Impact of Housing and Services Interventions for Homeless Families

Supporting Statement for Paperwork Reduction Act Submission- Part B

Contract No. C-CHI-00943
Task Order: CHI-T0001

Follow-up Survey Data Collection

June 21, 2011
Revised September 30, 2011
Revised October 14, 2011
Revised February 26, 2012

Prepared for:
Elizabeth Rudd
Anne Fletcher
U.S. Department of HUD
Office of Policy Development and Research
451 Seventh Street SW Room 8140
Washington, DC 20410

Prepared by:
Abt Associates Inc.
4550 Montgomery Avenue
Suite 800 North
Bethesda, MD 20814-3343

Follow-Up Survey Data Collection – Draft

Table of Contents

Part B: Collection of Information Employing Statistical Methods

B.1 Identification of Appropriate Respondents

B.1.1 Sample Recruitment and Random Assignment

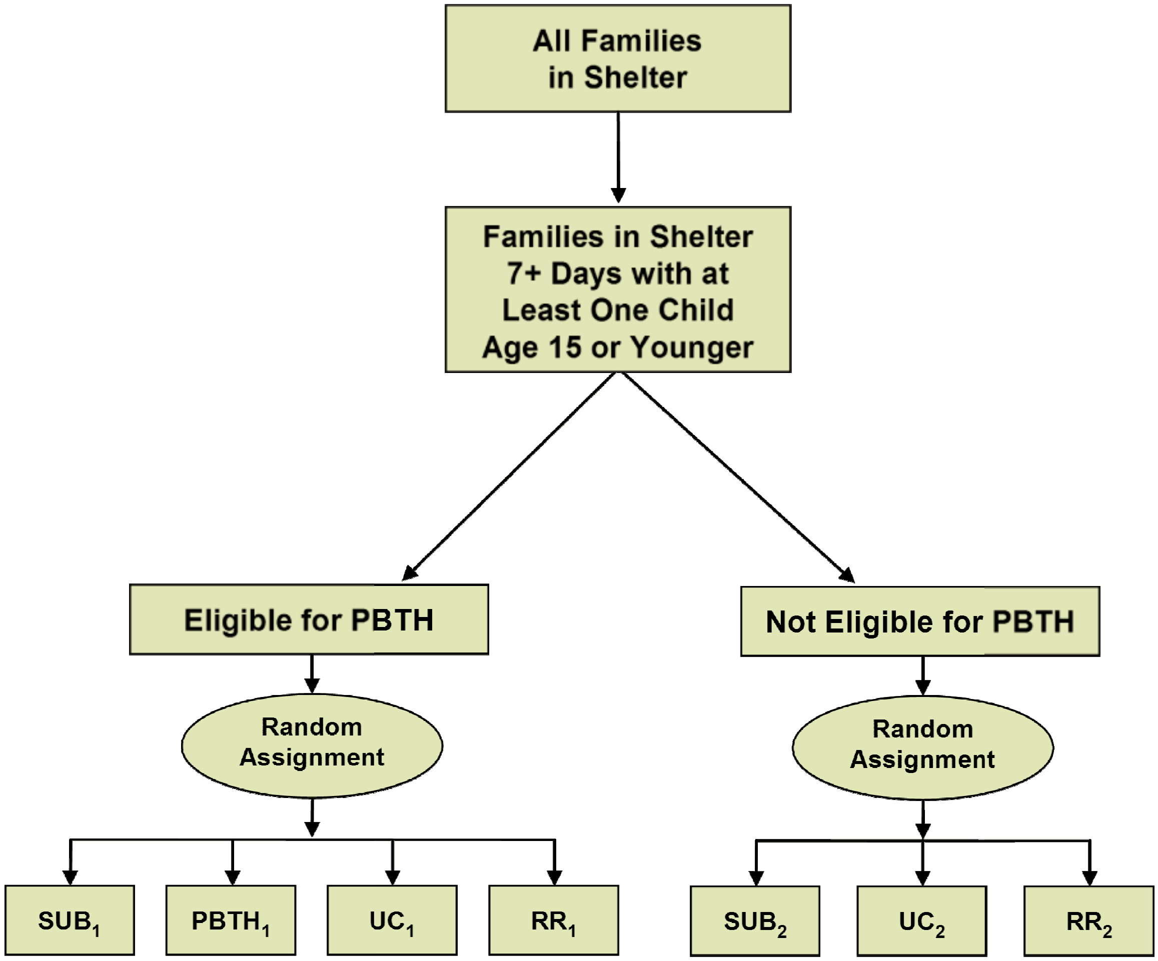
The study design is a randomized experiment. We recruited 2,305 homeless families who had been in emergency shelter for at least 7 days across 12 sites. We excluded families who leave shelter in less than 7 days because the more intensive interventions considered in this study are not considered appropriate for families with such transitory needs. We expect shelters to continue to provide all services and referrals they ordinarily provide to help families leave shelter up until the point of random assignment. Families are then assigned, as close to the 7-day mark as is feasible, to the Subsidy Only (SUB), Community-Based Rapid Re-housing (CBRR), Project-Based Transitional Housing (PBTH), or Usual Care (UC) interventions. Recruitment began in September 2010 and ended January 31, 2012.

