

**Supporting
Statement for
Paperwork
Reduction Act
Submission**

**Pre-Purchase
Homeownership
Counseling
Demonstration and
Impact Evaluation**

Contract # R-CHI-01108

Part B

Supporting Statement for Paperwork Reduction Act Submission

Table of Contents

B. Statistical Methods..... 1

- B.1 Potential Response Universe..... 1
- B.2 Statistical Methods..... 1
 - B.2.1 Sampling Plan..... 1
 - B.2.2 Analysis Plan..... 4
 - B.2.3 Justification of Level of Accuracy..... 5
 - B.2.3 Unusual Problems Requiring Specialized Sampling Procedures..... 8
 - B.2.4 Any Use of Periodic (less frequent than annual) Data Collection Cycles to Reduce Burden..... 8
- B.3 Maximizing Response Rates..... 8
- B.4 Tests of Procedures or Methods..... 8
- B.5 Statistical Consultation and Information Collection Agents..... 9

B. Statistical Methods

B.1 Potential Response Universe

The Pre-Purchase Homeownership Counseling Demonstration and Impact Evaluation has three potential response universes: (1) study participants; (2) participating lenders; (3) participating counseling agencies. Each of the data collections covered by this submission relate to the universe of study participants, which is defined by a previous submission (Control #2528-0293). The study design is be a randomized experiment. We are currently in the process of recruiting up to **6,000** low-, moderate-, or middle-income (LMMI), first-time homebuyers from **3 national lenders** across **28 jurisdictions** nationally.

The study will seek to collect the Interim Survey and tracking letter from all study participants who enroll in the study. The study team will also collect the counseling agencies' service tracking data for all study participants who receive services from the study's participating counseling agencies and the lenders' origination and servicing data from any study participant who originates a home purchase loan with one of the study's participating lenders. The study will seek to collect consent from each co-borrower listed on a home purchase mortgage loan of a study participant. The universe of respondents for the co-borrower consent collection therefore includes all co-borrowers of all enrolled study participants.

The set of focus group participants will be drawn from the set of study participants in three of the study's sites. The potential respondent universe is therefore the full set of study participants in these three sites. The statistical methods section describes the process used to select the sites and recruit focus group participants.

B.2 Statistical Methods

The evaluation requires sampling of jurisdictions, housing counseling agencies, lenders, study participants, implementation of focus groups, and lender and housing counseling staff interviewees. Below we discuss the sampling plan for each of these groups. The study's prior submission (Control #2528-0293) outlines the statistical methods used in the initial phase of this study

B.2.1 Sampling Plan

Jurisdictions

Participant enrollment is implemented in 28 jurisdictions. The primary criterion for selecting jurisdictions was the expected volume of eligible study participants. However, the process of selecting the sample of jurisdictions also considered the ability of local in-person housing counseling agencies to provide services to enrolled participants throughout the jurisdiction. Exhibit B.2.1 lists the study's 28 sites.

Exhibit B.2.1 28 Study Jurisdictions

List of 28 Study Sites
Atlanta-Sandy Springs-Marietta, GA
Boston-Cambridge-Quincy, MA-NH
Chicago-Joliet-Naperville, IL-IN-WI

Dallas-Fort Worth-Arlington, TX
Detroit-Warren-Livonia, MI
Houston-Sugar Land-Baytown, TX
Las Vegas-Paradise, NV
Los Angeles-Long Beach-Santa Ana, CA
Miami-Fort Lauderdale-Pompano Beach, FL
Minneapolis-St. Paul-Bloomington, MN-WI
New York-Northern New Jersey-Long Island, NY-NJ-PA
Orlando-Kissimmee-Sanford, FL
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD
Phoenix-Mesa-Glendale, AZ
Portland-Vancouver-Hillsboro, OR-WA
Raleigh-Cary, NC
Riverside-San Bernardino-Ontario, CA
Sacramento-Arden-Arcade-Roseville, CA
San Antonio-New Braunfels, TX
San Diego-Carlsbad-San Marcos, CA
San Francisco-Oakland-Fremont, CA
San Jose-Sunnyvale-Santa Clara, CA
Seattle-Tacoma-Bellevue, WA
St. Louis, MO-IL
Stockton, CA
Tampa-St. Petersburg-Clearwater, FL
Virginia Beach-Norfolk-Newport News, VA-NC
Washington-Arlington-Alexandria, DC-VA-MD-WV

Note: The 28 jurisdictions listed above are ordered alphabetically.

