OMB Clearance Package

Quality Control for Rental Assistance Subsidy Determinations

Section B. Collections of Information Employing Statistical Methods

Submitted by:

Office of Policy Development and Research
Department of Housing and Urban Development
Washington, DC 20410

Prepared by:

ICF Macro
11785 Beltsville Drive
Calverton, MD 20705-3119

July 1, 2010

Table of Contents

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

1. *Respondent universe and sampling methods*

Provide a numerical estimate of the potential respondent universe and describe any sampling or other respondent selection method to be used. Data on the number of entities (e.g., households or persons) in the universe and the corresponding sample are to be provided in tabular format for the universe as a whole and for each stratum. Indicate expected response rates. If this has been conducted previously include actual response rates achieved.

Currently, the allocations have not been completed for the drawing of the sample. However the methodology and the population are very similar to those used last year, so previous sample figures will be provided. The sample sizes, number of primary sampling units (PSUs) and total number of respondents will be the same. The stratification approach will be the same, but because of the use of implicit stratification and population changes, the exact number of units per stratum will vary.

The universe includes all assisted housing projects and tenants located in the continental United States, Alaska, Hawaii, and Puerto Rico. The following programs will be included in the sample:

- Public Indian Housing (PIH)-administered Public Housing (i.e., Public Housing)

- PIH-administered Section 8 projects
  - Moderate Rehabilitation
  - Vouchers

- Office of Housing-administered projects (i.e., owner-administered)
  - Section 8 New Construction/Substantial Rehabilitation
  - Section 8 Loan Management
  - Section 8 Property Disposition
  - Section 202 Project Rental Assistance Contracts (PRAC)
  - Section 202/162 Project Assistance Contracts (PAC)
  - Section 811 PRAC

The sample will be designed to obtain a 95 percent likelihood that estimated aggregate national rent errors for all programs are within 2 percentage points of the true population rent calculation error, assuming an error of 10 percent of the total rents (based on OMB criteria). In previous studies, we determined that a tenant sample size of 2,400 will yield an acceptable precision for estimates of the total average error.

In addition to the overall estimates, error rates will be estimated for each of the three major program types (Public Housing, PIH-administered Section 8, and Owner-Administered programs). Assuming each constitutes a third of the sample, we will require a 95 percent confidence interval within 5 percent of their population values. Assuming a design effect of 2.0, we multiplied that by 400, a number slightly larger than the number required for the desired

precision in case of a random sample, and obtained a tenant sample size of 800 per program, for a total sample size of 2,400. The design effect is the ratio of the variance of the estimate to the variance of the estimate for a random sample of the same size. Past experience has shown a design effect of 2 to be a reasonable assumption for this design.

The FY 2008 study found an average QC rent of $219.93 and an average error of $42.59 and a standard error of $2.59. This yields a 95 percent confidence interval of $5.41. This constitutes 2.5 percent of the QC rent. HUD considered this an acceptable confidence interval and hence the basic elements of the design and the sample sizes are being preserved.

As will be described later, the sample will be a three stage sample with 60 PSUs consisting of counties or groups of counties, ten projects within each PSU and four tenants per project. HUD regions will be used as implicit strata in PSU selection, and the three program types will sub-stratify the PSUs. Table B1.1 illustrates the classification of states, the District of Columbia, and Puerto Rico to HUD regions.

Table B1.1. Allocation of States to HUD Regions

HUD Region	States
1	CT, MA, ME, NH, RI, VT
2	NJ, NY
3	Washington DC, DE, MD, PA, VA, WV
4	AL, FL, GA, KY, MS, NC, Puerto Rico, SC, TN
5	IL, IN, MI, MN, OH, WI
6	AR, LA, NM, OK, TX
7	IA, KS, MO, NE
8	CO, MT, ND, SD, UT, WY
9	AZ, CA, HI, NV
10	AK, ID, OR, WA

Table B1.2 presents the population and expected sample size by region from the FY 2009 study.