Our design also recognizes that not all families are eligible for all interventions. Consistent with this consideration, families are screened as to their eligibility for each specific service provider in their site, prior to random assignment. Families are randomly assigned only to interventions for which they appear eligible, based on their responses to the screening questions. As long as one provider within each experimental intervention at a given site will accept a family with a particular profile, that family is considered eligible for that intervention. Exhibit B-1 shows the random assignment model used to allocate families to interventions, assuming all four interventions are included in the study design.

As shown at the top of the exhibit, the population of interest for this study is all families who have been in an emergency shelter for at least 7 days and who have at least one child 15 or younger. This restriction is included because child outcomes are important to the study, and we will not have a large enough sample to consider outcomes for youth who become young adults in the course of the follow-up period. Hence the age restriction of 15 and younger.

The study is not designed to capture the experiences of families who seek assistance directly from transitional housing programs without first entering emergency shelters. The design relies on emergency shelters as the point of intake for families in the study.

Exhibit B-1. Random Assignment Plan



In each site, this population was identified and randomly assigned in one of two ways. In sites where all such families qualified for assistance from at least one provider of each intervention and all interventions were available, families were randomly assigned to all four interventions, as shown in the left-hand stream of the diagram. In sites where some families are ineligible for all programs that make up a particular intervention, we will randomly assign those families only to the interventions for which they are eligible. The right-hand stream in Exhibit B-1 shows the random assignment design for families who are eligible to receive subsidies, with or without intensive services (interventions SUB and CBRR), but are not eligible to receive transitional housing (intervention PBTH) because no transitional housing provider in the site will accept them. This diagram and the resulting analysis can be generalized to the situation where some families are not eligible for other interventions, but for simplicity we illustrate the case where restrictions apply only to Transitional Housing. We assumed that all families were eligible for the emergency shelter (intervention UC). Note that both streams could be operative in the same site; i.e., families who are eligible for all interventions would be assigned as in the left-hand stream, while those who are not eligible for transitional housing would be assigned as in the right-hand stream. In a site with no transitional housing program, all families in that site would be randomly assigned to three interventions, as in the right-hand stream in Exhibit B-1.

As we describe below, this design assures that comparisons of interventions will involve well-matched groups in each intervention, even when some families are ineligible for a particular intervention program. The design thus assures that any observed differences in outcomes are caused by the differential treatment families receive, and not by any pre-existing differences among the families.

Although assignment to interventions was conducted at random, within interventions families need not be assigned at random to service providers that represent the intervention. Assignment was made instead on the basis of family characteristics, as is currently done. Thus, for example, if one or more of the transitional housing programs in a site specialized in families with a particular profile (only families with domestic violence issues, or only families where the mother has been clean and sober for some period), then among families randomly assigned to Transitional Housing, only those that fit that program were assigned to that service provider. If a site has vouchers available only to veterans, then among families randomly assigned to the SUB intervention, only families that include a veteran will be assigned to veteran housing. This preserves and studies programs as they currently operate.

