Public Needs for Library and Museum Survey (PNLMS)
Data Collection

Supporting Statement for PRA Submission

B	**Collection of Information Employing Statistical Methods**

B.1. Universe, Sample Design, and Estimation

The target population for the Public Needs for Library and Museum Services (PNLMS) Survey is the non-institutionalized population, aged 18 years and older, who live in the United States. A national probability sample of households generated using list-assisted random digit dialing (RDD) methodology will be employed by the survey. To provide randomization within the household, the adult in the household with the next birthday will be interviewed.

B.1.1. Universe and Sample Design

The universe for the survey includes all adults 18 years and older in the United States. A total of 3,500 interviews are planned, of which 2,975 will be conducted with participants over landlines and 525 interviews will be conducted with participants over cell phones. Random Digit Dialing (RDD) will be used to contact respondents for this survey. MDAC will interview respondents based upon their attendance at museums and libraries in the past 1 month. Respondents will be grouped into one of four categories: dual users, library-only users, museum-only users, and non-users. Library-only users will be administered the library module, museum-only users the museum module, and dual users will receive both the library and museum modules. The quantities of respondents who qualify for each module will provide an estimate of museum and library usage. There will be a separate module for the non-users. For households with children under the age 18 years, a child's module will be administered. If more than one child under 18 years resides in the household, the respondent will be asked to refer to the child with the next birthday to answer the questions.

The estimated count of phone numbers in the initial pool, from Marketing Systems Groups (MSG), is 167,370. IMLS anticipates a 20% response rate of completed interviews, equal to the response rates of similar federal collections. The predicted disposition of the calls is in Table B1. These numbers are only estimates. Once the field period has been completed, IMLS will have actual numbers for each cell, which will provide us with the information necessary to calculate an accurate response rate for the collection.

Table B1. Predicted Selection of Respondents within Households for the PNLMS

Forecasted Dispositions for PNLMS Landline Study Sample at 20% Response Rate		
Estimated Count of Total Phone Numbers Generated by Sample House (Marketing Systems Group - MSG)	167,370	
Estimated Count of Non-productive (i.e., on-working telephone numbers, business telephone numbers and fax/modems numbers) Phone Numbers Purged by MSG Process from Initial Sample Frame (between 50-55% - used mid-point of 52.5%). Note: This initial process does not purge 100% of the non-productive numbers that have been generated.	87,870	
Estimated Total Number of Phone Numbers Provided to MDAC for Dialing	79,514	100%
Predicted Results of Dialing:		
Not eligible Numbers	37,957	47.7%
Fax/data line	7664	9.6%
Non-working/disconnect	21664	27.2%
Number changed	312	0.4%
Business, government office, other organizations	8239	10.4%
No eligible respondent	78	0.1%
Estimated Remaining Sample	41,557	52.3%
Unknown eligibility, non-interview	32,827	41.3%
Always busy	1444	1.8%
No answer	11654	14.7%
Answering machine-don't know if household	5482	6.9%
Call blocking	33	0.0%
Technical phone problems	230	0.3%
Housing unit, unknown if eligible respondent	9914	12.5%
No screener completed	2593	3.3%
Other	1477	1.9%
Estimated Remaining Sample	8,730	11.0%
Eligible, non-interview	5,685	7.1%

Refusal and breakoff	**1,684**	**2.1%**
Refusal	**1,263**	**1.6%**
Known-respondent refusal	**421**	**0.5%**
Non-contact	**3,366**	**4.2%**
Respondent never available	**1,122**	**1.4%**
Answering machine household	**2,244**	**2.8%**
Other, non-refusals	**635**	**0.8%**
Physically or mentally unable/incompetent	**157**	**0.2%**
Miscellaneous	**478**	**0.6%**
Estimated Remaining Sample	**3,045**	**3.8%**
Interviews	**3,045**	**3.8%**
Partial	**70**	**0.1%**
Complete	**2975**	**3.7%**

Note: These estimated results are for example purposes. Due to percentage rounding, individual totals may not be exact.

B.1.2. Precision Requirements and Sample Sizes

The incidence of visitation/usage for museums and libraries will be determined by respondent's answers to the survey questions. Estimates of sample sizes and corresponding margins of error are illustrated in the table shown below.

This analysis assumes a random sample of 3,500 completed interviews with no group quotas and the population distribution shown in the upper left table. The expectation is that this will result in the group sample sizes shown in the lower left table. The corresponding margins of error, with 95% confidence, are shown in the lower right table. The original target margins of error are in the upper right table for reference.

95%	Confidence Level
3500	Completed Interviews

Population Percentages		Library		Row Sums
		Yes	No	
Museum	Yes	42%	18%	60%
	No	20%	20%	40%
Column Sums		62%	38%	100%

Target MEs		Library		Row Sums
		Yes	No	
Museum	Yes	3.0%	4.0%	3.0%
	No	4.0%	5.0%	
Column Sums		3.0%		

Sample Sizes		Library		Row Sums
		Yes	No	
Museum	Yes	1470	630	2100
	No	700	700	1400
Column Sums		2170	1330	3500

Actual MEs		Library		Row Sums
		Yes	No	
Museum	Yes	2.6%	3.9%	2.1%
	No	3.7%	3.7%	
Column Sums		2.1%		

B.2. Procedures for the Collection of Information

After OMB approval is received, Morris Davis and Company (MDAC), the company contracted by IMLS to conduct the PNLMS, will begin field operations. MDAC will program a CATI survey instrument to administer the questionnaire. Before respondents are called, MDAC will mail an advance letter describing the survey to all households with an address match on file. Within five days of the anticipated receipt of the advance letter, MDAC will make the initial call. MDAC will interview respondents as they are contacted. MDAC will interview respondents who have either or both a museum or a library within the past month and respondents who have visited neither within the last year. Data collection in the field will be terminated when 3,500 interviews are completed.

