

# **SUPPORTING STATEMENT FOR PAPERWORK REDUCTION ACT SUBMISSION**

## **Passport Demand Forecasting Study Phase III OMB Number 1405-0177 Form Number SV2011-0010**

### **B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS**

1. The main objective of this study is to estimate the overall demand for passports and passport cards. The monthly surveys will provide CA/PPT with a body of continuously updated and reliable statistics. The monthly surveys will be crucial to ensuring CA/PPT's ability to reliably determine short (3 months), medium (12 months), and long-term (2-3 years) passport demand on a continuous basis. The potential respondent universe (or the target population) will consist of all U.S. citizen and U.S. national persons (16 years of age or older) living in U.S. households in any of the 50 states and in the District of Columbia.

Gallup or a similar organization will use an Address Base Sampling (ABS) approach to generate a representative sample of the target population (U.S. citizen and U.S. national persons 16 years of age and older) every month. The sample will be created by using the United States Postal Service's (USPS) Second Generation Delivery Sequence File (DSF<sup>2</sup>) database, which is a service made available through a nonexclusive license agreement with private companies. DSF<sup>2</sup> is a computerized file that contains all delivery point addresses serviced by the USPS, with the exception of general delivery. Commercial vendors then have the option of joining additional information to the address record from other sources, including surname, age of head of household, income level and other relevant variables. Using this augmented sampling frame, a sample of approximately 16,000 persons will be drawn to complete around 4,000 surveys nationwide every month.

The primary mode of data collection will be by mail, however, Gallup or a similar organization will send out surveys via mail along with an internet link to the corresponding internet/web version of the surveys that will reflect the same identical questions for those respondents who prefer to respond via the internet. The survey will contain distinct variables to answer questions covering passport demand (books and cards), other topics of interest such as:

- interest in making an appointment to go in person to a local Passport Agency to complete a passport application;
- non-international travel use of passports;
- knowledge and impact of Real ID; and
- other socio-demographic variables of interest to CA/PPT.

2. The methodology used for the collection of information in the Monthly Passport Demand Forecasting Studies will be based on Address Base Sampling (ABS) through the use of the United States Postal Services Second Generation Delivery Sequence File (DSF<sup>2</sup>) database. The DSF<sup>2</sup> database is useful with Address Base Sampling (ABS) as an alternative design to Random Digit Dialing (RDD) telephone surveys. Address Base Sampling (ABS) is a cost efficient alternative to RDD ensuring better coverage (covering about 98 percent of U.S. households) and is of comparable speed and efficiency.

According to the United States Postal Service, The Delivery Sequence File (DSF<sup>2</sup>) is a service made available through a nonexclusive license agreement with private companies. The Delivery Sequence File (DSF<sup>2</sup>) is a computerized file that contains all delivery point addresses serviced by the USPS, with the exception of general delivery. On the file, each delivery point is a separate record that conforms to all USPS addressing standards. Each record contains the ZIP+4 code, carrier route code, delivery sequence, delivery type, and seasonal delivery information. The nonexclusive licensees offer this service to the mailing industry by comparing a mailer's address list against the DSF using USPS-approved Coding Accuracy Support System (CASS) address-matching software. The results of a match or no-match situation are conveyed via a series of footnote codes.

DSF<sup>2</sup> provides mailers with the following:

- Address validation
- Address standardization
- ZIP+4 coding
- Carrier route coding
- Delivery sequence
- Detection of addresses that are potentially undeliverable for reasons other than a change of address (nixies)
- Cost savings through improved address list quality
- Delivery-type code that indicates business or residential
- Seasonal delivery information

Commercial vendors then have the option of joining additional information to the address record from other sources, including surname, age of head of household, income level and other relevant variables. At the first stage of sampling, the monthly sample (of approximately 16,000 households) for this study will be drawn from this augmented database as the sampling frame. This sample will be a nationwide simple random sample and is, therefore, expected to represent different demographic subgroups roughly in the same proportion they are represented in the target population. However, if response rates differ significantly by subgroups, it may become necessary to oversample certain subgroups with relatively low response rate (for example, younger persons or specific ethnic/racial groups). Available information on the sampling frame

will allow such oversampling as part of the ABS (Address Based Sampling) approach. Once the first stage sample of households is selected, the second stage sampling will involve selection of one eligible (16 years of age or older) person from the selected household using the “most recent birthday” method.

As mentioned before, an initial sample size of 16,000 persons will be selected to generate about 4,000 completed surveys per month based on an anticipated overall response rate of 25 percent. Gallup previously tested the impact of monetary incentives (\$1.00 in the mailing) on mail survey response rates. Sampled persons were randomly assigned to the incentive or control group and a total of 6,600 surveys were completed. It was found that the overall response rate was significantly higher (25 percent) for the incentive group as compared to that of the control group (12 percent). The incentive was not found to have any adverse effect on sample representativeness. In view of these findings and the overall objective of maximizing response rate, the plan is to include a \$1 incentive in the mailing for the monthly surveys.

To achieve a high level of confidence on survey based estimates, an estimated 48,000 survey respondents (4,000 x 12) will have data collected annually. Each survey is estimated to be no more than 10 minutes in length, including any screening for qualification. Based on the time to conduct an interview and the total number of interviews (48,000), the hourly burden time for the responding public to complete the Passport Demand Forecasting survey is estimated to be 8,000 hours.