Through a combination of the baseline data collected under the previously approved data collection, and the follow-up data we will collect under the data collection activity submitted here for approval, the design will provide rigorous experimental answers with sufficient statistical power for the following broad questions:

- What is the relative effectiveness of homeless interventions in ensuring housing stability of homeless families?
- Are the same interventions that are effective for short-term housing stability of homeless families effective for longer-term housing stability as well?
- What is the relative effectiveness of homeless interventions in ensuring the well-being of homeless parents and self-sufficiency of homeless families?
- Do some interventions promote family preservation and benefit children’s well-being, in particular, more than other interventions?

The overarching research question for this study is the extent to which housing and/or intensive services influence housing stability, family well-being, and other non-housing outcomes. The study design will provide empirical evidence on each of these effects, separately and in tandem. Many families leave shelter on their own, but little is known about what happens to them in terms of either residential stability or other outcomes. By including a Usual Care group that does not receive a dedicated subsidy or targeted intensive services, we will understand not only the impacts of interventions relative to no special services, but also whether interventions that explicitly address homelessness produce superior results to temporary shelter and the mainstream poverty assistance system but no additional specialized assistance. In this section, we describe the specific impact estimates that we will generate to answer these questions.

B.1.2 Universe of Households and Survey Samples

The study sample will comprise families, defined as at least one adult and one child, who experience homelessness, receive assistance at an emergency shelter, and remain in the shelter for at least seven days. Exhibit B-2 summarizes the definition and sample sizes for all of the random assignment groups.

Exhibit B-2. Definition and Size of Randomly Assigned Groups in the Family Options Study

Group	Intervention Definition	# Assigned per Site	Total # Assigned
SUB	Subsidy only; defined as deep, permanent housing subsidy that may include housing related services but no supportive services.	0-76	604
CBRR	Community-Based Rapid Rehousing: Time-limited housing subsidy that may also include housing-related services and limited supportive services	8-83	577
PBTH	Project-Based Transitional Housing: Time-limited housing subsidy coupled with supportive services	0-66	370
UC	Usual Care: Other assistance available in the community	21-81	754
	Total, all Intervention Groups	**58-281**	**2,305**

B.2 Administration of the Survey

B.2.1 Sample Design

The enrollment for the study is 2,305 families. All randomly assigned families completed a baseline interview, and all of them will be included in the participant tracking. For the follow-up survey, interviews will be attempted with all members of the research sample. Therefore, no sampling is required for the tracking or follow-up surveys.

Data to analyze the impacts of the housing and services interventions will come primarily from the follow-up survey, which is submitted for OMB review under this supporting statement. Key topics included in the follow-up survey are related to housing stability (incidence of homelessness in the follow-up period, use of shelter, type of housing situations); self-sufficiency (employment and earnings over the follow-up period, income and receipt of public assistance); family preservation (changes in family composition over the follow-up period, placement of children into foster care); adult well-being (physical and behavioral health); and child well-being (academic performance; school attendance, health, and behavioral health for a focal child, defined as one child, selected at random from among children age 15 years old or younger who resided with the family head at baseline.

B.2.2 Estimation Procedures

The rigor of this study comes from random assignment of families to different treatment “arms,” or conditions, within sites. In keeping with this design, we will compute impacts for each of the policy comparisons defined above on a site-by-site basis. Then we will pool the impact estimates across sites to calculate the overall study-level impact findings for each comparison. A covariate-adjusted regression model will be used to derive site-level impact estimates, and the site-level estimates will be combined using appropriate weights consistent with commonly used meta-analysis techniques.