Lenders

The study team has partnered with three national lenders—Bank of America, Citibank, and Wells Fargo—for the implementation of this study. The participating lenders were chosen purposively, based on their volume of loan originations and willingness to partner for this study.

HUD-Approved Housing Counseling Agencies

The study team has partnered with ClearPoint, eHome America, and NeighborWorks America to provide housing counseling services to study participants. Each housing counseling agency was selected purposively, based on the adherence of the housing counseling curricula and delivery mechanisms to the study's defined interventions: remote and in-person counseling services.

Study Participants

The universe of potential study participants is defined in the previous submission (Control #2528-0293) and includes all eligible customers who communicate with a study participating lender during the enrollment period. The sample of study participants is a 100 percent sample of the universe,

excluding only those customers who cannot be contacted, refuse the initial offer of study participation, do not meet the study's eligibility requirements, or do not provide consent.

No statistical methods are involved in identifying study participants for completion of the interim survey. The study will administer the interim survey to each study participant in the potential respondent universe.

Co-Borrowers

No statistical methods are involved in identifying the co-borrowers for consent collection. The study will seek to collect consent from each co-borrower in the potential respondent universe.

Focus Groups

No statistical methods are involved in determining the set of study participants eligible for the focus groups. The set of focus group participants reflect a convenience sample of study participants—and not a statistically-representative sample of study participants. The study team will identify the study sites and study participants for the focus groups using the criteria described below.

Selection of study sites:

After the first 12 months of study enrollment, the study team will examine the distribution of enrolled participants and select three locations for conducting focus groups based on the following criteria:

- Number of enrolled study participants in each site;
- Diverse geographic regions—Northeast, Southeast, and West;
- Housing market diversity: for example, extent to which affected by the housing downturn, extent of recovery from same;
- Number of in-person counseling agencies to choose;
- Number of study participants referred by each of the participating lenders (i.e. the potential for all three lenders to be represented); and

The study team will purposively select three study sites that have a sufficient number of study participants to recruit participants for the focus groups.

Selection of focus group participants:

After the selection of focus group locations, a list of potential focus group participants will be generated to recruit 8 to 12 study participants per focus group based on the following criteria:

- Whether they were assigned to in-person or remote education and counseling or had a preference of in-person or remote education and counseling from assignment to the choice group;
- Whether they a) didn't take up the education and counseling offered; b) took up the education and counseling but didn't complete it; or c) completed the education and counseling;
- Which lender referred them to the study, to ensure that all three lenders are represented;
- Which in-person counseling agency they selected (if counseling was initiated), in locations with multiple participating agencies, with the objective of including participants from multiple agencies;

- Where they are in the home purchase process - whether they have purchased a home or not (to explore whether HEC influenced this decision); and

In each of the three selected sites, we will conduct a focus group with the four sub-groups outlined below.

1. *In-person Non-completers*: Study participants offered in-person pre-purchase education and counseling services through a housing counseling agency who either a) chose not to take up the pre-purchase education and counseling services; or b) started the pre-purchase education and counseling services but did not complete them;
2. *Remote Non-completers*: Study participants offered on-line pre-purchase education and telephone counseling services who either a) chose not to take up the pre-purchase education and counseling services; or b) started the pre-purchase education and counseling services but did not complete them;
3. *In-Person Completers*: Study participants who completed the pre-purchase education and counseling services through a housing counseling agency; and
4. *Remote Completers*: Study participants who completed the pre-purchase on-line education and telephone counseling services.

The final sample of focus group participants will include 8-12 participants in each of these four focus groups in three study sites (i.e. 12 total focus groups with an estimated 120 total respondents).

B.2.2 Analysis Plan

The previous submission explains the implementation analyses, analysis and documentation of random assignment, and the baseline analyses in detail. This submission discusses the impact analyses which will be conducted based on the follow-up data the study collects and the analysis to be conducted based on the data collected from the focus groups.

