B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

1.   Respondent Universe and Sampling Methods

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 VA plans to conduct a 20-year longitudinal study on three cohorts of Veterans and their dependents who are receiving educational benefits authorized by the Post-9/11 GI Bill. Eligible dependents include individuals receiving transferred benefits and those who receive benefits as a result of the Fry scholarship.   The dates of the three cohorts will be fiscal years 2010, 2012, and 2014. This request covers the first year of data collection for this Post-9/11 GI Bill Study, which will include the initial survey for Cohorts I and II. Therefore, the respondent universe for Cohort I includes all Veterans and dependents who receive educational benefits in FY 2010. The respondent universe for Cohort II includes all Veterans and dependents who receive educational benefits in FY 2012.

A sufficient number of Veterans and dependents will be recruited for the survey such that at the end of the 20-year study period, 1,190 respondents remain in each cohort sample. The cohort samples selected for the study are intended to represent the cohort populations receiving educational benefits during the respective fiscal years. Since fiscal year 2012 has not yet ended, the respondent universe for Cohort II has not yet been defined.

The number of Veterans and dependents who received educational benefits in FY 2010 is 335,170. Table 3 describes the univariate domains of interest (DOIs) that were used to develop sampling strata for selecting the random sample of Cohort I survey participants.

Table 3. Univariate domains of interest used to develop sampling strata for Cohort I

Tabulation variable	Value	Population size	Population prevalence	Smallest population prevalence
(none)	(entire population)	335,170		
Beneficiary size	Veteran / Active Duty*	283,487	85%	4%
	Child / Fry scholarship	37,955	11%	
	Spouse	13,728	4%	
Gender	Male	242,479	72%	28%
	Female	92,691	28%	
Age	< 30	159.522	48%	17%
	30-39	117,983	35%	
	40+	57,665	17%	

*Includes unknown.

In order to ensure that a statistically valid sample remains at the end of the 20-year study period, we anticipate completing 3,500 initial surveys for each cohort. This initial set of completed surveys will be the result of recruiting participants from a recruitment sample randomly selected from the universe for each cohort.

The recruitment sample will be organized into a main group and a reserve group which will be further partitioned into multiple release groups within each sampling stratum. We will only contact the reserve sub-sample cases if recruitment response rates are less than assumed. For Cohort I, a stratified recruitment sample of 30,000 Veterans and dependents will be randomly selected from the Cohort I universe and then randomly portioned into one initial sample of 10,000, with five additional reserve samples of 4,000 each. Table 4 contains the stratum sample sizes for the initial recruitment sample and the five associated reserve samples. The initial sample of 10,000 program participants will be released in the first wave of recruitment and the target recruitment response rate will be 35%. The additional samples will be

held in reserve and released as necessary if the targeted response rate is not achieved.

The recruitment response rate of 35% is based on recent VA experience in recruiting Veterans for the Congressionally-mandated Vocational Rehabilitation and Employment (VR&E) Longitudinal Study. Similar to the EDU Longitudinal Study, the aim of the VR&E Longitudinal Study was to recruit 3,500 Veterans for the FY 2010 cohort and 3,500 Veterans for the FY 2012 cohort. VA was successful in doing this. The participation rate for the FY 2010 cohort was 34 percent and that for the FY 2012 cohort was 26 percent.

Table 4.   Stratum sample sizes for recruitment sample

Stratum #	Beneficiary type	Gender	Age Group	Population size	Sample Size Main	Sample Size Reserve**	Over-Sample?
1	Spouse			13,728	1,452	581	Yes
2	Child /Fry scholarship			37,955	1862	745	Yes
3	Veteran / Active duty*	Male	30-39	86,834	2266	906	
4			< 30	95,745	1588	635	
5			40+	43,339	1150	460	
6		Female	30-39	42,224	768	307	Yes
7			< 30	23,373	594	238	
8			40+	9,982	320	128	Yes
Total					10,000	4,000	

 * Includes unknown
**Single reserve sample (out of 5)

2.   Information Collection Procedures

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.

One can use the data from a longitudinal survey to calculate estimates of level at individual time points or estimates of change between pairs of time points. For the survey variables that are being collected in the longitudinal survey of Chapter 33 benefits, there will generally be a high correlation between data collected at two different time points. Consequently, it is expected that the survey's estimates of change will be more precise than its estimates of level. As a result, to develop the proposed sample design we specify target precisions for the estimates of level, with the expectation that associated estimates of change will generally be more precise than the estimates of level.

The study will recruit three cohorts of Veterans and dependents who are receiving educational benefits in FY 2010, FY 2012, and FY 2014, respectively. In FY 2010, there were 335,170 Veterans and dependents who received educational benefits. Not all Veterans and dependents who are invited to participate in the study will agree to participate, nor will every participant complete each annual survey for the full 20 years. Hence, in addition to specifying target precisions, it is important to specify target response rates to ensure enough Veterans and dependents are recruited to result in a statistically valid sample with sufficient power.