To understand the estimation method for each site, consider two interventions q and r (e.g., project-based transitional housing and subsidy only), where we treat the first option (q) as the base case. For each site j, we will estimate the impact on an outcome Y (e.g. housing stability, earnings) for intervention r relative to intervention q by estimating the equation:

$$(1)\ Y_{ij} \quad = \quad \alpha_j^{q,r} + X_{ij}\beta_j^{q,r} + T_{ij}^{q,r}\delta_j^{q,r} + e_{ij}^{q,r}$$

for those families who could have been randomly assigned to both options q and r, and were assigned to one of them. In words, outcome Y for family i in site j is modeled as a constant α, adjustments for observed covariates[1] X with regression coefficients β, a dummy variable for assignment to intervention r where the corresponding regression coefficient δ gives the impact of intervention r (relative to intervention q), and a residual e. The overall impact, $\bar{\delta}^{q,r}$, would then combine impact estimates from all sites where the pairwise policy comparison could be made using the equation[2]:

$$(2)\ \bar{\delta}^{q,r} \quad = \quad \sum_{j=1}^{J} w_j\, \delta_j^{q,r}$$

where w_j is the weight to be applied to each site's impact.
There are several ways that weights may be established: equal weights on all sites, weighting proportional to the sample size in each site, or weighting inversely proportional to the variance of the impact estimate in each site. The choice of how to weight the impacts will be made during the evaluation analysis design phase in fall 2011.

The computation of the standard error for the overall impact $\bar{\delta}^{q,r}$ would involve the same set of weights used to calculate the overall impact estimate itself:

$$(3)\quad SE\left(\bar{\delta}^{q,r}\right) \quad = \quad \sqrt{\sum_{j=1}^{J} w_j^2\, VAR\left(\delta_j^{q,r}\right)}$$

---

1 The covariates will be variables constructed from the baseline survey data. These covariates will serve to improve precision of the impact estimates to the extent that they are related to later outcomes and they will adjust for chance differences in baseline characteristics between random assignment arms.

2 This equation for the overall impact is the weighted average of impacts from independent sites. Computing an overall impact in this manner is a method used in meta-analysis. Many meta-analyses first convert impacts to effect sizes by dividing the impact from each site by the standard deviation of the outcome in order to have a common metric across studies. We omit the conversion of impacts to effect sizes since the metric of impacts from the study sites will already be the same across sites, due to uniform data collection across all sites.

The estimation of equation (1) will utilize another set of weights called “analysis weights” that serve two main purposes. First, the weights will account for the probability of being assigned to the family’s intervention group. Second, the weights will adjust for survey non-response.

In addition to the estimation of overall impacts described above, the analysis will also estimate impacts for subgroups based on a “challenge” index, an “instability” index, and potentially other characteristics of interest. The challenge index will be constructed using data from the baseline survey on behavioral health and trauma. The instability index will be empirically derived from baseline predictors of subsequent instability in the usual care group. These indices may be used in one of two manners to examine how impacts differ depending on level of challenge or instability. First, the sample could be divided using cut-points on the indices which divide the sample evenly into high and low challenge subgroups and high and low instability subgroups. Impact estimates would then be separately estimated for each subgroup. Alternatively, the estimation method could directly incorporate the index. This second method would produce an estimate of how changes in the index affect the size of the impact, which is useful information for policy simulations. This way of examining the “moderating” effects of subgroup characteristics provides more statistical power to detect differences in impacts by moderating variables but requires assumptions about the functional form of the moderating relationship.

The impact analysis just described will involve a large number of hypothesis tests due to the inclusion of six impact comparisons, many outcome measures, and multiple subgroups or moderators. Testing such a large number of hypotheses heightens the danger of “false positives” arising in the analysis, i.e., of obtaining statistically significant impact findings where true impact is zero. This danger is called the “multiple comparisons problem”; the risk of false positives rises above the desired 5 or 10 percent chance as the number of hypothesis tests performed rises above one. To address the multiple comparisons problem we will separate the hypothesis tests into “confirmatory” tests and “exploratory” subsets. Only the most important outcomes will be included in the confirmatory group—a set to be decided during the evaluation analysis design phase in fall 2011. All other impact estimates, including all estimates for subgroups, will be considered exploratory. We will characterize findings of statistical significance for confirmatory outcomes as the *proven* impacts of the policies being compared, and findings of statistical significance for exploratory outcomes as merely suggestive of the impacts that *may have* occurred.