B.2.1. Data Collection

Programming and Pretesting the CATI
MDAC will develop a draft CATI instrument using the final version of the survey questionnaire approved by OMB for formal pretesting. This draft survey instrument will include formal interviewer instructions, section breaks, help-screens, appropriate branching for survey questions (skip patterns), with built-in edit, range and consistency checks. Only a few questions are open-ended responses. MDAC will provide IMLS with access to the draft CATI program for testing. The testing phase will include the importation of simulation data to test all skip patterns and edit checks.

MDAC will conduct seven (7) to nine (9) telephone pretests of the draft CATI survey instrument with each user type (e.g. dual (library/museum), single (library or museum) and non-user of the draft CATI survey instrument). The pretest will be used to replicate the data collection process and to identify any problem areas related to the process, the survey instrument in total, specific questions, answer choices, questionnaire instructions or question format. All facets of data collection, including questionnaire flow, CATI systems, sample management, quality assurance, and data processing will be tested by MDAC telephone supervisors. Pretests will also be used to confirm the average time to complete interviews. Households used for the pretest will be representative of the intended target population of the PNLMS. MDAC will compile the results from the telephone pretest in order to make recommended changes to address any problems with the draft survey instrument. MDAC will allow IMLS to listen in while the pretests are being conducted.

Interviewer Training
All interviewers will undergo intensive training and orientation regardless of their level of experience prior to being hired. Detailed briefing sessions are conducted prior to the start of every study and audio-taped for follow-up and review. In response to normal interviewer turnover and/or increased staffing needs, all interviewers new to the project will receive the full complement of training prior to beginning their interviewing for this survey program.

Before a telephone survey is fielded, MDAC conducts four (4) to six (6) hour project-specific classroom training for each project that includes:

Background and purpose of the survey	Successful completion of a total of eight (8) hours of training spread across multiple sessions
Use of correct respondent selection procedures	A question-by-question review of the hard copy and CATI versions of the questionnaire
Intensive hands-on training on the basics of interviewing—skip patterns, probing and clarifying techniques, sample administration, use of Computer Assisted Telephone Interviewing (CATI) technology, avoiding refusals, etc.	Question-by-question training of the questionnaire
Observing and listening to experienced interviewers	Feedback from interviewers, supervisors;
Conducting "real" interviews, during which each trainee's performance is closely monitored and evaluated under actual interviewing conditions	Interviewer/trainer role play
Repeated reference to the importance of accuracy, quality and courtesy	Problems and issues emerging from previous waves of data collection

Once the telephone survey fielding period begins, briefing sessions are continued on a daily basis at the start of each interviewing shift to address issues that arise during the first days of fielding. These sessions are repeated on an as-needed basis, through completion of the field period. This process ensures adherence to survey program specifications and procedures from start to finish, and enhances work performance and efficiency.

MDAC will establish a specific refusal-conversion team comprised of supervisors and lead interviewers who have been with MDAC for at least 3 years. In addition, all interviewers will be trained in effective practices of refusal avoidance and "holding" respondents in mid-survey who threaten to terminate the interview. This will increase survey response rates and the quality of the data.

Contacting Respondents
Initially, a designated amount of phone numbers (total estimate = 167,370 – see Table B1) will be released to the call center. MDAC will use techniques to facilitate the highest possible

response rate using the numbers released and emphasizing reaching in-scope households, where i*n-scope* is defined to mean numbers where contact has been achieved and eligibility determined. Sample numbers will be added based upon past calling history, the quantity of numbers determined to be ineligible, and projection of completes based upon past and current experience, number of callbacks achieved, and refusal conversion rates.

IMLS has provided the text and format of an Advance Letter for the collection. The advance letter will be sent out prior to data collection to all known addresses within the sample. MDAC will obtain a residential name and address for each sample phone number by conducting a "reverse match" operation. MDAC will utilize Genesys for the address matching. Genesys employs an address match process which utilizes four residential databases—Acxiom, Experian, InfoUSA and Targus. The benefit of using four databases is a higher match rate. This service is called "Enhanced Match" because of the number of databases utilized. Each database includes close to 200 million or more records. Generally, the enhanced match will achieve a 50% to 60% match rate.

MDAC will mail the Advance Letters so they arrive one (1) to five (5) days before the first call attempt. MDAC will develop a tracking system for the Advance Letter that will allow IMLS to match any returned undeliverable letter with the original mailing list.

MDAC will establish a toll-free telephone number that will be used to arrange callbacks with respondents and conduct interviews when respondents call the contractor. The toll-free telephone number will also be included in the Advance Letter. The toll-free telephone number will be listed in each letter with the specific times to call (Monday through Friday 9:00 am – 11:00 pm EST; Sat and Sun, 11:00 am to 11:00 pm EST).   Should the respondent call outside these times, they will receive a phone message asking them to leave their name and number and someone will contact them as soon as possible to conduct the interview.