Table B1.2. Number of Projects and Units in Sampling Frame by HUD Region
FY 2009 Study

HUD Region	Projects				Units				Expected PSU Sample	Actual PSU Sample
	PIH-Admin Sec 8	Public Housing	Owner Administered Sec 8	Total	PIH-Admin Sec 8	Public Housing	Owner Administered Sec 8	Total		
US	11,738	6,135	18,449	36,322	1,820,800	949,287	1,337,339	4,107,426	60.00	60
1	839	380	1,598	2,817	117,123	60,354	119,659	297,136	4.35	5
2	1,607	593	1,557	3,757	281,501	240,843	157,469	679,813	10.52	10
3	929	535	1,813	3,277	143,446	72,462	151,371	367,279	5.37	6
4	2,261	1,786	3,402	7,449	328,115	252,435	240,076	820,626	12.52	13
5	1,783	996	4,002	6,781	269,074	131,796	301,192	702,062	10.24	10
6	1,338	895	1,611	3,844	197,813	88,657	103,932	390,402	5.60	5
7	626	442	1,118	2,186	78,481	33,535	60,571	172,587	2.48	2
8	476	161	784	1,421	60,447	14,825	37,647	112,919	1.54	2
9	1,508	257	1,759	3,524	284,979	45,486	130,531	460,996	6.04	6
10	371	90	805	1,266	59,821	8,894	34,891	103,606	1.34	1

Response Rates

Two types of non-response may effect this data collection: that by PHAs/owners and tenants.

PHAs/owners

Project-Specific Information
Participation by selected PHAs/owners is mandatory such that their contracts with HUD require their participation in studies of this type. In the FY 2009 study all PHAs/owners completed the Project-Specific Information Form resulting in a 100 percent response rate. We anticipate a similar response rate for the upcoming studies.

In an effort to ensure PHA/owner participation, the initial mailing is conducted using an overnight delivery service to catch their attention. PHAs/owners are given a date by which the information is needed and if that time elapses, follow-up telephone calls are made to obtain the needed information. If further follow-up is required, a list of the non-responsive PHAs/owners are provided to HUD and contacted by them as well.

Project Staff Questionnaire
Participation by selected PHAs/owners is mandatory such that their contracts with HUD require their participation in studies of this type. For the FY 2009 study, all of the 552 PHAs/owners completed the Project Staff Questionnaire resulting in a 100 percent response rate.

In an effort to ensure PHA/owner participation, the initial mailing is conducted using an overnight delivery service to catch their attention. PHAs/owners are given a date by which the information is needed and if that time elapses, follow-up telephone calls are made to obtain the needed information.   If further follow-up is required, a list of the non-responsive PHAs/owners are provided to HUD and contacted by them as well.

Tenants

Participation by selected tenants is mandatory; refusal to participate could result in their termination of assistance. In the FY 2008 study, 104 tenants were non-responsive out of 2,401 total tenants, resulting in a 96 percent tenant response rate.

The most common reason for tenant non-response was serious illness. Other common reasons for replacement included: 1) the tenant moved out of the sampled unit between the file abstraction and household interview phases of the study; 2) the tenant refused to participate in the study, and 3) the tenants were away for extended periods and could not be contacted for an interview during the four month data collection window. Field interviewers are required to make at least four in-person contacts with the tenant to conduct interviews with individuals who try to evade the interview. For the FY 2010 study a similar tenant non-response rate is anticipated. Study time limits and budget constraints do not allow us to further pursue tenants who evade, refuse or are away during the data collection period.

2. *Procedures for collection of information*

Describe the procedures for the collection of information, including: Statistical methodology for stratification and sample selection; the estimation procedure; the degree of accuracy needed for the purpose in the proposed justification; any unusual problems requiring specialized sampling procedures; and any use of periodic (less frequent than annual) data collection cycles to reduce burden.

Basic Cluster Design

Two levels of clustering will be used in this study:

- Projects clustered within PSUs, which are generally groups of counties

- Tenants clustered within projects

The optimum number of tenants per project is based on a cost ratio of two additional tenants for each additional project, PSU intraclass correlation (δ), project cost (C), and tenant cost (c):

$$opt.\ n=[(C(1-\delta))/(c\delta)]1/2$$

References for this formula can be obtained in Hanson, Hurwitz and Madow, Vol I., 1953, formula 16.2. We estimate that adding a project would result in a cost comparable to adding two tenants. In the FY 2003 study, we applied this formula and determined that a sample size of 2.74 tenants per project would be optimal. We chose four tenants per project in order to preserve an acceptable measure of intra-project variance and to take advantage of the fact that errors appear to be concentrated in projects. In fact, in the FY 2007 study we found that the projects accounted for almost 6 percent of the variance in gross error, and this was statistically significant (p<.001). Since the FY 2003 study we used the same basic design with minor modifications.

The optimal number of projects and tenants per cluster is a function of logistics. The same two to one ratio that was applied to calculate the optimal number of tenants per project can be used to define cost units. A cost unit is the cost of including a tenant in the survey. Cost units are a

function of the data collector's time and other factors. Ten projects and four tenants per project in a PSU produces sixty cost units ($2*10 + 1*10*4 = 60$). A design with six projects and eight tenants per project would also have sixty cost units ($2*6 + 1*6*8 = 60$). Experience has shown that greater than sixty cost units results in an impractical amount of work for one data collector to handle. We believe that sixty cost units provide the best balance between logistical requirements and design effect. Given these issues, we decided to sample four tenants per project, ten projects per cluster, and sixty clusters, for a total of 2,400 tenants.