Additional analytic techniques will be needed to deal with missing data, investigate which program attributes contribute most to the impact of a given intervention approach (called “mediational analysis”) and to address the likelihood that some families will receive an intervention different than the one to which each was randomly assigned (the problem of “randomization non-compliance”). We will draw on state-of-the-art methodologies from other random assignment studies—including new and emerging methods from the literature and other Abt studies—to address these challenges. The particular methods adopted will be chosen during the evaluation analysis design phase under Task Order 3 in fall 2011.

B.2.3 Degree of Accuracy Required

The research team has estimated the minimum detectable effects for this evaluation that will be available through the impact analysis. The analysis of statistical power is presented here.

Power Calculations for Binary Outcomes

In this section, we consider statistical power to estimate impacts of interest. Specifically, we report minimum detectable effects (MDEs). MDEs are the smallest true effects of an intervention that researchers can be confident of detecting as statistically significant when analyzing samples of a given size. The power analyses are computed based on actual numbers of families assigned to the interventions and available for each pairwise comparison. These sample sizes differ somewhat from the planned design due to constraints on families' eligibility and availability of slots by site.

Our analysis indicates that the proposed design will have sufficient statistical power to detect impacts of the magnitude we might expect to occur for two of the central outcomes of the study—housing stability and child separation from the family for some pairwise comparisons. As discussed below, we will be able to detect effects on these outcomes as small as 8.0 percentage points for the CBRR vs. UC and SUB vs. UC pairwise comparisons and as small as 10.2 percentage points for the PBTH vs. UC comparison..

Exhibit B-3 shows the MDEs by pairwise comparison for the pooled study sample of 1,729 which is 75 percent response of the full sample of 2,305 families. The MDEs presented are the minimum detectable differences in outcomes (in percentage points) between two randomly assigned groups with 80 percent power when we perform a two-sided[3] statistical test at 10 percent level of significance, assuming a regression R^2 of 0.10[4] and no finite population correction.[5] The differences are shown for various average outcome levels for second assignment group in each comparison.

The last column of the first row of Exhibit B-3 shows that for a mean group outcome of 0.5, the MDE for the CBRR vs. UC comparison is 8.0 percentage points. This means that if the true effect of CBRR compared to UC is to change the prevalence rate of an outcome measure—such as return to shelter housing, or percent of families whose head is a leaseholder at 18-month follow-up—from 50 percent to under 42 percent (for return to shelter) or above 58 percent (for lease holding), we would have an 80 percent likelihood of obtaining an impact estimate that is statistically significant. If the true effect is less than 8 percentage points, there is a lower likelihood that differences between these assignment groups will be statistically significant, though many might still be detected.

---

3 While one-sided tests would decrease MDE's, we believe one-sided tests are inappropriate because we care about negative impacts; i.e., they are in a substantive sense not equivalent to a finding of no impact. To see this consider comparing Transitional Housing to Subsidy Only. There a negative point estimate implies that one of the interventions is worse than the other. We care about that, above and beyond the idea that the other intervention is not better.

4 Since we will estimate regression-adjusted impact estimates, we assume an amount of explanatory power for the regressions. An R^2 of 0.10 is conservatively assumed. This is the pseudo-R^2 for the general health outcome probit regression in the Effects of Housing Vouchers on Welfare Families evaluation. Outcomes with higher regression R^2's will have smaller MDE's.

5 Applying the finite population correction (FPC) would reduce the MDE's. However, we believe not applying the FPC more accurately represents our uncertainty as to results holding true in future similar applications of the intervention approaches.

