possible approach to analyze the bias from nonresponse is to investigate response to the screening and extended questionnaires separately. Separating the two stages of response permits comparison of respondents and nonrespondents using data available for both groups. Unfortunately, the majority of nonresponse is expected during the screening phase, and the address-based sampling frame is a limited resource for a bias study. The majority of sample cases will be drawn from stratum 3, where most addresses will not include a Veteran, and where the only available information comprises location and address characteristics. Thus, analysis of nonresponse to the screening questionnaire based on data available for both respondents and nonrespondents will be limited to strata 1 and 2.

The remaining phase of nonresponse occurs after people sampled on the basis of the screening questionnaires are mailed extended questionnaires. The sample design assumes 70 percent response in this phase. The screening questionnaire will establish the presence or absence of a Veteran, so analytic comparison of extended survey respondents to nonrespondents will be less problematic than for the screening phase. For each of the characteristics measured by the screening questionnaire, it will be possible to compare extended questionnaire respondents to nonrespondents (method 2 and 3). As in a number of other nonresponse bias studies, the association between different variables and response status could be summarized with multivariate methods.

The frame-based approach should provide a fairly clear statistical account of nonresponse to the extended questionnaire during the second phase of collection because the list frames contain information about Veterans and both respondents and nonrespondents to the extended questionnaire are also Veterans. It will be more problematic to evaluate nonresponse to the screening questionnaire because nonrespondents can be either Veterans or non-Veterans, the screener response rates will most likely be different for Veteran and non-Veteran households, and the list frames contain information only about Veterans.

Although there are clear differences, the similarities between selecting the Veteran sample and selecting the sample of Veteran spouses leads to a number of parallels in the design of a nonresponse bias study for spouses. The target sample size is far smaller than that for Veterans, and no analysis of subgroups of Veteran spouses has been explicitly targeted in the design.

Following the analysis for veterans, as a first step, nonresponse rates will be computed according to standard formulas, recognizing that the rates will confound nonresponse by households containing Veterans or spouses of Veterans with nonresponse by households without Veterans or spouses of Veterans.

Although the VetPop model does not track spouses of Veterans, the American Community Survey is a reliable source for estimates of spouses currently married to and living with a Veteran. The margins from the ACS will be incorporated into the weighting. Following a strategy parallel to the veterans, it is possible to investigate the joint effect of coverage and nonresponse, that is, *effective coverage rates*, by deriving alternative estimates based on the base weights only. To parallel the method for Veterans, the following rates will be computed for married, spouse-present Veteran spouses:

- Standard nonresponse rates overall and by sampling stratum, and
- Effective coverage rates for each of the margins used in raking-ratio estimation as specified in Section 6.2. The effective coverage rates compare base-weighted NSV estimates to corresponding estimates from the ACS.

It is also possible to examine how closely the NSV weighted estimates approximate other estimates available from the ACS.

As previously noted, most Veteran spouses are currently living with their spouses. Estimation and analysis of spouses married to but not living with a Veteran is more problematic. Spouse-absent group based on a single cell adjustment to the weights for married, spouse-present group. The adjustment stems from an assumption that the coverage of the rarer group of spouses is likely to be less than the net adjustment for married, spouse-present group.

As with Veterans, the second analytic approach for Veteran spouses could examine the separate phases of response. Because Veterans and spouses of Veterans are sampled in identical manner during the screening phase of the sampling, there is complete overlap in nonresponse at the screening phase for the two populations. Thus, the analysis for Veterans does not need to be repeated for Veteran spouses. Because they are sampled and respond separately for the extended questionnaire, however, Veteran spouses require a separate nonresponse analysis for the extended phase. This nonresponse analysis will be conducted by cross-tabulating extended response status by individual characteristics available from the frame (including responses to the screening questionnaire) and fitting multivariate models to the pattern of extended response.

Although the number of spouses married to but living apart from a Veteran may prove too few to support a meaningful nonresponse analysis, it might be possible to study the response of this specific group to the extended survey. Individuals in this category can be identified by the screening questionnaire. At a minimum, then, a nonresponse rate to the extended survey can be estimated for this group.