Target rates. There are two target rates of interest: the recruitment response rate, denoted r_c , and the annual attrition rate, denoted r_a . Table 5a shows for various fielded sample sizes the recruitment response rates needed to obtain 3.500 completed surveys in the first year. Table 5b shows the number of responses after 20 years, which equals $3,500*(1-r_a)^{19}$. From

Table 5b, it follows that an initial sample of 3,500 Veterans and dependents should result in a final sample of at least 1,190 Veterans and dependents at the end of 20 years as long as the annual attrition rate is no more than 5.5%.

Table 5a. Recruitment rate needed to obtain 3,500 completed surveys in the first year

Table 5b.   Number of responses after 20 years

Table 5a

Fielded sample size	Needed response rate
10,000	35%
11,667	30%
13,400	25%
17,500	20%
23,333	15%

Table 5b

Annual attrition rate	# responses after 20 years
1.0%	2,892
2.0%	2,384
3.0%	1,962
4.0%	1,611
5.0%	1,321
5.5%	1,190

The reasons for attrition in the number of participants include deaths of some participants, unable to contact, and unwillingness to participate. An analysis based on SSA's life tables predicts that the annual attrition rate due to deaths of participants will be approximately 0.3 percent. It is difficult to predict the annual attrition rates due to other causes. In Table 5c, we compute the annual attrition rates from the total attrition rates for several longitudinal surveys, summarized in Watson and Wooden (2009)[1]. Since the proposed study will have multiple cohorts, if attrition rates are higher than expected then pooling across cohorts is one way to increase the precision of calculated estimates.

Table 5c. Annual attrition rates for selected longitudinal

Survey	# waves or years	Total attrition	Annual attrition
U.S. National Longitudinal Study of Youth - 1979 cohort	8	8%	1.0%

[1] Watson, w. and Wooden, M. (2009). Identifying Factors Affecting Longitudinal Survey Response, in Lynn, P. (Ed),   *Methodology of Longitudinal Surveys*, (pp. 287-302). Wiley.

U.S. National Longitudinal Study of Youth - 1979 cohort	21	20%	1.1%
U.S. National Longitudinal Study of Youth - 1997 cohort	5	12%	2.5%
Univ of Michigan Panel Sudy of Income Dynamics	8	25%	3.5%
European Community Household Panel (Portugal)	5	18%	3.9%
German Socio-Economic Panel	8	34%	5.1%
British Household Panel Survey	8	34%	5.1%
Dutch Socio-economic Panel	11	58%	7.6%
European Community Household Panel (Ireland)	5	43%	10.6%

As we recruit and enroll Veterans and dependents into the study, we will need to monitor the recruitment and enrollment process to ensure we have recruited a sufficient number of Veterans and dependents with specific characteristics that will allow us to compare sub-populations of interest at the end of the 20-year study period. For example, the agency may be interested in comparing the outcomes of male Veterans versus female Veterans. In order to make those comparisons, we need to be sure the sample includes a sufficient number of male and female Veterans.

Target precisions. A measure of the precision of an estimate is the half-width of the 95% confidence interval about the estimate, which is referred to as the estimate's margin of error (MOE). The relative size of MOEs by DOI varies inversely with DOI sample prevalence—that is, the proportion of the sample contained in the DOI. The DOI's with large sample-prevalence values will have small MOEs, and DOIs with small sample-prevalence values will have large MOEs.   In an unstratified sample in which response rates are the same across the DOIs, sample prevalence equals population prevalence. Thus, in an unstratified sample, a DOI with a small population prevalence may have a large MOE. Column 3 of Table 6 contains the MOEs for DOI estimates of p = 50% for an overall completed sample size of n = 1,190. For example, for Veterans and Servicemembers the expected MOE is 3.1% and for spouses the MOE is 14.0%. If the sample is unstratified and the response rates are the same for these two types of beneficiaries. By oversampling spouses, the MOE for spouses can be decreased, but this will increase the MOE for Veterans and Servicemembers. Column 4 of Table 6 contains the projected MOEs resulting from planned oversampling of spouse and

child/Fry-scholarship beneficiaries, females, and beneficiaries aged 30 or older.

Table 6. MOEs for DOI estimates of p = 50% for unstratified and stratified samples yielding 1,190 completed cases and equal response rates in each DOI

Tabulation variable	Value	Unstratified sample		Stratified sample	
		n	MOE	n	MOE
(none)	(entire population)	1,190	2.8%	1,190	3.1%
Beneficiary size	Veteran / Active Duty*	1,007	3.1%	796	3.5%
	Child / Fry scholarship	135	8.4%	222	6.6%
	Spouse	49	14.0%	173	7.4%
Gender	Male	861	3.3%	695	3.9%
	Female	329	5.4%	495	4.9%
Age	< 30	566	4.1%	512	4.9%
	30-39	419	4.8%	448	4.9%
	40+	205	6.8%	230	6.9%

*Includes unknown

3. Maximizing Response Rates

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.