Our hypothesis is that the interventions to be tested in relation to the Usual Care intervention—all involving housing assistance or subsidy of some sort—will have fairly large effects on housing stability. Drawing on the longitudinal HMIS analysis of shelter utilization (AHAR, 2008; Culhane et al., 2007), we estimate that of families who remain in shelter for at least seven days without any special assistance, approximately 50-60 percent find housing that keeps them from returning within a multi-year follow-up period. There is substantial potential for the proposed interventions to expand this percentage, by using subsidies to eliminate the risk of shelter return for many families in the other 40-50 percent of the population. Housing subsidies remain available to families many months after first receipt, during which time they should provide a sufficiently stable and improved housing option compared to shelters that, for most families, precludes the need for returns to shelter. Research in St. Louis, Philadelphia, and New York City (Stretch & Krueger, 1993; Culhane 1992; Shinn et al., 1998) tends to support this projection. For example, in St. Louis just 6 percent of families who left shelter with a housing voucher returned, compared to 33 percent of those without subsidized housing.[6] Housing stability differed by more than 60 percent between those who received a subsidy (80 percent in stable housing at five years) and those who did not (18 percent stable at five years) in the New York study. Thus, we conclude that an MDE of 8.0 to 10.2 percentage points assures confident detection of the type of impact on housing stability we would expect from the tested interventions (CBRR, SUB, and PBTH) when compared to the Usual Care group.

Exhibit B-3. Minimum Detectable Effects for Prevalence Estimates by Pairwise Comparison

Sample	Expected Number of Completed Follow-up Survey Interviews		MDE if Mean Outcome for the Second Assignment Group is:		
	First Assignment Group	Second Assignment Group	0.1 (or 0.9)	0.3 (or 0.7)	0.5
CBRR vs. UC	433	435	4.8 pp	7.3 pp	8.0 pp
SUB vs. UC	453	411	4.8 pp	7.3 pp	8.0 pp
PBTH vs. UC	272	257	6.1 pp	9.4 pp	10.2 pp
CBRR vs. SUB	290	329	5.7 pp	8.7 pp	9.5 pp
CBRR vs. PBTH	177	175	7.5 pp	11.5 pp	12.6 pp
SUB vs. PBTH	180	194	7.3 pp	11.2 pp	12.2 pp

Notes: (1) The MDE's are based on calculations which assume that two-sided tests are used at the 10 percent significance level, the desired power is 80 percent, and the regression R^2 is 0.10. (2) All MDE's assume a 75% survey response rate, with no finite population correction.

A similar conclusion holds for the prevalence of child separation from the family during the follow-up period. This is likely to be a less common occurrence, making the column of Exhibit B-3 labeled

6 Note that this observational pattern is not a direct measure of the impact of subsidized housing on shelter return. Likely the families who exited shelter with subsidies differed from the without-subsidy group on other factors that led to their better outcomes. But even if the difference in unadjusted shelter return rates exaggerates the true impact of a subsidy by an extreme amount—say, 2 or 3 times—the observed 27 percentage point difference would mean an impact of 9 to 13 percentage points.

"MDE if Mean Control Group Outcome is: 0.3" likely the most relevant one.[7] Here, a slightly smaller true impact can be detected with 80 percent assurance. The MDEs in Exhibit B-3 are for analyses that are performed with the entire pooled sample. MDEs for split-sample subgroups are larger than those shown here. As noted elsewhere, the study will be best equipped to explore how impacts differ by family characteristics using the Family Need Index whose role in producing larger or smaller impacts can be examined without dividing the sample into pieces.[8]

Power Calculations for Earnings

Exhibit B-4 shows the MDEs for earnings impacts by pairwise comparison. These MDE's are based on the adult earnings outcomes from the Moving To Opportunity (MTO) Demonstration (Orr, et al., 2003), a study of families who were living in distressed (i.e., barely better than emergency shelters) public housing or private assisted housing projects in high poverty neighborhoods at baseline. The first row of the exhibit shows that the analysis will be able to detect a difference between mean annual earnings of the CBRR and UC groups of $1,170 with 80% likelihood. Given that only two of the interventions tested have a partial focus on the labor market—though better, more stable housing may enable steadier employment and resulting greater earnings—the study design is weaker for detecting these effects. On the one hand, it is by no means assured that even an intervention directly focused on employment and training could produce an earnings impact of over $1,200 per year. On the other hand, a true impact substantially smaller than this amount—say, an impact on annual earnings of $600—would have little potential to move families out of poverty and hence may not be important to detect with high confidence.