Because of resource constraints, the 75 completed extended questionnaires targeted for surviving spouses will not support a detailed statistical analysis of nonresponse. A credible source for the aggregate size of this group also appears unavailable. Nonetheless, it is appropriate to compute and report a standard response rate to the extended survey for this group, because surviving spouses will be identified by the screening questionnaire. The analysis would compute weighted response rates by subgroups, except the only likely interest would be in the group of surviving spouses as a whole.

Analysis of express letter subsample. One disadvantage of the above non-response bias analyses is that it does not provide outcome data for those who did not respond to the survey at all (e.g., awareness of VA benefits). In order to do the latter, it is necessary to collect data from non-respondents. The use of the express mailing for a subsample of 10,000 screener non-respondents provides one opportunity to do this. Tabulations of responses to the NSV for those getting the express package to those that got a USPS package will result in a comparison between two groups with different response rates. If there are significant differences between the two groups along outcome measures, there would be evidence of non-response bias.

This analysis is dependent on the success of the express package procedure producing a significantly higher coverage/response rate. To the extent this is true, the above analysis will provide complementary information to the bias analysis using auxiliary information. However, if effective coverage/response rates are not different or they are only marginally different, this analysis will not be as informative.

Methods for List Frame

The list frame consists of three surveys, including those for Active Duty Military, Demobilized National Guard/Reserves, and Spouse of Active Military (Attachments 17, 18, and 19). The procedures to be used for these surveys will also involve four different contacts. Because this sample is from a list, the mailing package will be addressed to each individual sampled. The first contact will be a letter that requests they fill out the survey by going to the web (Attachment 20). This letter will also include an insert encouraging respondents to go to the web to fill out the questionnaire. The second contact will be a reminder postcard referring respondents to go to the web to fill out the questionnaire (Attachment 21). In both the first and second contacts, sample members will be informed that if they cannot complete the survey over the Internet, a paper version of the survey will be sent in approximately 2 weeks.

Two weeks after the reminder postcard, a third request to complete the extended interview will be mailed. This request will include both a web address as well as a paper survey (Attachment 22 for letters). The fourth, and final, request will be sent 2 weeks after the second request (Attachment 23). This request will also include both a web address and a paper survey as well as an option to complete the survey over the telephone. Exhibit 2 illustrates the data collection procedures for the list sample.

Exhibit 1 - Address Sample Screening Procedures
Veterans, Spouses of Veterans, Surviving Spouses of Veterans