Impact Analyses

The primary impact analyses for the evaluation will be conducted using outcomes collected from administrative data and the Interim Survey. These analyses will rely on the data from the baseline data described by the study's prior submission for baseline measures of time-varying outcomes and to define covariates and subgroups for the impact analyses. For example, subgroups may be defined with respect to borrower personal characteristics (from the baseline survey); financial characteristics (from the lender intake form and baseline survey); neighborhood characteristics (which requires knowledge of the borrower's address from the baseline survey); and dosage levels (from the counseling services data).

The experimental evaluation design balances the pre-existing characteristics of the study participants assigned to each of the treatment groups and to the control group. Therefore, the impact of being offered a treatment can be estimated by simply comparing mean outcomes of those offered each treatment, relative to those in the control group. For example, if we find that study participants who receive Online Education + Telephone Housing Counseling have higher homeownership rates than control group members, we can interpret the difference in homeownership rates as the causal impact of receiving Online Education + Telephone Housing Counseling. We will estimate this difference in means with multivariate regression, as controlling for additional study participant characteristics will increase precision. This same approach can be used for each outcome of interest.

The randomization of participants to treatment and control groups produces an estimate that reflects the impact of being offered housing counseling—referred to as the Intent to Treat (ITT) estimate. Commonly, adjustment of the ITT allows estimation of the impact of housing counseling on those who actually complete the prescribed housing counseling—referred to as the Treatment on the Treated (TOT) estimate. The two will differ because of incomplete take-up (i.e., some people who are offered a treatment will choose not to take it up). We will compute TOT estimates by two-stage least squares regression.

As an extension to the analyses described above, we will perform subgroup analyses based on post-intervention borrower characteristics. For instance, we will compare outcomes for participants who fully complete the housing counseling intervention, relative to those who complete only part, relative to those who do not initiate housing counseling at all. To capture the effect of varying “dosage,” we will use a propensity score matching approach that identifies subgroups based on their pre-existing traits. This allows us to estimate an unbiased impact on these dosage-defined subgroups of interest.

Focus Group Analysis

The focus groups will contribute to the Implementation Analysis described by the previous submission, replacing the role of the study participant interviews. The focus groups will examine participants’ experience with the study, including enrollment, monitoring, and any interactions with counseling agencies. In particular, the focus groups will examine why study participants in each treatment group decided to complete or not complete pre-purchase counseling and education. Focus group discussions will be transcribed and then analyzed using NVivo software to examine common experiences and opinions of study participants.

Similar analysis techniques will also be used to analyze the focus group responses to learn about study participants’ experiences with the two types of pre-purchase education and counseling offered. While the study’s design does not contain a robust qualitative study, analysis of the focus group transcripts using NVivo software will be used to supplement the findings of the quantitative impact analysis by extracting qualitative information about how the pre-purchase homeownership education and counseling services may have affected study participants’ behaviors.

B.2.3 Justification of Level of Accuracy

This section presents minimum detectable effects for the impact analyses discussed in the previous section. Minimum detectable effects (MDEs) are the smallest true effects of an intervention that researchers can expect to detect as statistically significant when analyzing samples of a given size. The calculation of MDEs assumes that the recruitment process successfully enrolls 6,000 study participants and that the response rate to the final follow-up survey reaches 75 percent of the initially recruited sample. The anticipated sample size for each experimental treatment arm is as follows: 2,550 (43 percent) control group; 1,725 (29 percent) remote counseling group; 1,183 (20 percent) choice group; and 542 (9 percent) in-person counseling group. However, note that we do not plan to use the in-person group sample for confirmatory impact analyses. Exhibit B.2.3 presents the MDE values based on these anticipated sample sizes. The MDEs in Exhibit B.2.3 are calculated using standard assumptions regarding power, significance level, and R-squared values. These assumptions are:

- 80 percent power;

- A two-tailed test at the .10 significance level; and,
- An R-squared value of .10.

The MDE and Impact Size figures in Exhibit B.2.3 are shown for dichotomous outcomes—variables like foreclosure that take on a value of 0 or 1. The initial column of estimates presents the MDE and Impact Size figures when the mean value of the outcome is .10 (i.e., 10 percent of control group members experience foreclosure). This value also applies if the mean value is .90. The final two columns present similar values when the mean value of the outcome is .30 (or .70) and when it is .50.