<u>Definition, Allocation and Sampling of Clusters</u>

A sample of 60 PSUs will be designed, with ten projects per PSU and four tenants per project (allowing PSUs and projects to be selected more than once if sufficiently large). The design calls for equal allocation of the three HUD programs: Public Housing, PIH-administered Section 8, and owner-administered projects. The earlier samples were designed to yield the expectation of the same number of households for each program type, but for the last several years the design was modified so it would select exactly the same number of households per program. One additional project has been added to this design to insure contractual compliance in the event that at the last minute something prevents data from one project to be properly processed.

Source files used for sample selection

The source files for the FY 2010 study are currently being reviewed. Base on previous experience with the types and numbers of files typically provided, we expect to receive similar information.

OWNER-ADMINISTERED PROJECTS. HUD provided two files of information on owner-administered projects. One file had a record for each property, including the address. The second listed the contract numbers that corresponded to each property. Certain types of contract were excluded from the files because the rent calculation rules used for these contracts are outside the scope of this study; these include SUPP, RAP, service coordinator, and expired contracts.

VOUCHER AND MODERATE REHABILITATION PROJECTS. HUD provided one file that contained information on Voucher and Moderate Rehabilitation households, including geographic information. Out-of-state households (households with transport vouchers who used them in another state) will be eliminated from the frame.

PUBLIC HOUSING PROJECTS. One Public Housing file was provided by HUD, and included geographic information for all but a few projects. Since they are out of scope for the study, Move-to-Work PHAs were not included in the file. As needed, we will use the county of the PHA or the county from a previous year file to classify these Public Housing projects into counties. As in past years, Louisiana parishes affected by Hurricane Katrina (i.e., Jefferson, Orleans, Plaquemines, St. Bernard, St. Charles, St. John the Baptist, St. Tammany, Calcasieu, Cameron) will also be excluded from the frame.

Across all program types, project covering fewer than 14 units will be excluded so as to not unduly burden especially small projects and to increase the efficiency of the data collection by

decreasing travel to numerous small projects to collect the 2,400 cases. The number 14 was chosen at a time when seven households were selected per project, and was preserved in order to insure comparability of the frames for subsequent years. In addition, any projects that are located in Guam, the Northern Mariana Islands, and the Virgin Islands will be removed from the frame. Once the above files are processed, it will be possible to estimate the number of tenants in each program in each county.

Sample cluster size

The clustering procedure will use counties as the initial cluster. Clusters will be restricted to those with a minimum number of tenants and projects. In the FY 2009 study, the requirements were 40 projects and 2,000 tenants, and at least two PHA/county combinations. For these purposes, vouchers will be counted as one project for the first 300 tenants, and as an additional project for every 200 tenants above that (e.g., 500 tenants would count as two voucher projects, but 501 would count as three). When a county does not meet the criterion, we will identify the nearest county in the same state and merge the two. A total of 472 PSUs were created for the FY 2008 study, and 321 were created for the FY 2009 study. The decrease was due to a change in the requirements for a cluster. In FY 2008, we required 30 projects, at least three of each kind, to form a cluster. That led to a few cases where replacing from within the cluster was difficult, so the minimum was augmented to 40 projects, at least four from each program type This meant bigger clusters, and hence fewer of them nationwide.

The clustering program has been highly effective in previous years' efforts, except that from time to time the resulting PSUs have been unnecessarily large. This has been resolved in the past by a manual revision of PSUs after selection. We will use the new files to create PSUs anew, and will examine the resulting PSUs to determine whether it is desirable to modify the resulting parameters.

We will select PSUs with probabilities proportional to size (PPS), a standard approach followed in most national surveys. However, the study calls for an equal number of tenants to be selected from each of the three major classes of programs. In order to accomplish this, we will select PSUs with a size measure calculated as the average of the proportions of tenants from each of the three programs found in the PSU. The number of tenants in each program within a PSU will be divided by the number nationwide. The three values will be averaged to create a measure of size that sums to one.

The size measure will then be multiplied by 60—the number of PSUs to be selected—to obtain the expectation of selection for each PSU. If this expectation is less than one it will be interpreted as the probability of selection of the PSU. If it is greater than one, the PSU will be selected with certainty. The integer part of the expectation will indicate the minimum number of times the PSU can be selected and the fractional part will indicate the probability that the PSU will be selected one additional time.