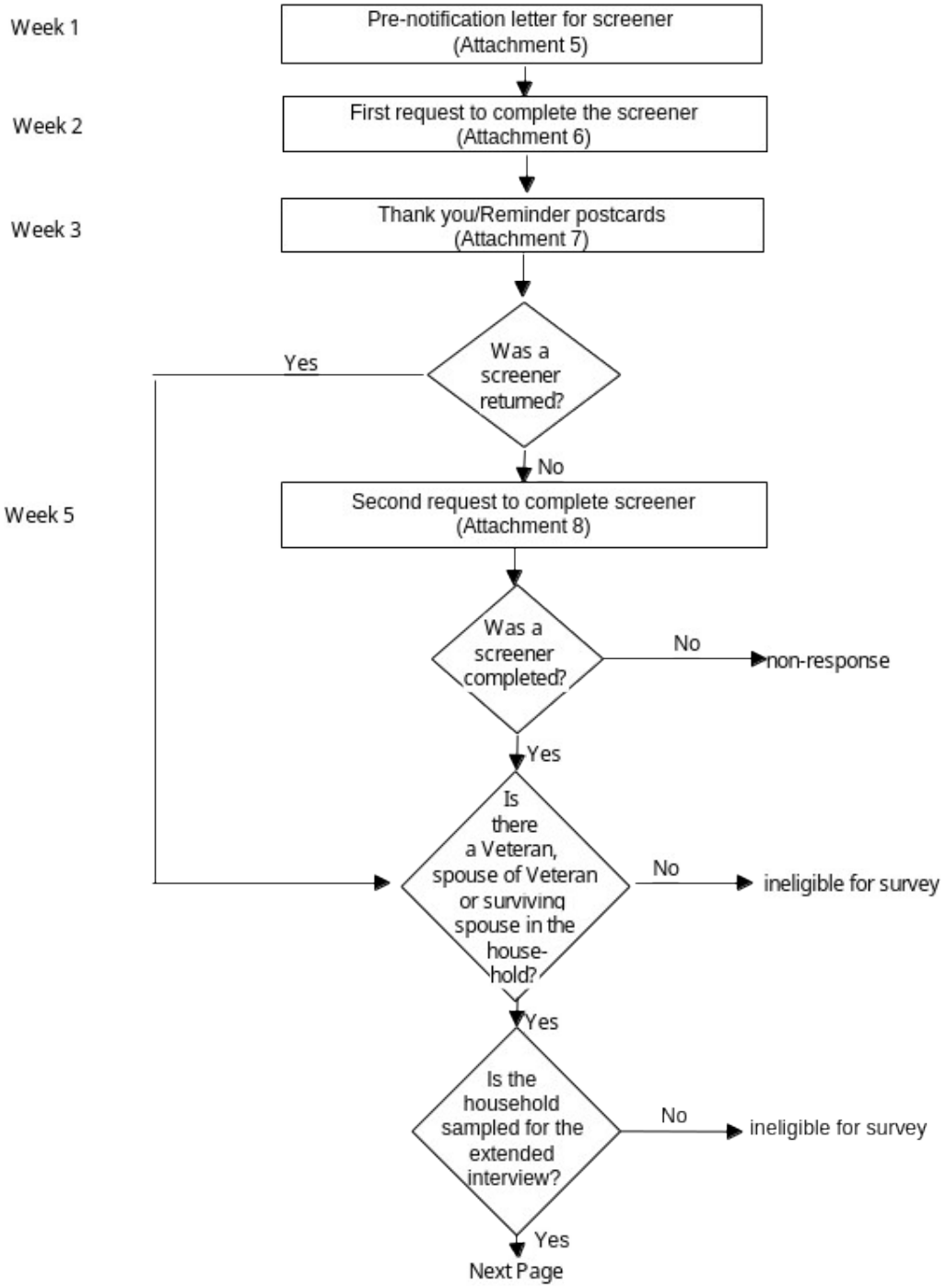


Exhibit 1-Address Sample Extended Survey Procedures (cnt'd)