Panel A presents MDE values for comparing the pooled sample of remote and choice counseling group members to the control group. For a dichotomous outcome variable with a mean control group value of .5, the MDE estimate when there is a 50 percent rate of no shows and crossovers (our expected rate) is 7.7 percentage points. This value implies that the average value for the treatment group would have to increase to 57.7 percent for the impact of counseling to be significant. The required average value is 56.5 and 59.7 percent if 40 percent and 60 percent of sample respondents are no shows.

The actual impact on treated participants must be proportionally larger than the ITT estimate if some treatment group participants do not receive treatment. The row labeled ‘Impact Size’ shows the necessary size of the actual impact on the set of participants who take up counseling. The impact size increases non-linearly as no shows increase. For an outcome variable with a control group mean of .5, the actual impact of housing counseling when 50 percent of treatment group members do not complete housing counseling is 15.5 percentage points—compared to 7.7 percentage points for the MDE. If 40 percent of treatment group members do not complete housing counseling, the actual impact of housing counseling on counseled individuals must be 10.8 percentage points. If 60 percent of treatment group members do not complete housing counseling, the necessary impact size is 24.2 percentage points.

Comparing Types of Counseling: In addition to comparing the treatment groups to the control group, it may be informative to compare the relative impact of different counseling types by comparing across treatment groups. Panel 2 of Exhibit B.2.3 shows the MDE estimates for comparing the remote and choice treatment groups. For a dichotomous outcome variable with a mean control group value of .5, the MDE estimate when there is a 50 percent rate of no shows and crossovers is 11.6 percentage points. This value implies that the average value for the treatment group would have to increase to 61.6 percent for the impact of counseling to be significant. The required average value is 59.7 and 64.5 percent if 40 percent and 60 percent of sample respondents are no shows.

Subgroup Analysis: The impact analyses will include subgroup analysis for key exogenous subgroups of interest, such as credit score and income. For these measures, the study will define subgroups that divide the sample approximately in half, defining subgroups of ‘high’ and ‘low’ credit scores (or incomes). Panel 3 of Exhibit B.2.3 shows the MDE estimates for these subgroups, assuming that the subgroups evenly split the sample into two groups of 3,000 study participants. For a dichotomous outcome variable with a mean control group value of .5, the MDE estimate when there is a 50 percent rate of no shows and crossovers is 11.5 percentage points. This value implies that the average value for the treatment group would have to increase to 61.5 percent for the impact of counseling to be significant. The required average value is 59.6 and 64.4 percent if 40 percent and 60 percent of sample respondents are no shows.

The analysis will also examine endogenous subgroups and exogenous subgroups with small sample sizes. Such analyses will be exploratory in nature, and the study is not designed to confidently detect impacts for these subgroups.

Exhibit B.2.3. MDE Estimates for Impact Analysis of Dichotomous Follow-Up Outcomes

	Mean Outcome for Control Group		
	.10	.30	.50
Panel A: Pooled Treatment/Control Comparisons			
40% No Show & Crossover:			
MDE	0.038	0.059	0.065
Impact size	0.064	0.099	0.108
50% No Show & Crossover:			
MDE	0.046	0.071	0.077
Impact size	0.093	0.142	0.155
60% No Show & Crossover:			
MDE	0.058	0.089	0.097
Impact size	0.145	0.222	0.242
Panel B: Remote Treatment/Choice Treatment Comparisons			
40% No Show & Crossover:			
MDE	0.058	0.089	0.097
Impact size	0.097	0.148	0.161
50% No Show & Crossover:			
MDE	0.070	0.106	0.116
Impact size	0.139	0.213	0.232
60% No Show & Crossover:			
MDE	0.087	0.133	0.145
Impact size	0.218	0.333	0.363
Panel C: Pooled Treatment/Control Comparisons for Subgroup Analysis			
40% No Show & Crossover:			
MDE	0.057	0.088	0.096
Impact size	0.095	0.147	0.160
50% No Show & Crossover:			
MDE	0.069	0.106	0.115
Impact size	0.138	0.212	0.231
60% No Show & Crossover:			
MDE	0.087	0.132	0.144
Impact size	0.216	0.331	0.361

