Supporting Statement for Paperwork Reduction Act Submission
Department of Housing and Urban Development
Evaluation of the Office of Public and Indian Housing's (PIH) Energy Performance
Contracting (EPC) Program

PART 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.

Description of Population and Sample

The web survey collects information from three groups of public housing authorities (PHAs): Group 1: PHAs that have executed an EPC, Group 2: PHAs that started the EPC process, but did not complete it, and Group 3: PHAs that never executed an EPC. Our proposed sample design approach is to stratify the entire population of PHAs under group 1 (EPC population) and group 3 (non-EPC population) by the size and region of the PHA. Group 2 will not be stratified as the population size is very-small. Moreover, group 2 is not intended to be compared against groups 1 and 3 via performance measures. The survey questions for group 2 primarily focus on the HUD EPC application process and related topics.

The population counts for the three size categories are shown in the table below. The size distribution in the population is skewed, with fewer number of large PHAs compared to medium and small PHAs. However, large PHAs account for a majority of units affected by EPCs. For any moderately sized sample, an equal probability sample design would give extremely small numbers of larger PHAs. So, to allow for separate analysis by size, we are selecting a reasonably large sample size per size-band. This will be achieved by systematically increasing the sampling fraction as the size of PHA increases.

Group Population Counts

Small	Group 1	Group 2	Group 3
South	24	10	1308
West/Midwest	33	15	961
Northeast	43	7	251
Total	**100**	**32**	**2520**
Medium	Group 1	Group 2	Group 3
South	26	2	56
West/Midwest	21	5	35
Northeast	27	3	29

Total	**74**	**10**	**120**
Large	Group 1	Group 2	Group 3
South	15	7	19
West/Midwest	18	2	16
Northeast	30	0	8
Total	**63**	**9**	**43**
TOTAL	**237**	**51**	**2683**

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.

Description of procedures for data collection

Statistical methodology for stratification and sample selection. For the purpose of this data collection, HUD anticipates utilizing a stratified sampling design, with size and region being the stratification groups. The sizes of the PHAs are categorized as small, medium and large. We will categorize regions according to U.S. Census regions: South, West/Mid-west and Northeast. Since the West region has smaller sub-group counts, we have combined the West and Midwest regions into a single category for the purpose of getting a nationwide representation for this survey. We assume a response rate of 70% and anticipate receiving approximately 299 responses from the 427 web survey invitations that we will send out.

EPC (Group 1): For the EPC population, we will employ a disproportionate stratified sampling strategy for the three size categories defined above. We will apply an incremental sampling fraction for each size category from small to large PHAs. Large PHAs will be selected with certainty; i.e., the sampling fraction will be 1. In addition, the sample will be stratified by region, although within each size category, the sampling fraction will be kept roughly constant and the sub-samples will be selected proportional to corresponding population counts.

EPC Sampling Strategy

		EPC Population	EPC Sample	Sampling Fraction
	South	24	16	0.6
	West/Midwest	33	21	0.6
Small	Northeast	43	28	0.6
	Sub-Total	100	65	0.6
	South	26	21	0.8
	West/Midwest	21	17	0.8

Medium	Northeast	27	22	0.8
	Sub-Total	74	60	0.8
	South	15	15	1
	West/Midwest	18	18	1
Large	Northeast	30	30	1
	Sub-Total	63	63	1
	TOTAL	**237**	**188**	

Non-EPC (Group 3): For the non-EPC population, we will try to match the EPC population in all the sub-groups to make this group as similar as possible to the EPC population. Large PHAs will be selected with certainty. There are only 43 large PHAs in group 3 compared to 63 in group 1. We plan to slightly increase the small and medium size-bands in group 3 to compensate for the smaller number of large PHAs in group 1.

Non-EPC Sampling Strategy

		Non-EPC Population	Non-EPC Sample
Small	South	1308	20
	West/Midwest	961	25
	Northeast	251	32
	Sub-Total	2520	77
Medium	South	56	24
	West/Midwest	35	20
	Northeast	29	24
	Sub-Total	120	68
Large	South	19	19
	West/Midwest	16	16
	Northeast	8	8
	Sub-Total	43	43
	TOTAL	**2683**	**188**

Group 2: Group 2 members will be selected with certainty as there are only 51 PHAs in the population.

Summary of Main Web Survey Sample Design

Group	Population	Sample	70% Response
EPC (Group 1)	237	188	132
Started, not completed (Group 2)	51	51	35
Non-EPC (Group 3)	2683	188	132
Total	**2971**	**427**	**299**

Telephone Interviews: The study will include in-depth interviews of approximately 28 participants who completed the web survey. These interviews will be based on the responses provided by the participants. The sample will be strategically identified based on their responses

to the main web survey. These interviews will be conducted to gather qualitative data and not used for quantitative analyses.