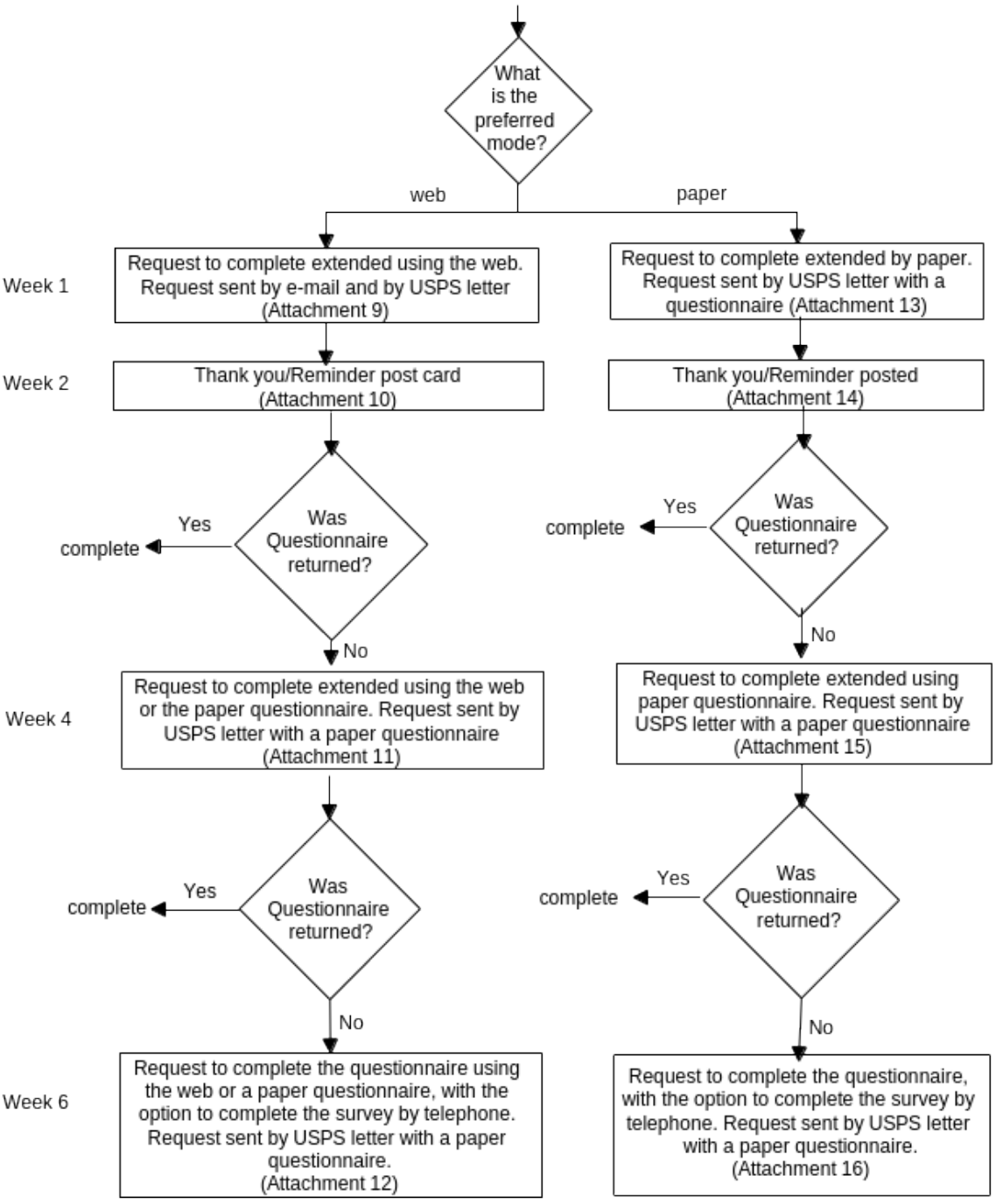
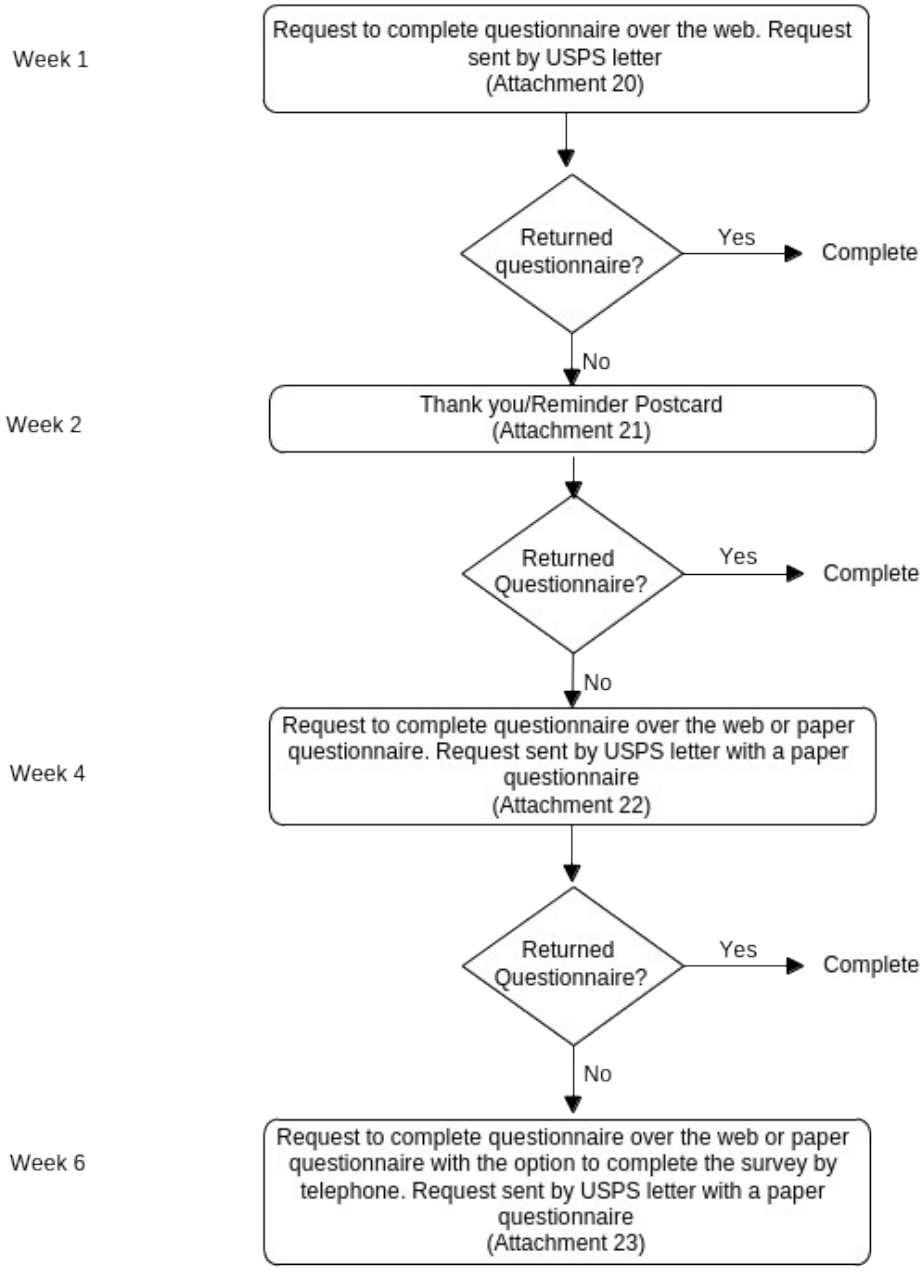


