# Part B
# Statistical Methods

HUD has contracted with Concentrance Consulting Group to conduct a study aimed to assess the foreclosure rates among ADDI and HOME program downpayment assistance recipients. For each of the years, 2000 through 2005, the study design aims to produce a sample of homebuyers from the combined HOME and ADDI programs, in which each homebuyer has an equal chance of being selected. The sampling strategy is described below.

## B1 Respondent Universe and Sampling Method

While the survey seeks information on homebuyers assisted by the ADDI and HOME programs, the respondent universe are the PJs that administer these programs. The respondent universe in each year from 2000 to 2005 consists of all PJs that assisted at least one homebuyer in that year and that were eligible to receive an allocation of ADDI funding in Fiscal Year 2004. PJs are eligible for ADDI funding if they have a total population of at least 150,000 and would receive an allocation of $50,000 or more based on the ADDI Allocation Formula. Fiscal Year 2004 had the largest allocation of funding for the ADDI program, so it is the year when the largest numbers of PJs were eligible for funding. The total universe of PJs in the respondent universe ranges from a low of 302 in 2000 to a high of 371 in 2005.

The sampling approach was designed to minimize the number of PJs that were selected to provide information while also meeting goals related to the variance of the estimated foreclosure rate (described more below). To accomplish these goals the design uses a sample of 33 PJs in each of the 6 years, and separates the PJs into three groups according to the number of buyers assisted in the particular year:

- Group 1: Over 700 buyers;
- Group 2: From 51 to 700 buyers; and
- Group 3: 50 or fewer buyers.

PJs that belong to Group 1 are selected with certainty; each of them will be asked to provide information on a random sample of 50 buyers selected from its total list of assisted homebuyers for that year. From the PJs that belong to Group 2 a sample of PJs is selected with probability proportional to PJ size (PPS), where the measure of size is

the number of buyers; as with Group 1, each of them will be asked to provide information on a random sample of 50 buyers selected from its total list of assisted homebuyers for that year. From among the PJs in Group 3 a simple random sample of PJs is selected, with each selected PJs asked to provide data on all the buyers that it assisted in that year. The number of PJs sampled in Groups 2 and 3 is proportional to the number of total PJs in these two groups. For example, if the PJs in Group 2 account for one third of the total PJs in Groups 2 and 3 in a given year, then one third of the sample PJs will be from Group 2 and two thirds will be from Group 3. In sampling PJs for Groups 2 and 3, PJs are first sorted by type of PJ (state, city, county, or consortium of local governments) and by region of the country to be sure the selected sample provides reasonable coverage in each of these dimensions. A systematic random sample is then drawn from these two groups for each year.

The resulting sample includes 129 PJs, including 92 that were sampled once, 22 that were sampled twice, 5 that were sampled 3 times, 6 that were sampled 4 times, 2 that were sampled 5 times, and 2 that were sampled in all 6 years. However, since the sample is limited to a maximum of 50 buyers per year, the largest number of buyers any one PJ will be asked to report on is 300. Most of the PJs sampled for 4 or more different years were states. In these cases PJs will be housing finance agencies that are expected to have much of the information available electronically. The total number of buyers sampled in each year ranges from 983 to 1,219.

The sample has been designed to provide an estimate of the foreclosure rate with a standard error of 1 percentage point. In selecting the sample size, we have assumed that the foreclosure rate in the HOME program is 5 percent. This is a conservative estimate of the foreclosure rate based on findings from an earlier study of the HOME program.[1] If a simple random sample of assisted homebuyers in the HOME and ADDI programs were selected, we would need a sample size of 475 buyers to achieve this level of accuracy. However, the sample sizes must be increased to account for both non-response and the sample design. We have assumed a 90 percent response rate based on our experience in previous surveys of PJs to collect additional data on HOME program operations in which we have achieved response rates above 90 percent. The sample size also has to be increased to account for the fact that the sample of buyers is clustered

[1] Study of Homebuyer Activity Through the HOME Investment Partnerships Program, U.S. Department of Housing and Urban Development, January, 2004

within a sample of PJs. Even after taking into account these factors, the sample sizes for each year are more than sufficient to achieve the desired level of accuracy of the estimates. For example, in accounting for both non-response and the clustering of the sample, the effective sample size for the year 2000 is 587, which is much larger than the sample of 475 that is needed to achieve a standard error of 1 percentage point. Thus, we can have a higher non-response rate or greater degree of clustering in the results for assisted homebuyers within a PJ and still achieve the desired level of accuracy.

Finally, the sample sizes have also been increased to account for the sample design

## B2 Information Collection Procedures

The research team will contact all PJs by email using the contact information on HUD's website for HOME program contacts (ADDI programs are managed by the same personnel). The initial contact will introduce the project using the letters from HUD and Concentrance. Attached to the email will be the two survey instruments:

> Questionnaire on General ADDI/HOME Program Characteristics (in Microsoft Word the contents of which are shown in Appendix B)
> Data Verification and Collection of ADDI/HOME Program Participants (in Microsoft Excel the contents of which are shown in Appendix A)

The two instruments have been designed to be self-explanatory; however, the research team will have staff available to assist PJs with questions. PJs vary in size and technological sophistication and we expect some to be capable of creating data extracts from their database systems. The research team will encourage any such time-saving methods of data collection and transmission and will be accommodating of PJs' technical requirements.

Upon receipt of the data, the research team will examine the responses for completeness and face validity. When there are questions the team will follow up with the PJ by email and telephone.

## B3 Maximizing Response Rates

Data collection will take place over a twelve week period. The research team will maintain a record of all PJs surveyed. That record will include whether the PJ has responded to the survey, whether that response was complete, and whether the PJ was working with the team in developing its response. One week after the initial email distribution of the survey the research team will send a follow-up email with the letters and survey instruments to those PJs that had not responded. After another week, those PJs that have not responded will be contacted by telephone and will be encouraged to respond. We will continue to contact those PJs that have not responded to the survey over the remaining weeks of the data collection period. In addition, research team members will offer assistance to PJs in the development and transmission of their responses.

As participants in the ADDI and HOME programs, respondents will have an incentive to participate in this survey to help learn about program outcomes and help build Congressional support for the program. In previous studies of the HOME program, we have achieved very high response rates from PJs to requests for information about program characteristics and updating data on individual households assisted.

## B4 Tests of Procedures or Methods

Initial design of the data collection process was informed by two efforts. First we consulted a previous study performed by one of HUD's contractors to obtain essentially the same information from a sample of PJs.[2] Second we conducted exploratory phone calls with a small sample of PJs to inquire about the nature of information maintained in their files and the degree of effort that would be required to gather this information. The specific data collection instruments will be pre-tested during June and July 2007 to refine the instruments to improve respondents' ability to provide the requested information.

## B5 Statistical Consultation and Information Collection Agents

HUD has contracted with Concentrance Consulting Group Inc. and Abt Associates Inc. to design and conduct the data collection for this study.

[2] Study of Homebuyer Activity Through the HOME Investment Partnerships Program, U.S. Department of Housing and Urban Development, January, 2004

The data collection procedure will be similar to that used by these firms in other efforts to collect data on homebuyers from PJs and lenders.  The sampling plan was developed by Dr. David Hoaglin, Statistician, of Abt Associates Inc