When a phone number is called initially, the interviewer will determine whether it is a household. Then the interviewer will request to speak with an adult 18 years of age or older (if the person on the phone is not an adult). Once an adult is on the line, the interviewer will randomly select the actual survey respondent by asking for the adult in the household who has the next upcoming birthday. When the adult with the upcoming birthday is on the phone, the interviewer will conduct the survey. Should the interviewer not be able to complete the survey, the disposition will be recorded.

Cell Phone Sub-sample
MDAC is experienced in the administration of cell phone surveys. With regard to within-unit selection in the cell phone survey, MDAC, as most other research organizations, uses the approach that in practice cell phones in the U.S. are considered a personal device. Cell phone surveys require special procedures that will be implemented for this survey:
- The person answering the call will be immediately asked whether they have been contacted on a cell phone, and if so, if it is convenient and safe to continue talking, or better to schedule a callback. Interviewers will be instructed to avoid interviews in situations which could prove dangerous to the respondent or others;

- Respondents reached on a cell phone will be offered a $5.00 incentive for completing the interview (a customary practice to compensate those who are charged according to minutes of usage); this will require adding survey questions to obtain the respondent's name and mailing address at the end of the interview to respondents contacted on cell phone.;
- Consistent with the way interviews with persons on cell phones are usually handled, we will not need to ask respondent selection questions—the underlying assumption being that cell phone users do not share their cell phone with other household members, because it's considered a personal communications device;
- Because of the portability of cell phone numbers, we will need to obtain the location of the place of residence—the respondent's zip code, or, failing that, the city and state; this information will be used in weighting and analysis;
- In order to appropriately blend interviews from the landline RDD and the cell phone samples via weighting, we will have to determine whether the cell phone respondent could also have been reached on a home landline phone; and, if so, as in *all* interviews, how many landline phones in the home are used for non-business voice conversations.

Techniques to Enhance Response

During the data collection period, MDAC will attempt at least ten (10) call attempts for non-contacts and ten (10) callbacks for arranging interviewing appointments with respondents. If MDAC is unable to make contact with a member of the household, the toll-free telephone number will be mentioned at the seventh attempt in messages left for potential respondents that have answering machines. No message will be left on answering machines before the seventh phone attempt; this is due to the concern that people might avoid the call or feel harassed if they were away for a few days and find four to six messages on their answering machine upon returning home. Given that a household with an answering machine might be called two to three times per day during the PNLMS survey, there must be a balance between perceived harassment and encouraging participation.

Callbacks will be scheduled and prioritized by the CATI software. The callbacks will be prioritized based upon the following criteria: first priority--scheduled callback to qualified household member; second priority--scheduled callback to "qualify" household (includes contact with Spanish language barrier households); third priority--callback to make initial contact with household (includes answering machine, busy, ring no answer); and fourth priority--callbacks that are the seventh or higher attempts to schedule interview.

An interview will be considered "complete" only if the respondent was cooperative and provided reliable and valid answers for at least 80% of the PNLMS questions. Should the interviewer not be able to complete the interview the following procedures will be followed:

Scheduled callbacks can be dialed at any time during calling hours and as frequently as requested by the callback household up to seven (7) times. Callback attempts in excess of seven are at the discretion of the interviewer, based upon his/her perception of the likelihood of completing the interview. The basis of the interviewer's perception, in part, will be determined by how vigorously the interviewer is being encouraged to call back to complete the interview by the

potential respondent or another member of the household. The interviewer will then confer with a supervisor and a final determination will be made as to whether the interviewer will continue calling.

Callbacks to Spanish language barrier households will be conducted by Spanish-speaking interviewers. Non Spanish-speaking interviewers who reach a Spanish-speaking household will schedule a callback that will be routed to a Spanish-speaking interviewer.

Callbacks for initial contact with potential respondents will be distributed across the various calling time periods and weekday/weekend to ensure that a callback is initiated during each time period each day. Two (Saturday and Sunday) to three (Monday through Friday) callbacks per number will be initiated per day assuming the number retains a callback status during the calling. There will be up to ten (10) callback attempts. This protocol is designed for ring no answer and answering machines. When an interviewer reaches a household with an answering machine on the seventh call, the interviewer will leave a message with toll-free telephone number.

Callbacks to numbers with a busy signal will be scheduled every 30 minutes until the household is reached, disposition is modified, maximum callbacks are achieved or the survey program is completed.

For landline telephone numbers, MDAC will use predictive dialing technology and automated number recognition (ANI) to improve the productivity and data quality. (Note: it is illegal to use predictive dialing technology with known cell phone numbers, so this applies only to the landline sample.) Predictive dialing systems enhance productivity by dialing numbers; recognizing busy signals, answering machines, non-working numbers, and ring no-answers results; and communicating the dispositions of those calls to the CATI system without expending interviewer time. By relieving interviewers of these unproductive dials, productivity can increase by up to 50 percent, delivering cost savings to the government. ANI helps mitigate the Caller ID screening that has resulted in a decline in response rates for telephone surveys. Potential respondents are typically less likely to answer a call when the Caller ID display doesn't identify the incoming number. To increase the likelihood of a respondent answering the call, MDAC's dialer system will work with the local telephone system to provide an actual name or telephone number on the Caller ID display.