This proposed monthly sample size of 4,000 completed surveys will ensure reasonable precision for survey-based estimates at the national level as well as for different subgroups of potential interest. For the estimation of an unknown population proportion (for example, proportion applying for new passports), the sampling error associated with estimates based on a sample size of 4,000, ignoring design effect, is about  $\pm 1.5\%$  at the 95% level of confidence. Even after taking into account an anticipated design effect of about 1.5, the sampling error is not likely to exceed  $\pm 2\%$ . It may also be noted that this is based on the most conservative assumption that the unknown population proportion is around 50 percent and hence the actual sampling error for most estimates based on similar sample sizes are likely to be even lower. At the subgroup level, the precision will obviously depend on the sample size (number of completed surveys). For any subgroup with a sample size of 1,000, for example, and an anticipated design effect of about 1.5, the sampling error is not likely to exceed  $\pm 3.8\%$ .

The sample data will be weighted to minimize bias associated with the monthly estimates. In order to adjust for unequal selection probabilities and non-response, it will be necessary to use appropriate weighting procedures to generate unbiased estimates. The first stage of weighting to correct for unequal selection probability will involve weighting each case by the ‘number of eligible (16 years of age or older) persons’ in the selected household. This number may be truncated (at 3, for example) to avoid extremely large weights. If oversampling of certain subgroups is used, appropriate weighting adjustments will be carried out to compensate for unequal selection probability resulting from disproportional sample allocation across suitably defined strata. At the next stage, post-stratification weighting will be carried out to project the sample results to known characteristics for the target population. The target data for post-stratification weighting for demographic variables may be derived from the latest Census or CPS estimates. The following demographic variables (along with their categories) may be used for post-stratification adjustments:

- Race: White Only, Black Only, Other
- Region by Gender by Age: Four Census Regions; Male/Female; 16-34, 35-44, 45-54, 55+
- Ethnicity: Hispanic/Non-Hispanic
- Education: Less than college, Some college, and College grad

The post-stratification weighting process will be iterative and will be repeated using combing algorithms until all weighted numbers are reasonably close to their corresponding targets. Finally, some trimming of weights may become necessary to avoid extreme weights and its effect on variance of estimates.

3. To maximize coverage of the target population, the contractor will use the Delivery Sequence File (DSF<sup>2</sup>) database for Address Base Sampling. The Delivery Sequence File (DSF<sup>2</sup>) contains all delivery point addresses serviced by the USPS, with the exception of general delivery. On the file, each delivery point is a separate record that conforms to all USPS addressing standards.

Gallup or a similar organization will then have to purchase additional information to the address record from other sources and attach to the address such as: last name, age of head of household, and income level. The use of Address Base Sampling will ensure that the survey reaches the appropriate audience and leverage the best research method to obtain responses for the given survey topic. When used in conjunction with mail surveys, the DSF<sup>2</sup> provides survey researchers with the advantages of cost effective sampling designs as well as a means of reducing survey non-response and reaching cell phone-only households and households without telephones which are missed by RDD designs.

To examine the potential for nonresponse bias, Gallup or a similar organization will conduct a nonresponse bias study for the monthly surveys and complete up to 400 interviews semiannually for up to 800 surveys annually. The contractor will use mail and telephone for the data collection mode. For the mail portion, they will use an approach similar to the monthly survey protocol described above. For the telephone portion, a seven-call design will be employed where telephone interviewers will make up to seven calls to contact and complete a telephone interview.

The questionnaire for the nonresponse follow-up study will include a selected subset of questions from the main monthly Passport study. The analysis plan for the nonresponse bias study will include comparing the respondents and the non-respondents on key variables including demographic and attitudinal/behavioral questions and examining the nature of nonresponse pattern and its effect on the demand estimates. In order to ensure cooperation from the non-respondents, the Contractor will offer all sample members (for the non-response follow-up study) a \$10 cash incentive to complete the telephone interview.

4. Survey questionnaires and procedures will be tested in several ways. The questionnaire will be internally pre-tested by Department of State and Gallup or a similar organization's personnel for timing, content and clarity. Gallup has previously conducted a review of the questionnaire using cognitive testing methods (viewing the respondent as he or she answers questions, follow-up

probing questions to ensure understanding as intended of the questions, etc.) involving nine or fewer participants to examine the comprehensibility, structure and order of survey items.

Gallup or a similar organization will also engage in a formal pre-test of the survey instrument, to be conducted with a sample of 50 households to confirm that the screening questions and procedures, as well as all items in the survey questionnaire are working as intended. Although the pre-test will be designed as a confirmatory procedure, if any issues are uncovered with survey instructions, item wording, or response categories during the process, revisions will be proposed and incorporated into the final survey materials upon receipt of agency approval.

5. Gallup previously developed the survey design and Gallup, or another contractor, will be responsible for collecting, processing, and analyzing the data and presenting findings to CA/PPT. The following individuals were consulted in developing the survey design, the sampling plan, and statistical aspect of the study.

Dr. Manas Chattopadhyay  
 Chief Statistician  
 Gallup, Inc.  
 901 F Street, NW  
 Washington, D.C. 20004  
 202-715-3179

Dawn Royal, M.A.  
 Engagement Manager  
 Gallup, Inc.  
 901 F Street, NW  
 Washington, D.C. 20004  
 949-474-2707

Edward M. Worthan  
 Management and Program Analyst  
 Passport Services  
 Bureau of Consular Affairs  
 U.S. Department of State  
 2100 Pennsylvania Ave, NW  
 Washington, D.C. 20037-3202  
 202-663-1073

Eric Olesen, M.S.  
 Sr. Statistical Analyst  
 Gallup, Inc.  
 901 F Street, NW  
 Washington, D.C. 20004  
 202-715-3208

Dr. Don Dillman  
 Gallup Senior Scientist  
 Department of Sociology and  
 The Social & Economic Sciences Research Center  
 Washington State University  
 Pullman, WA 99164-4014  
 509-335-1511