7 We note that Cowal, et al, (2002), finds 44 percent. In as much as that estimate applies here, we will have slightly lower power.

8 "Challenge score" can be entered into the impact regression equation interacted with indicator variables for the different random assignment groups to see if the magnitude of effect from being assigned to a particular service package changes as the degree of family challenge rises, and the equation then estimated using all the data. Impacts on categorical subgroups will be estimated by splitting the sample and doing separate analyses for each category.

Exhibit B-4. Minimum Detectable Effects for Annual Earnings Impacts by Pairwise Comparison

Sample	Expected Number of Completed Follow-up Survey Interviews		MDE (dollars)
	First Assignment Group	Second Assignment Group	
CBRR vs. UC	433	435	1,170
SUB vs. UC	453	411	1,172
PBTH vs. UC	272	257	1,498
CBRR vs. SUB	290	329	1,385
CBRR vs. PBTH	177	175	1,837
SUB vs. PBTH	180	194	1,783

Notes: (1) The MDE's are based on calculations which assume that two-sided tests are used at the 10 percent significance level, the desired power is 80 percent, and the regression R^2 is identical to the MTO adult annual earnings impact regression. (2) All MDE's assume a 75% survey response rate, with no finite population correction. (3) The variance of earnings is derived from the standard error of the ITT impact estimate for the experimental group (n=1,729) vs. the treatment group (n=1,310) in the MTO Demonstration: $254.

B.2.4 Procedures with Special Populations

In this study we may encounter interview respondents whose first language is Spanish. As we did with the baseline survey, we will translate the follow-up survey instrument into Spanish, for administration in the language most comfortable for the respondent. The participation agreement also will be made available in Spanish.

All baseline interviews were conducted in either English or Spanish, with no need for other languages.

B.3 Maximizing the Response Rate

During the data collection period for the participant tracking component of the study, non-response levels and the risk of non-response bias will be minimized in the following ways:

- The Contractor will rely on the local site interviewers to lead the continued tracking and follow-up survey efforts. They are already established in the local communities and have an existing rapport with the study families.
- The Contractor will support the local site interviewers through recruitment of additional interviewers skilled at working with this population.
- Respondents will have a choice of time for the data collection.
- Additional field tracking and locating steps will be taken, as needed, when sample members are not found at the phone numbers or addresses previously collected.
- The use of the Abt Associates Field Management System will permit interactive sample management and electronic searches of historical tracking and locating data.

- For the follow-up survey, the Contractor's survey director and field supervisors will manage the sample release and monitor response rates in a manner that allows us to work the sample groups for each of the study interventions evenly.

By these methods, the Contractor anticipates being able to achieve a 75 percent response rate for the follow-up survey. The contractor team has extensive experience conducting longitudinal studies involving participant tracking and follow-up survey data collection. The tracking approach in place for the Family Options Study reflects that experience and track record for achieving high response rates to follow-up survey data collection efforts. The proactive and rigorous approach to tracking respondents is designed to maintain contact with the study sample over the follow-up period which we believe is essential to achieving the highest possible response rate for the 18-month follow-up survey.

For the Family Options study, we use a combination of passive tracking methods such as searches of proprietary database and National Change of Address as well active contacts with the respondents every three months after random assignment. Active contacts consist of a brief telephone call three months after random assignment, a short survey six and twelve months after random assignment, and tracking letters nine and fifteen months after random assignment. In addition to these frequent contacts, the active tracking protocol benefits from having field interviewers located in the community conduct the tracking. Because the field interviewers are local, they can contact respondents in person when telephone response is not obtained.

The intent of the active contacts is two-fold:

- □ Collect information directly from the respondent about his/her most up to date address and best way to contact her/him along with dependable secondary contacts;
- □ If we are unable to contact the respondent directly during one of the active contacts, the attempted contact provides information about the quality of the contact information available for each respondent and as well as information about the next steps to take on the harder to locate cases.