English-Spanish Language Telephone Survey Administration
According to a 2010 AAPOR Cell Phone Task Force report, when conducting surveys only in English, a proportion of contacts may be deemed out of scope (ineligible) due to a language barrier. In 11 national Pew surveys, an average of 6 percent of their cell phone contacts did not speak English versus 4 percent for the landline sample. In the one survey that was conducted in both English and Spanish (the 2009 Religion & Public Life Survey) only 0.5 percent of cell phone contacts and 2.3 percent of cooperating cell phone numbers were dispositioned as Language Barrier, about the same proportions as for the landline sample. This difference is primarily a function of the disproportionate number of minorities among the U.S. wireless-only and wireless-mostly populations – a difference that is also a factor of survey geography. The NCHS report for the second half of 2009 estimated that 30.4% of Hispanic adults and 20.6% of

Asian adults are wireless-only and that 16.9 percent of Hispanic adults and 18.5% of Asian adults are wireless-mostly. In the 2007 California Health Interview Survey (CHIS), which was conducted only in English in California (which has a high incidence of Hispanics), 8% of the respondents contacted by cell phone did not speak English. Consequently, language barrier rates can be reduced if a survey, including both landline and cell phones, is conducted in Spanish as well as English.

To administer IMLS' Public Needs Library and Museum Survey in Spanish, MDAC will translate verbatim the original questionnaire to Spanish using a team consisting of a bilingual translator (English and Spanish) and a bilingual peer reviewer to perform a quality control review and revision of the survey, as needed. The Spanish questionnaire will be programmed into its CATI version. This version of the survey will be used by bilingual interviewers to minimize language barriers. At the start of each shift, bilingual interviewers log on to a special account which enables them to receive phone numbers for Spanish speaking households through the CATI system's call handling system. If an English-language only interviewer encounters a language barrier other than Spanish, either with the person answering the phone or with the designated respondent, the interviewer will thank the person and terminate the call. However, if a called household requires a Spanish language interviewer, the household is assigned to a special file accessible only to Spanish speaking bilingual interviewers. The phone number is placed in the call handling queue so that the first bilingual interviewer to access a new sample record is forwarded the record for the Spanish speaking household. This results in callbacks by Spanish speaking interviewers being made almost immediately after the initial call. If the call to the Spanish speaking household is not completed on the call from a bilingual interviewer and requires additional calls to be made at a later time, the call record remains in the special file and can be accessed by a bilingual interviewer throughout the field period or until the record reaches a final disposition. MDAC ensures that there are a sufficient number of bilingual interviewers staffed on each shift to handle call volume for Spanish speaking households. MDAC also maintains bilingual supervisors on staff who are involved in training the bilingual interviewers, monitor Spanish language calls for quality assurance, and assist with issues pertinent to Spanish language administration of the questionnaire.

B.2.2. Sample Weight Development

Overview of Sample Weight Development

Two types of sample weighting and adjustments will be used for the survey data: 1) inverse-probability weights to correct for unequal selection probabilities as a result of various factors (e.g., non-response, multiple telephone lines, and persons-per-household; and 2) post-stratification weights to correct for known discrepancies between the sample and the population, based on various demographic characteristics. The final analysis weight will reflect both types of adjustments and will be the weight that MDAC uses when analyzing survey data.

The final analysis weight can be developed by making adjustments for:
• Base sampling weights;
• Unit non-response;
• Households with multiple voice telephone numbers;

- Cell phone only households;
- Post-stratification adjustments to the target population, such as the expected overrepresentation of females.

The product of the above adjustments is the final analysis weight. A final analysis weight is calculated for each survey respondent and applied to all of the respondent's survey responses. Below, we present a process for determining the final analysis weight for each respondent.

Base Sampling Weights

The first step in determining the final analysis weight is to calculate the base sampling weight for each telephone number in both the landline and cell phone only (CPO) samples. The sampling rate is the inverse of the telephone number's probability of selection, or:

$$W_S = \frac{N}{n}$$

When N represents the total number of telephone numbers in the population and n represents the total number of telephone numbers in the sample.

Adjustment for Unit Non-Response

After calculating base sampling weights, sampled telephone numbers can be classified as responding or non-responding households according to Census division and MSA status (i.e., inside or outside a Metropolitan Statistical Area). The non-response adjustment factor for each telephone number can be calculated as follows:

$$ADJ_{NR} = \frac{1}{CASRO\ response\ rates}$$

The denominator represents the Council of American Survey Research Organizations (CASRO) response rate for each Census division (c) and MSA status (s). The response rate for a specific Census division-MSA combination is given by the ratio of the estimated number of telephone households included in that cell to the number of completed surveys in that cell.

The non-response adjusted weight (W_{NR}) is the product of the sampling weight (W_S) and the non-response adjustment factor (ADJ_{NR}) within each Census division - MSA combination.

Adjustment for Households with Multiple Voice Telephone Numbers

Some households have multiple voice telephone numbers, including landlines and cell phones, and thus have multiple chances of being selected into the sample. The sample weights will be adjusted to account for these households. The adjustment for multiple voice telephone numbers can be calculated as follows:

$$ADJ_{MT} = \frac{1}{Min\ (Nb\ telephone\ lines, 3)}$$

As shown in the formula above, the multiple voice telephone number adjustment is limited to a maximum of three telephone numbers. In other words, the adjustment factor ADJ_{MT} is one over one (1.00) if the household has one telephone number, one over two (0.50) if it has two telephone numbers and one over three (0.33) if it has three or more telephone numbers.