Sample cluster selection

The PSUs will be grouped within states and then within HUD-defined regions. States will be sorted in a random order within regions, and PSUs will be randomly sorted within states. As the

frame is prepared for the selection of PSUs, PSUs will be arranged in order and each assigned an expectation value. A random number will be generated as a starting point to select the PSUs. A cumulative distribution of the expectations will be calculated by adding the expectation of a PSU to the cumulative expectation of the previous one (starting with the random number). Thus the real numbers between 0 and 60 will be divided into segments where each PSU is represented by the segment between the cumulative expectation of the previous PSU (or 0 for the first PSU) and its cumulative expectation. A random number x between 0 and 1 will be selected, and the integers from 0 to 59 will be added to the random number. The numbers x, 1+x, 2+x ... 59+x will define the selected PSUs and a PSU will be selected as many times as one of these numbers falls into its corresponding segment.

This is essentially the Goodman-Kish approach (1950) but using sampling with minimal replacement (Chromy, 1979)[1,2]. This procedure results in sample sizes roughly proportional to the number of tenants in each region, but counting tenants in smaller programs more than those in larger programs. Rather than allocate a number of clusters to this region, this method implicitly stratifies the sample and essentially allows a fractional allocation. In other words, if the expectation for a region should be 4.6 PSUs, it would have a 40 percent chance of getting 4 and a 60 percent chance of getting 5.

In addition, once the PSUs are selected, the larger PSUs will be divided and one of the parts will be selected with PPS. The decision to divide or not will be implemented subjectively, using a map to determine data collection burden. Once a division is made, one of the parts will be selected with PPS using the same combined size measure used in selecting the PSUs.

Allocation and Sampling of Projects

Over the last few years of quality control studies, different methodologies have been used in the allocation and sampling of PHAs/projects. These methodologies have been employed to identify an approach that continues to improve the evenness of probabilities of selection. As has been done since the FY 2006 study, the project sample will be selected such that there will always be 10 projects in each PSU. These will be selected by first allocating a fractional number of projects to each cell and then using controlled rounding to make the rows add up to ten projects per PSQ and the columns to 200 projects per program. After obtaining the allocations for FY 2010, a sample of projects will be selected from each sampling cell (program type/PSU combination) with probabilities proportional to the number of households. As in previous years, our methodology will allow PHA-administered Section 8 projects to be selected more than once, but Public Housing and owner-administered projects will be selected only once. The same PPS systematic approach used to select PSUs will be used to select projects. Projects will be sorted by program type, county and PHA prior to selection in order to assure diversity.

---

[1] Chromy JR. Sequential sample selection methods. In Proceedings of the Survey Research Methods Section, American Statistical Association, pp 401–406, 1979.

[2] Goodman R. and Kish, L. (1950) "Controlled Selection—A Technique in Probability Sampling" J. Americ. Statist. Assoc. 45, 350–372.

<u>Selection of Tenants</u>

The initial tenant sample will be approximately self-weighting because the measures of size used in selecting PSUs will not always correspond to the sum of the measures of size of projects within the PSU. The term *self weighting* refers to a sample where all units being samples (in this case, households) will have the same weight assuming that the frame is accurate and 100 percent response is achieved. In addition, the number of occupied units found in a project may not correspond to the number of units listed in the frame. To compensate for this issue, we will make individual decisions by project once the project is sampled and its real size determined.

Consider the initial theory behind the sample. Let f be the fixed sampling rate desired for all tenants in the nation. Let p be the overall probability that a project with N tenants is selected. The needed number of tenants to be sampled (n_j) from the project to equalize weights is given by: $n_j = fN/p$. (We note that n_j may be greater or less than n, the desired fixed sample size.) As a practical matter, project sample size will not be permitted to vary in accordance with this formula, as this would create highly disparate interviewer workloads. It will, however, be allowed to vary if more than a two to one ratio between projected and actual weight is discovered.

Because the selection of tenants will be completed at the PHA/project site, the sampling procedures need to accommodate a variety of possible situations related to the availability of tenant lists. Some lists are computer generated and include optimum information; other lists are manually prepared by project staff and include minimal information. Interviewer procedures will provide instruction on how to select the sample and ORC Macro headquarters staff will be available to provide sampling assistance to the field interviewers by telephone.

<u>Weighting</u>

The probability of selection of a tenant will be the product of the following:

1) The probability of selection of the PSU.

2) The probability of selection of the sub-PSU when the PSU was divided.