Option period survey: If the option to conduct a study on smaller PHAs is exercised, there will be a survey of very-small to small PHAs. For this survey, it is anticipated that the survey instrument will remain the same. It is anticipated that we will select 86 participants from each of the above-mentioned groups 1 and 3. We will stratify groups 1 and 3 of very-small to small PHAs by region, to get a nationwide representation. In addition, we will randomly select 22 very-small to small PHAs from Group 2 that had started the EPC process, but not completed it. The table below shows the population and sample counts for the option survey for groups 1 and 3. Assuming a response rate of 70%, we anticipate a total of approximately 136 responses to the option survey. Due to the low number of small to very-small PHAs that have executed an EPC, there will be an overlap with the sample that was surveyed in the first year. To eliminate re-surveying a small PHA and to reduce burden, we will use the data collected from the base web survey in the analyses of small PHAs. If a small PHA is selected in the sample for the option survey and had already completed the first survey, we will not survey them again. Telephone interviews will be conducted with approximately 18 very-small to small PHAs that participated in the web survey to capture qualitative data.

Sampling Strategy for Very-small to Small PHA Web Survey

	EPC Population	EPC Sample (Group 1)	70% Response
South	24	21	
West/Midwest	33	28	
Northeast	43	37	
Total EPC	**100**	**86**	**60**
	Started EPC, but not completed (Group 2)	Started EPC, but not completed (Group 2)	
	32	**22**	**16**
	Non-EPC Population	Non-EPC Sample (Group 3)	
South	1308	21	
West/Midwest	961	28	
Northeast	251	37	
Total Non-EPC	**2520**	**86**	**60**
TOTAL	**2652**	**194**	**136**

Data Collection. It is anticipated that the web survey invitations will be sent to 427 respondents for the first survey; and to 194 respondents for the optional follow-up survey of very-small to small PHAs. For this study, the primary mode of data collection will be web surveys. Our survey administration protocol will consist of the following steps:

- Each participating PHA will have a point of contact who will see to the response from that PHA.

- Each respondent from a participating PHA will receive a pre-notification e-mail a week prior to the launch of the survey; the pre-notification e-mail will identify the participant, invite the participant to complete the survey and explain the survey purpose.

- Each respondent will then be sent a customized web survey link based on their PHA's EPC execution status, as described in the 3 groups above.

- Each respondent will receive a weekly reminder after the launch of the survey to help boost response rates.

- The data collection will end after 8 weeks of survey administration.

Degree of accuracy needed for the purpose described in the justification. The sample design mentioned above will yield a nationwide representation with a margin of error of ±3% at a confidence level of 95% among the EPC population and a margin of error of ±6% at a confidence level of 95% among the non-EPC population.

Unusual problems requiring specialized sampling procedures. At this point, HUD does not anticipate any unusual problems that require specialized sampling procedures.

Any use of periodic (less frequent than annual) data collection cycles to reduce burden. HUD does not anticipate any use of additional data collection cycles at this point.

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.

Methods to Maximize Response Rates

Several steps will be taken to boost participation and ensure maximum response rates. The steps will include:

- Constructing the survey instrument to be as respondent-friendly as possible

- Minimizing respondent time by providing branching options on the web survey to automatically route respondents to relevant questions

- Sending a pre-notification e-mail informing potential respondents about the survey

- Sending weekly reminders to help boost response rates

- Validating and legitimizing the survey via signed letters from a Department of HUD official

- Use of Department of HUD graphics and color scheme

- Providing a toll-free helpdesk phone number and an email address to address respondents' issues and questions

Dealing with non-response

Data from all scientific surveys are weighted before data can be used to produce reliable estimates of population parameters. There are some practical limitations of a sample survey,

such as unequal selection probabilities, differential non-response and under-coverage. Typically, the weighting process consists of three major steps. The first consists of computation of base weights as reciprocal of the selection probabilities. The second involves adjustment of the base weights for non-response bias. In the third, the weights are further calibrated to known population parameters to compensate for sampling frame inadequacies. For this survey, we will use statistical software tools to perform weighting on collected survey data. To adjust for non-response bias, we will estimate response propensities using logistic regression by dividing the sample into classes of respondents and non-respondents.

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.

   Testing

   The survey instrument has already been tested on usability, functionality, survey instructions, communication materials and protocol. We pilot-tested the survey with 9 PHAs selected from the sample pool. We selected 3 PHAs that have executed an EPC (Group 1), 3 that have begun the EPC application process, but not completed it (Group 2), and 3 that have never executed an EPC (Group 3). In the pilot test conducted, 6 participants completed the survey. Of those who completed the survey, one (1) respondent was from a PHA in Group 1, three (3) were from PHAs in Group 2, and two (2) respondents were from PHAs in Group 3. All participants were able to complete the survey successfully and submit their responses. The feedback we received was mostly positive, and the only comment we received was regarding the communication which implied that the survey was for only those PHAs that executed the EPC. We updated the communication materials and the survey title to ensure clarity.

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.

   Contact Information for Statistical Consultants and Analysts

   Jyothsna (Jo) Prabhakaran, (571) 633-7796, LMI – Sampling, Design and Analysis

   Mike Canes, (703) 927-0076, LMI – Design and Analysis