It is becoming more difficult to conduct telephone surveys as people move from traditional landline telephones and are less willing to accept calls on their cell phones. There is increasing evidence that web and paper questionnaires, using accepted contact procedures (e.g., followup mailings; Dillman, et al., 2009) can achieve response rates as high (or higher) than a telephone survey.

A web questionnaire, by itself, will not provide the coverage that a paper-mail or telephone survey will provide. There are many households that do

not have ready access to the web and other that will not take a survey via the web. Westat's recent experience on the National Survey of Veterans found that only about 15 percent or 20 percent of Veterans were willing to complete the survey on the web (Cantor, et al., 2010). Consequently, it is imperative for this survey of Veterans and dependents who are receiving educational benefits to combine the use of the web with another mode, such as a telephone or paper survey, to maximize response rates and minimize respondent burden. Therefore, a multi-modal data collection approach will be used that combines web, mail, and telephone administration.

Literature on survey research suggests that response rates in fact decrease when respondents are presented with multiple administration options at the same time. Instead, a more effective strategy is to offer respondents one survey mode at a time, and then implement a different mode with non-responders. For the EDU Longitudinal Study, survey participants will be first asked to complete the survey online via the Internet. The initial survey invitations to complete the survey online will be mailed and emailed to sampled Veterans. An email reminder for completing the online survey will be sent to non-responders at the beginning of the 2nd and 3rd weeks of the data collection field period. The invitation letter will be printed on VBA letterhead and signed by the Director of the EDU program. A personalized, official letter of invitation from a high-ranking executive within the VA will make the study appear to be more credible to participants, and is likely to increase participation.

A paper-pencil survey will be mailed to those who do not respond via web. In addition to the survey itself, the mailed package will also include another personalized invitation letter, a pre-addressed postage-paid return envelope, and possibly a pen or pencil; and will be mailed to those who did not complete the web survey at the start of week 4 of the data collection field period. Weekly email reminders will be sent to nonrespondents with valid

email addresses. A similar data collection and contact protocol will be used for the annual followup surveys.

Because of the longitudinal design of the survey, it is important to maintain accurate contact information and keep participants engaged throughout the study period to gain their cooperation for completing the annual surveys. To achieve a high retention rate and minimize attrition, we propose to maintain contact with participants between data collection activities. Approximately 3 months after completing an initial or annual survey, we will email a note to all participants thanking them for their participation and reminding them to provide us with any updated contact information using the toll-free study number or the study-designated email address. In the next quarter, we will use various methods to send minimally burdensome reminders to all participants to send us updated contact information; including using automated methods such as email, interactive voice response, voicemail messages, and text messages. We will send the final reminder and request for updated contact information via mail in the quarter immediately preceding the next data collection effort.

High participation rates are not sufficient to rule out possible biases due to non-response. Therefore, response rates by strata will be examined and if needed, we will adjust survey weights to account for non-response.

For each cohort, the goal will be to recruit 3,500 Veterans to participate in the first year of data collection. Recruitment will begin for each cohort by fielding 10,000 invitations to participate. If the recruitment rate with respect to this initial fielding is at least 35%, the goal of recruiting 3,500 Veterans will be achieved. If necessary, however, additional recruitment samples will be fielded in order to yield 3,500 participants in the first year of data collection for each cohort.

If the recruitment rate with respect to all the fielded recruitment samples is less than 80%, a non-response bias study will be conducted. This study will use the VA administrative data available for Veterans and dependents who are receiving benefits authorized by the Post-9/11 GI Bill, including individuals receiving transferred benefits and those who are receiving benefits as a result of the Fry scholarship program. One way VA administrative data will be used is to compare the respondents with nonrespondents to the invitations to participate. Another way the administrative data will be used is to estimate the bias caused by nonresponse to the invitations to participate that is present in weighted estimates computed from the administrative data. It will be possible to estimate such biases because the VA administrative data is available for both respondents and nonrespondents to the invitations to participate.

4.   Testing of Instruments

Describe whether, and to what extent, the collection of information involves the use of automated, electronic, mechanical, or other technological collection techniques or other forms of information technology, e.g., permitting electronic submission of responses, and the basis for the decision for adopting their means of collection. Also describe any consideration if using information technology to reduce burden.

No testing of procedures or methods will be undertaken for completion of this survey. However, we did use our experience from the administration of the initial survey for the VR&E Longitudinal Study to make improvements to the proposed survey for the EDU Longitudinal Study. In the future, survey refinement will occur based on the results of conducting the initial survey to minimize burden and improve utility wherever possible.

5.   Individuals Consulted on Statistical Issues

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.

VR&E has contracted with Westat for conducting the survey and analyzing the data for the study.

Statistical Consultation:

        Richard Sigman
        Senior Statistician
        Westat
        1600 Research Boulevard
        Rockville, MD 20850
        (240) 453-2783

Data Collection and Analysis:

        Mustafa Karakus

Principal Investigator
Westat
1600 Research Boulevard
Rockville, MD 20850
(301) 294-2874