MDAC will collect information about multiple voice telephone numbers from each respondent as part of the survey. We will provide summary statistics regarding the number of telephone lines among sampled households (e.g., mean and standard deviation). For respondents who do not provide that information, MDAC will infer the number of landlines in the household and cell phone ownership based on eligible respondents who do answer this question.

The weight that is adjusted for non-response and for households with multiple voice telephone numbers (W_{NRMT}) is the product of the non-response adjusted weight (W_{NR}) and the adjustment factor for households with multiple telephone numbers (ADJ_{MT}).

Adjustment for Number of Eligible Household Members

The probability of selecting a survey respondent depends in part on the number of eligible respondents in a particular household. Thus, it is important to account for the total number of eligible household members when constructing the final analysis weights. The adjustment for the number of eligible household members can be calculated as follows:

ADJ_{EM} = *Number of Eligible Household Members*

MDAC will collect information about eligible household members from each respondent as part of the survey. We will provide summary statistics regarding the number of eligible household members (e.g., mean and standard deviation). For respondents who do not provide that information, a value for ADJ_{EM} will be imputed based on the distribution of the number of eligible household members from responding households within the corresponding age, gender, race/ethnicity, education, and income samples.

The weight that is adjusted for non-response, for households with multiple voice telephone numbers, and for the number of eligible household members (W_{NRMTEM}) is the product of the adjusted weight for non-response and households with multiple voice telephone (W_{NRMT}) and the adjustment factor for the number of eligible household members (ADJ_{EM}).

Adjustment for Cell Phone Only Households

Survey response rates may differ between the cell phone only (CPO) and landline samples. To account for that possibility, MDAC can apply an additional adjustment to each sample. The following adjustment will be applied to each telephone number in the landline sample:

$$ADJ_L = \frac{p_{LP}}{p_{LS}}$$

The numerator (p_{LP}) is the proportion of households in the population with a landline, and the denominator (p_{LS}) is the proportion of telephone numbers in the total sample (i.e., landline and CPO samples considered together) with a landline.

The corresponding formula for the CPO sample is as follows:

$$ADJ_{CPO} = \frac{p_{CP}}{p_{CS}}$$

The numerator (p_{CP}) is the proportion of households in the population with a cell phone only, and the denominator (p_{CS}) is the proportion of telephone numbers in the total sample (i.e., landline and CPO samples considered together) with a cell phone only.

Estimates of the population proportions of landline households CPO households will be taken from current U.S. Census Bureau or Center for Disease Control and Prevention (CDC) statistics.

To adjust for differences in response rates for landlines and CPOs, the adjusted weight for non-response, for households with multiple voice telephone numbers, and for the number of eligible household members (W_{NRMTEM}) is multiplied by either the adjustment factor for landline households (ADJ_L, resulting in $W_{NRMTEML}$) or by the adjustment factor for CPO households (ADJ_{CPO}, resulting in $W_{NRMTEMCPO}$), depending on the sample from which a particular telephone number originated.

Post-Stratification Adjustments

Adjusting survey weights to correspond with demographic population counts provided by the Census Bureau can account for the varying response rates of different demographic subgroups. Doing so will: 1) increase the precision of survey estimates; and 2) reduce potential biases in the survey data that may result from the inclusion of only those households with a telephone number. To account for demographic population counts, MDAC will use post-stratification adjustments that will result in final analysis weights that sum to the target population (U.S. non-institutionalized adults 18 years of age or older) by age, gender, and, education.

The outcome of post-stratification adjustments are multipliers (M) that scale $W_{NRMTEML}$ and $W_{NRMTEMCPO}$ within each demographic cell, so that the weighted marginal sums for age, gender, and education correspond with demographic population counts provided by the Census Bureau. MDAC will calculate post-stratification adjustments using simple ratios based on the appropriate national population totals for given cells that are defined by the intersection of age, gender, and education characteristics. [1] The ratios will then be applied to the sample weights corresponding to each demographic cell.

The general method for calculating the ratios is as follows:

- MDAC will create a table of the sum of sample weights for each cell defined by the intersection of age, gender, and education. There will be six levels for age, two levels for gender, and four levels for education. The combination of those levels results in a total of 48 cells (i.e., 6 x 2 x 4). [2] Each cell is denoted by S_{ijk}, where i is the indicator for age, j is the indicator for gender, and k is the indicator for education.
- MDAC will also create an analogous table of national population controls, where each cell will be denoted by P_{ijk}, where the subscripts are as above.
- The ratio $R_{ijk} = P_{ijk} / S_{ijk}$ is calculated for each cell.
- Each respondent's weight is multiplied by the appropriate cell ratio of R_{ijk} (which serves as the

---

[1] The Census Bureau provides a detailed breakdown of population counts by age, gender, and education.

[2] Some categories may be merged, if the number of completed interviews within the corresponding cells falls below 30.

M), resulting in a final analysis weight, W_{FINAL}. The final analysis weight is the number of population members that each respondent represents.

MDAC will collect demographic information about each respondent as part of the survey. We will provide summary statistics regarding respondent demographic characteristics (e.g., mean and standard deviation). MDAC will exclude those respondents who do not provide necessary demographic information from the post-stratification adjustment process (i.e., we will assign them an M value of 1).

Variance Estimation Methodology

The survey data is obtained through a complex sample design involving stratification and weighting, and the final analysis weights are subject to several adjustments. The methodology that MDAC uses to estimate the variance must involve some simplifying assumptions about the sample design. Some simplified conceptual sample design structures are provided in the sections below.