3) The probability of selection of the project from the set of projects in the PSU. This is the probability described in the appendix, but capped by 1.0 for tenant-based Section 8 projects.

4) The probability of selection of the tenant from the set of in-scope tenants in the project—this is the total number of tenants sampled from the project divided by the estimated number of tenants in scope. The estimate is obtained by multiplying the total number of tenants by the proportion of tenants selected who are in scope. As an example, if a total of six tenants are reviewed to find four tenants who are in scope, one is out of town and one is no longer subsidized, though his name is still in the list, then the estimate would be 120x(5/6)=100 tenants.

The four probabilities will be multiplied together to form the preliminary weight. The weights will then be adjusted to sum to estimates of the national total of tenants in each program. The

final step will be trimming the weights. Extreme weights will be reduced and the weights will be re-adjusted so that they sum to the same national totals.

3. *Maximization of response rates*

Describe methods used to maximize the response rate 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.

Two types of non-response may effect this data collection: that by PHAs/owners and tenants.

PHAs/owners

Participation by selected PHAs/owners is mandatory such that their contracts with HUD require their participation in studies of this type. In an effort to ensure PHA/owner participation, the initial mailing is conducted using an overnight delivery service to catch their attention. PHAs/owners are given a date by which the information is needed and if that time elapses, follow-up telephone calls are made to obtain the needed information. If further follow-up is required, a list of the non-responsive PHAs/owners are provided to HUD and contacted by them as well. Appendix B contains study letters that are provided to PHAs/owners at the outset of the study (i.e., Phase I) and again in Phase IV.

Tenants

Participation by selected tenants is mandatory; refusal to participate could result in their termination of assistance. Field interviewers will make at least four in-person contacts with the tenant to conduct interviews with individuals who try to evade the interview. Appendix C contains the letter that is provided to tenants regarding this study. In addition, the following letter is occasionally used to encourage tenant participation.

Tenant Encouragement Letter

[Date]

Dear

On [date] we provided you with a letter from the Department of Housing and Urban Development (HUD) which explained the study ICF Macro is conducting for HUD, and informed you that you have been randomly selected to participate in this study. Since then, our field interviewer has been attempting to get in touch with you to schedule an interview.

HUD and the Federal Office of Management and Budget (OMB) have determined that persons who receive housing assistance are ***required to participate*** in this study. For your information, the OMB clearance number for this study is 2528-0203. **Failure to participate is a basis for terminating housing assistance**. Your local HUD office has been informed of, and is assisting with, this study.

It is essential that you contact us immediately to schedule an appointment for an interview. If you do not contact us by [date], we will be forced to report your lack of cooperation to HUD. Please call the telephone number identified below to schedule your appointment with the field interviewer directly. If the field interviewer is not available, call the supervisor listed below for assistance.

The purpose of the study is to learn more about the types of errors that occur during determinations of eligibility and tenant rents. This information will be used to meet Congressionally mandated reporting requirements related to the accuracy of rent calculations. The interview will take from 40-60 minutes. Information collected by this study will be reported as statistical summaries; however, individual information is shared with HUD headquarters and may be made available to those normally responsible for your income and rent determinations.

If you have any general questions about the study, please call me at the toll free number listed below. If you have questions about our authorization to conduct this study, you may call Dr. Yves Djoko, the government project office, at 202-402-5851.

Thank you for your cooperation with this study.

Sincerely,

Laura Webb
Survey Manager

Field interviewer:			
	Name		Phone Number
Supervisor:			877 - 392 - 9776
	Name		Toll Free Number
Use this ID # when calling:			Affix C/P/C label here

4.   *Tests of procedures or methods*

Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions of 10 or more individuals.

Previous iterations of this data collection serve as the pretests for this data collection effort. As mentioned previously, similar studies have been conducted in 2000 (data was collected for actions taken in 1999 and early 2000) and enhanced for the FY 2003 through FY 2009 studies. Before each data collection cycle, all changes or enhancements to the study are tested in an in-house procedure that evaluates the administrative and computer systems-related aspects of the study. Prepared case examples (those used in training our field interviewers) are abstracted and entered into our data collection system. Additionally, mock household interview data is entered into our data collection system and all associated administrative paperwork is created and processed. Finally, tracking reports are produced to determine that our reporting system is in place and accurate.

5.   *Individuals consulted on statistical aspects of design*

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.

ICF Macro Staff—Design and Data Collection

- Mary K. Sistik, Project Director, (301) 572-0488
- Dr. Sophia Zanakos, Deputy Project Director, (301) 572-0239
- Dr. Pedro Saavedra, Senior Sampling Statistician, (301) 572-0273