Exhibit 2 - List Sample Data Collection Procedures
Active Duty, Spouse of Active Duty, National Guard/Reserve



Additional Methods to Maximize Response Rate

As discussed above, the methodology proposed is based on principles that have proved to be effective in maximizing response rates for a mail survey. Prior to conducting the survey, we will conduct a series of pre-tests to refine the survey materials. This will also enhance the overall response rate and data quality (see question 4 below). In addition, a Pilot Study was conducted to assess the effectiveness of the two-stage approach to generating national estimates for the Veteran population. It answered several questions that informed the design of the survey of the address-sample populations: 1) What are the response rates when screening for Veterans, spouses of Veterans, and surviving spouses of Veterans? and 2) What are the best methods to implement the screener. The Pilot Study's results for the first question provided information on the sample sizes for the address sample. Results for the second question provided data on the best mode of interview to use (web vs. mail), the best method to measure Veteran status, and the effectiveness of the insert to increase response rates.

At the final request to complete the extended interview, respondents in both the list sample and address sample will be given the opportunity to complete by telephone. This relies on standard survey practice of offering an alternative mode of administration when following up with respondents who had not complied with prior requests to complete the survey (Groves, et al., 2004: 104).[1] Respondents will be able to call a toll-free number to schedule the telephone interview. Trained Westat interviewers will administer a screener to validate their eligibility and will schedule survey administration times that are convenient for them. Interviewers will administer the survey over the telephone but will enter responses via the web instrument and will flag the survey as a telephone survey in the survey database.

4. Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of tests may be submitted for approval separately or in combination with the main collection of information.