Design Information for Variance Estimation

MDAC will base the survey sample on a national probability sample of households using a list-assisted random digit dialing (RDD) methodology. We will initiate the development of the sample by first imposing an implicit stratification based on telephone prefixes from Census divisions and metropolitan status. Within each Census division, samples from counties and their associated telephone prefix areas will be ordered by the size of their corresponding MSAs. Counties and associated telephone prefix areas that are not located within an existing MSA will be ordered by state. Within each state, the counties and their associated prefix areas will be ordered by geographic location.

Software

A software package such as SAS® or Stata® will be used for computing standard errors. These packages are commonly used to analyze data from complex survey samples and used to perform econometric analyses. They are capable of incorporating complex probability-based weights and stratification variables into variance estimates.

Method

MDAC will use three variables—DIVISION (Census Division), METRO (metropolitan status), and FINALWGT (final analysis weights)—to estimate sample variance. We will begin by using the intersection of the variables DIVISION (nine levels) and METRO (two levels) to create an 18-strata matrix (i.e., 9 x 2). We will then use a single stage selection procedure and the variable FINALWGT to estimate the sample variance. The method will provide somewhat conservative variance estimates.

The first step will be to create a variable, STRATA, from the intersection of the DIVISION and METRO variables such that each unique combination of DIVISION and METRO will be identified by a unique integer. After creating the STRATA, we will report the means and variances of the appropriate variables.   When sampling weights are post-stratified, the variance

of an estimate is reduced because the totals associated with each demographic characteristic are known without sampling variation.[3]

Degrees of Freedom of Precision

A rule-of-thumb for degrees of freedom associated with a standard error estimate is the quantity:

Number of Unweighted Records in the Dataset – Number of Strata

For practical purposes, any degrees of freedom exceeding 120 will be treated as infinite (i.e., one would use a normal *z*-statistic instead of a *t*-statistic for statistical significance testing). Note that a one-tailed critical *t* at 120 degrees of freedom is 1.98, whereas a one-tailed critical *t* at infinite degrees of freedom (a *z*-value of 0.025) is 1.96. If a variable of interest covers most of the sample strata, this limiting value will probably be adequate for analysis.

Alternate Variance Estimates

MDAC will derive standard error estimates using the stratifying variables DIVISION and METRO. Alternative variance estimates may be calculated if necessary using a replicate procedure. Possible reasons for using an alternate variance estimate include a small sample size for a given cross-section of the dataset; or a desire to provide additional privacy protection when releasing a dataset to the public.

All the variables that MDAC uses for estimating standard errors will be provided in public- and internal-use data files.

B.3. Methods to Secure Cooperation, Maximize Response Rates, and Deal with Non-Response

B.3.1. Gaining Cooperation

Establishing contact and gaining cooperation is the first step in any interview. The rigorous general training of call center interviewers and the specific training for this study will enable interviewers to make a good first impression by sounding knowledgeable, professional, confident, and courteous. Here are three examples:

- Interviewers will be trained to respond to questions and depart from the script if necessary, but then to resume the script as quickly as possible.

- Interviewers will be trained to handle objections raised by respondents. For example, if a respondent says that s/he is too busy, the response might be, "If you are too busy now, I can call back at a better time. When would be a good time?"

---

[3] For a discussion of the impact of post-stratification on the variance of survey estimates see, "Sampling and Weighting in the National Assessment", Keith F. Rust and Eugene G. Johnson, Journal of Educational Statistics, 17(2): 111-129, Summer 1992

- Respondents who initially seem reluctant to complete the interview may only need a little persuasion. Interviewers will emphasize the study's importance to museums and libraries and stress that all answers will remain anonymous.

B.3.2. Methods to Maximize Response Rates

MDAC will use a response rate calculation based on guidelines established by the Council of American Survey Research Organizations (CASRO). MDAC will calculate the final response rate for the sample using the following formula:

$$Response\ Rate = \frac{Completed\ HH\ Interviews}{\left\{ HHs\ In\ Scope + \left[Scope\ Undetermined * \frac{HHs\ In\ Scope}{HHs\ In\ \&\ Out\ of\ Scope} \right] \right\}}$$

MDAC will use the table on the following page to present the distribution of household telephone numbers by disposition categories. We will then use the number of household cases in each disposition category in the formula above to calculate an overall response rate.

Maximizing the overall Response Rate
MDAC will take the following measures to maximize the overall response rate for the survey. We will:
1. Match sample telephone numbers against commercial files and against residential directory-listed numbers.
2. Send advance letters to potential survey participants stating clearly the aims, objectives and importance of the survey. We will include a toll-free number in the letter that potential survey participants can call back for information. MDAC will work closely with IMLS to write an IMLS-approved advance letter.
3. Carefully coordinate the mailing of advance letters with survey calling.
4. Provide interviewer calling hours of Monday through Friday 9:00 am – 11:00 pm EST and Sat and Sun, 11:00 am to 11:00 pm EST; however, we might adjust the calling times based on fielding data to maximize productivity and lower cost.
5. Develop answers for questions and objections that may arise during survey calls.
6. Leave messages on answering machines with a toll-free number that potential survey participants can call back.
7. Have bilingual interviewers to minimize language barriers.
8. Remove non-residential telephone numbers from the survey sample.
9. Attempt a minimum of 10 callbacks of respondents who initially refused or broke-off survey.
10. Minimize the turnover of personnel during the project.
11. Submit a weekly production report to the IMLS PO, along with a staffing plan for the next weeks interview scheduling and shifts.