--	--	--	--

$\alpha = 0.10$; two-tailed test. For Pooled Treatment/Control Comparisons in Panel A, baseline remote treatment group members = 1,725; baseline choice treatment group members = 1,183; baseline control group members=2,550. For Remote/Choice Comparisons in Panel B, baseline remote treatment group members = 1,183; baseline choice treatment group members = 1,183. For subgroup analyses in Panel C, sample sizes are half the size of those used for Pooled Treatment/Control Comparisons in Panel A (assumes equal 50/50 split across two subgroups). For all analyses, follow-up sample is assumed to be 75 percent of baseline sample size.

B.2.3 Unusual Problems Requiring Specialized Sampling Procedures

There are no unusual problems associated with this sample.

B.2.4 Any Use of Periodic (less frequent than annual) Data Collection Cycles to Reduce Burden

Not applicable to this study.

B.3 Maximizing Response Rates

The study team will undertake extensive efforts to maximize the response rates for the collected data. The collection of lenders' loan origination and performance data and counseling agencies' service tracking data will continue the processes begun during the initial phase of the study, collecting the data for each study participant with administrative data from these sources.

The collection of the Interim Survey and co-borrower consent each include layered data collection strategies in order to maximize the response rate. The Interim Survey will use a telephone plus field approach to collect the survey responses. The telephone call center will conduct interviews with all study participants that can be reached by telephone. For those study participants who are not reached by telephone, the data collection team will conduct in-person interviews using a field data collection approach.

To collect co-borrower consent, the study team will first mail an advance letter prior to the start of telephone calls for the Interim Survey, which will occur 12 months after enrollment. The advance letter will include a copy of the co-borrower consent form and a postage-paid return envelope for each co-borrower identified during the baseline survey that did not provide consent previously.

The first follow-up telephone survey occurs 12 months after enrollment and will ask the study participant to confirm whether they purchased a home with one or more co-borrowers. If any co-borrower has not returned the copy of the co-borrower consent form provided in the advance letter, the interviewer will ask the study participant whether the co-borrower is available, and attempt to collect co-borrower consent verbally or arrange a time to callback to collect the consent. The study team will then conduct outbound calls to any co-borrower identified during the first follow-up telephone survey who did not provide consent during the Interim Survey.

For the focus groups, the study team will take several steps to ensure the participation of 8-12 study participants in each group. First, the selection of study sites will directly consider the number of potential respondents, purposively selecting sites where we are likely to recruit the target number of focus group participants. Second, the focus groups will be held in a location that is centrally and conveniently located for the study participants. Third, the study team will provide reminders and

directions to study participants prior to the scheduled focus groups, encouraging participation and asking participants who cannot make the focus group to contact us so that we can find a replacement.

B.4 Tests of Procedures or Methods

The study team relies on review of each instrument by Abt Associates and other study team staff, HUD personnel, and the study's advisory panel to ensure that the instruments are clear, flow well and are as concise as possible and collect the data necessary for analysis. Before the Interim Survey is implemented in the field, the study team will conduct up to nine pre-tests of the data collection instrument.

B.5 Statistical Consultation and Information Collection Agents

HUD's Office of Policy Development and Research will work with the contractor, Abt Associates, to conduct the proposed data collection. Marina L. Myhre, Ph.D., a Social Science Analyst in HUD's Office of Policy Development and Research, Program Evaluation Division, serves as Government Technical Representative (GTR). Her supervisor is Ms. Carol Star. Dr. Myhre and Ms. Star can be contacted at (202) 402-5705 and (202) 402-6139, respectively. The study's Principal Investigators are Dr. Laura Peck and Dr. Jonathan Spader from Abt Associates and Dr. Roberto Quercia from the Center for Community Capital at the University of North Carolina at Chapel Hill. Dr. Peck can be reached at (301) 347-5537. Dr. Quercia can be reached at (919) 843-2493. Dr. Jonathan Spader also serves as the study's Project Director. Dr. Spader can be reached at (301) 347-5789.