Results of the Pilot Study have already informed the basic methodology planned for the main survey including response rate assumptions, screener mode, and use of an insert in the survey package

In addition to the Pilot Study, a number of other pretesting activities will take place. First, we will implement a series of cognitive interviews for each of the six populations of interest. We will conduct 10 interviews for each group (60). These interviews will last approximately 1 hour each. The purpose of the interviews is to test

[1] Groves, R.M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E., and Tourangeau, R. (2004*). Survey Methodology.* Hoboken, NJ: Wiley.

the communication materials (letters) and the self-administered questionnaires. This will result in 60 hours of respondent burden (60 x 1).

Second, we will conduct pretest interviews by mailing the survey package to 26 Veterans and 50 spouses and surviving spouses identified during the Pilot Study but who were not sent the extended survey. After questionnaires are returned, the respondents will be called on the telephone to participate in a debriefing on issues related to filling out the instrument. They will also be asked for their reactions to the overall mailing package materials. This debriefing will last approximately 30 minutes. In combination with filling out the extended questionnaires, this results in approximately 1 hour of burden for each respondent. We anticipate that approximately half of the 76 individuals mailed an extended questionnaire will return it. Consequently, this pre-testing activity will result in 38 hours (76/2 = 38 x 1 hour) of respondent burden. All together, these testing activities will result in 98 hours (60 + 38) of respondent burden.

5. Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.

Table 6. Name and Telephone Number of Individuals Consulted on Statistical Aspects of the Design

Agency	Name Title	Phone Number	Statistical or Analytical
Westat	John Helmick NSV Project Director	301-294-2010	Analytical
Westat	Kimya Lee Senior Study Director	301-610-5522	Analytical
Westat	Wayne Hintze NSV Associate Project Director	301-517-4022	Analytical
Westat	Robert Fay Senior Statistician	240-314-2318	Statistical/Analytical
Westat	Richard Sigman Senior Statistician	240-453-2783	Statistical
Westat	David Cantor Senior Methodologist	301-294-2080	Analytical
Westat	Pamela Giambo Senior Study Director	240-453-2981	Analytical
Westat	Michele Harmon Senior Study Director	301-294-3814	Analytical
Westat	Marianne Winglee Senior Statistician	301-517-4169	Analytical
Westat	David Morganstein Vice President, Director of Statistical Staff	301-251-8215	Statistical
Westat	J. Michael Brick Vice President, Statistical Group	301-294-2004	Statistical
Westat	Wendy Hicks Senior Research Associate	301-251-2299	Statistical
Westat	Douglas Williams Research Associate	240-453-2934	Statistical
Westat	Brett McBride Research Analyst	301-517-8068	Statistical

List of Attachments

Address Sample

1. Address Sample: Screening Survey
2. Address Sample: Veteran Survey
3. Address Sample: Veteran Spouse Survey
4. Address Sample: Surviving Spouse Survey
5. Address Sample: Pre-notification letter for screener
6. Address Sample: Letter and insert for first attempt to complete the screener
7. Address Sample: Postcard reminder for the screener
8. Address Sample: Follow-up letter for screener
9. Address Sample: E-mails, letters, and insert for the first request for those asking for the survey by the Internet
10. Address Sample: Reminder postcard for extended interviews for those asking for the survey by the Internet
11. Address Sample: Second survey request for extended interviews for those asking for the survey by the Internet
12. Address Sample: Third survey request for extended interview for those asking for the survey by the Internet
13. Address Sample: Letter for first request for those asking for the survey by mail
14. Address Sample: Reminder postcard for extended interviews for those asking for the survey by mail
15. Address Sample: Second survey request for extended interviews for those asking for the survey by mail
16. Address Sample: Third survey request for extended interview for those asking for the survey by mail

List Sample

17. List Sample: Active Duty Service Member Survey
18. List Sample: Active Duty Spouse Survey
19. List Sample: Demobilized National Guard/Reserve Survey
20. List Sample: Letter and insert for first survey request for extended interview
21. List Sample: Reminder postcard for extended interview
22. List Sample: Letter for second survey request for extended interview
23. List Sample: Letter for third survey request for extended interview