In addition, MDAC will use predictive dialing technology and automated number recognition (ANI) to improve the productivity and data quality. Predictive dialing systems enhance productivity by dialing numbers; recognizing busy signals, answering machines, non-working

numbers, and ring no-answers results; and communicating the dispositions of those calls to the CATI system without expending interviewer time. By relieving interviewers of these unproductive dialings, productivity can increase by up to 50 percent, delivering cost savings to the client. ANI helps mitigate the Caller ID screening that has resulted in a decline in response rates for telephone surveys. Potential respondents are typically less likely to answer a call when the Caller ID display doesn't identify the incoming number. To increase the likelihood of a respondent answering the call, MDAC's dialer system can work with the local telephone system to provide an actual name or telephone number on the Caller ID display. Alternatively, the number displayed could be an IMLS telephone number that has been designated for this purpose.

We believe that the above measures will allow MDAC to obtain a 50 percent response rate for the base sample using the CASRO calculation presented above. *MDAC has achieved response rates of 50% or above for two Federal Government Telephone Survey Projects: Bureau of Transportation Statistics (BTS) Omnibus Household Survey and the HUD Fair-Market-Rent Telephone Survey.*

B.3.3. Statistical Approaches to Non-response

The effects of nonresponse bias are a major concern in most surveys of the general population regardless of the mode used to recruit respondents. Thus, guidance on how to study the size and nature of nonresponse has been developed in the past decade.

The last ten years has seen a rise in the use of various methodological and statistical approaches to investigate nonresponse bias in surveys. It has become a best practice to incorporate nonresponse bias studies into survey designs. The U.S. Office of Management and Budget (OMB, 2006)[4] issued a directive that any federal survey that expects to achieve less than an 80% response rate should contain a study of nonresponse bias.

As seen in Table B1, we estimate that we will get a response rate around 20 to 25 percent from the landline sample and a response rate of 15 to 20 percent for the cell phone sample. Because of this, we have also prepared to conduct a non-response bias analysis.

IMLS, with its contractor MDAC, propose to use a follow-up study of non-responders – a methodological approach that actually can be the most robust approach among all the possible methods that can be used (e.g., benchmarking). This approach involves a scientific follow-up survey of the sampled cases from the original survey that ended their field period either as refusals, noncontacts, and language barriers or due to another form of nonresponse (e.g., temporarily incapacitated).

The primary goals of such a follow-up survey are to gain as high a response rate as funding will allow and ask select items from the original survey questionnaire to allow knowledge to be gained on the size and nature of possible bias in the original survey on these key statistics due to nonresponse in that original survey.

---

[4] http://www.whitehouse.gov/sites/default/files/omb/assets/omb/inforeg/pmc_survey_guidance_2006.pdf

Sampling Design. Stratified random samples of non-responding cases in both the landline and cell RDD samples from the original survey will be drawn. Stratification will take place on geography, the type of nonresponse in the original survey (e.g., noncontact, refusal, other), and possibly some other variables available for each case (e.g., total effort made on the case to complete the questionnaire in the original survey).

The size of the total sample of completed questionnaires in the follow-up study will be predicated on a consideration of the size of the likely nonresponse bias that is suspected to be associated with the key variables, selected by IMLS to be included in the follow-up non-response survey questionnaire. The precision at which the size of the possible nonresponse bias will need to be measured will also factor into the decision of what sample size of completions will be needed in the follow-up survey. However it is anticipated that at least 200 completions (100 from each from the landline and cell phone frames) will be needed. Our target will be to have at least an equivalent response rate of the non-responder subsample that is similar to that of the response rate from the whole collection. Thus, if the collection response rate is 20%, we should attempt to attain 20% of non-responders in the follow-up non-response sub-sample.

Follow-up Questionnaire. A much shorter version of the original question will be used. The questionnaire for the non-response bias analysis has 16 items and takes 5 to 8 minutes to complete. By deploying a much shorter questionnaire, response rates in this follow-up study will be considerably enhanced. IMLS has identified the items from the original questionnaire considered to be key items to include in the follow-up non-response survey.

Response/Nonresponse within the Follow-up Non-Response Survey. In order to gather data from a representative sample of non-responders from the original study, a multi-pronged and muscular approach to gaining cooperation from the cases sampled for the follow-up survey will be utilized. IMLS intends to employ the following techniques to increase response:

1. **Advance letter**. Use of an advance letter for those cases for which an address is known. The text of the letter will include an explanation about why gaining cooperation from the household is crucial, that a very brief questionnaire is all that is required, and that a contingent incentive will be given to the household once the questionnaire is completed.

2. **Mode of Recruitment Contact.** All cases sampled for the follow-up study for which a mailing address is known will receive a printed letter in the mail for first contact with them. All cases for which an address is not known will be contacted via telephone.

3. **Within-Unit Eligible Respondent.** Unlike the original survey in which a systematic method was used to select one and only one age-eligible designated respondent within the household for the landline sample, the first age-eligible respondent who answers the telephone in this follow-up survey will be the designated respondent. And if this first person is not willing to complete the questionnaire but there is another age-eligible respondent available at that time that is willing to cooperate, that other person will be permitted to complete the questionnaire. This should eliminate the problem of low response rates associated with within-household "handoffs" to another person that is

associated with the low cooperation in surveys that select one and only one person as the designated respondent.[5]

4. **Contingent Incentives upon Completion of the Follow-up Questionnaire.** A $10 incentive will be promised to each sampled respondent. Research using contingent cash incentives[6] has shown that $10 is the smallest contingent incentive that will be effective in raising respondent rates.

5. **Length of Field Period.** Data collection for this follow-up study should be carried out over a one month period and that period should begin no sooner than one month after the end of the field period for the original study. This lag will provide adequate time to select and train interviewers on the new protocol, and it will provide respondents with a "cooling off" period between the original study and the follow-up non-response study.

6. **Interviewers and their Training.** There always is a considerable variation in the individual response rates achieved by different interviewers in RDD surveys of the general public. And it is not unusual for interviewers with the highest response rates to be three or more times as effective as interviewers with the lowest response rates.[7] For this follow-up survey only interviewers with high respondent rates in the original survey will be deployed. These interviewers will receive a special two-hour training developed and administered by MDAC. This training will include a module on how best to approach the original non-respondents. For example, experience suggests that it is a preferable for an interviewer to acknowledge that the household had previously been sampled in the original survey rather than "hiding" this fact. This special training also will include targeted Avoiding Refusal Training (ART). [8]

Data from the follow-up survey will be analyzed to determine whether or not the sample yielded from the original collection is significantly different, thus indicating potential bias.

B.4.   Tests to Minimize Burden and to Improve Utility

Pre-testing the CATI survey instrument with at least 25 people is planned. The testing will assess the overall time it takes to complete the interview as well as specific timing for sections of the questionnaire. This information will be used to cut questions if necessary. In addition, the pre-

---

[5] "Investigating the Errors that Occur with Within-Unit Respondent Selection." *2010 American Association for Public Opinion Research Conference*, Chicago. (authors: P. J. Lavrakas, T. Tompson and R. Benford)

[6] "Experimenting with Noncontingent and Contingent Incentives in a Media Measurement Panel." *2012 American Association for Public Opinion Research Conference*, Orlando. (authors: P.J. Lavrakas, J. D. Dennis, J. Peugh, J. Shan-Lubbers, E. Lee, and O. Charlebois.)

[7] Lavrakas, P. J. (2010). Telephone surveys. Chapter 14 in J. D. Wright & P. V. Marsden (eds.), *Handbook of Survey Research*. San Diego: Elsevier.

[8] "The Development and Experimental Testing of an Innovative Approach to Training Telephone Interviewers to Avoid Refusals." *2002 American Association for Public Opinion Research Conference*, St. Petersburg. (authors, C. Shuttles, J. Welch, J. B. Hoover, and P.J. Lavrakas).

testing will be used to test questionnaire flow, any difficulties reading specific questions, and how well specific questions are understood by respondents. Improvements in question order and wording will be made as necessary.

In addition, for landline telephone numbers, MDAC will use predictive dialing technology and automated number recognition (ANI) to improve the productivity and data quality. (Note: it is illegal to use predictive dialing technology with known cell phone numbers, so this applies only to the landline sample.) Predictive dialing systems enhance productivity by dialing numbers; recognizing busy signals, answering machines, non-working numbers, and ring no-answers results; and communicating the dispositions of those calls to the CATI system without expending interviewer time. By relieving interviewers of these unproductive dials, productivity can increase by up to 50 percent, delivering cost savings to the government. ANI helps mitigate the Caller ID screening that has resulted in a decline in response rates for telephone surveys. Potential respondents are typically less likely to answer a call when the Caller ID display doesn't identify the incoming number. To increase the likelihood of a respondent answering the call, MDAC's dialer system will work with the local telephone system to provide an actual name or telephone number on the Caller ID display.

B.5. Individuals Responsible for Study Design and Performance

The following individuals are responsible for the study design and the collection and analysis of the data on PNLMS.

Table B2. Personnel Involved with PNLMS

Person	Address	Email / Phone
Institute of Museum and Library Services (IMLS)		
Carlos Manjarrez Director, OPRE IMLS	1800 M Street, NW, 9th floor Washington, DC 20036	cmanjarrez@imls.gov 202-653-4671
Deanne W. Swan, Ph.D. Senior Statistician IMLS	1800 M Street, NW, 9th floor Washington, DC 20036	dswan@imls.gov 202-653-4769
M. Davis and Company, Inc. (MDAC)		
J. Scott Osberg, Ph.D. Project Manager MDAC	3000 Market St, Suite 202 Philadelphia, PA 19104	scott@mdavisco.com 202-393-2376
Michael Campbell Assistant Project Manager MDAC	3000 Market St, Suite 202 Philadelphia, PA 19104	michael@mdavisco.com 215-790-8900 x132
McKenzie Ballou, Ph.D. Research Assistant MDAC	3000 Market St, Suite 202 Philadelphia, PA 19104	mckenzie@mdavisco.com 215-790-8900 x126
David Ferree Call Center Manager MDAC	3000 Market St, Suite 202 Philadelphia, PA 19104	david@mdavisco.com 215-790-8900 x241
Thomas Sexton, Ph.D. Statistician State University of New York at Stonybrook	College of Business 317 Harriman Hall Stonybrook, NY 11794	Thomas.Sexton@StonyBrook.edu 631-632-7181
Herbert F. Lewis, Ph.D. Statistician State University of New York at Stonybrook	Dept of Technology & Society 339 Harriman Hall Stonybrook, NY 11794	Herbert.Lewis@StonyBrook.edu 631-632-7172