Our current experience with participant tracking has been encouraging. Currently, tracking completion rates for cohorts in which active tracking efforts have been closed are:

- 70 percent for the 3-month tracking contact;
- 64 percent for the 6-month tracking contact; and
- 65 percent for the 12-month tracking contact.

We have also have achieved a 26 percent completion rate to the tracking letter sent to participants nine months after random assignment, which exceeds typical response of 20 percent from our previous experience with this type of contact. In addition, we find it encouraging that in some cases, respondents who we have not reached for previous tracking efforts have returned the tracking letter sent nine months after random assignment.

It is important to note that the response rates for tracking activities, for this and other studies, are not meant to serve as an indicator for response rates that can be achieved in future follow-up survey data collection activities. This is because tracking efforts do not follow the full data collection protocol that is implemented for follow-up survey data collection. Follow-up survey data collection includes multiple phone calls to a respondent, multiple in-person attempts and attempts to contact all of the secondary contacts provided in the baseline interview, in addition to a range of passive locating methods. Follow-up survey data collection involves all of these things, thereby achieving higher response rates than earlier, less extensive attempts at contacting families.

Our current completion rates for tracking efforts are quite encouraging but it is important to keep in mind that even when we are unable to contact the respondent directly at any of the scheduled tracking contacts, we are still obtaining valuable locating information about each sample member. This information obtained during tracking—regardless of whether or not we reach the respondent—improves the likelihood of reaching them for the 18-month follow-up survey. For example, in some cases secondary contacts have informed us that respondents are temporarily in treatment centers or prison and unreachable to us For these types of cases, we collect as much information as we can (e.g.: release date, name of facility). We capture all of this information in the tracking database and will call on it for future tracking and for the follow-up survey data collection effort. Another encouraging indicator is that only 15 percent of the tracking letters sent nine months after random assignment have been returned as undeliverable. All the information gathered during the tracking efforts are captured in the study's tracking database. This information will be available to the field staff at all times.

Prior to the start of data collection for the 18 month survey, the data collection team will review the study sample to identify cases that do not have tracking updates from any point of tracking. These cases will be classified as "high priority cases" and will be assigned to a Senior Field Interviewer. In addition to any information obtained through the tracking process (e.g. returned letters, information from secondary contacts), the research team will also contact the program providers in the study sites to which the respondents were referred for assistance, to request any information available about the family's location, to ensure that all possible sources of locating information are available to the team. High priority cases will be reviewed by the Field Manager regularly to make sure all leads are followed.

Field interviewers who conduct the 18-month follow up survey will receive a comprehensive document for all released cases containing respondent's information history collected through the tracking components (all address, home/cell phone numbers and emails), secondary contacts and any relevant notes collected during the tracking efforts. The information will include the date of the update as well as the source of the update to help staff prioritize the locating data to determine which information to use first. After several weeks of data collection and if we still have not been able to locate the respondent, we will contact the providers to follow the same procedure as for the high priority cases.

The ongoing completion rates on the tracking components are an important indicator of sample productivity. To put these results in context, in previous longitudinal studies that have achieved a response rate of 75 percent or more for a follow-up survey, interim tracking efforts often yield only 30 to 40 percent response rates. Based on our experience on projects with similar populations and our response to ongoing tracking activities we feel confident that our approach will achieve the 75 percent response rate in the 18 month follow up survey.

B.4 Test of Procedures

Prior to commencing the follow-up survey data collection, HUD's evaluation contractor Abt Associates will conduct a pretest of the questionnaire with no more than nine respondents for any given survey item. Pretest respondents will be selected from members of the ***Family Options Study***. The pretest will allow the contractor to test the appropriateness of language level and word usage in the questionnaire. The pretest also will allow the contractor to confirm the estimates of interview length. Experienced interviewers will conduct the pretest, and senior survey staff will supervise this activity. Abt Associates will prepare a pretest report that describes the problems encountered and recommends solutions, as necessary, to shorten the survey instruments to conform with the planned length, to simplify the language to ensure that respondents understand the questions, and to modify question order or skip patterns to make sure that items flow smoothly and logically for respondents. Abt Associates will coordinate closely with HUD to coordinate the schedule for the pre-test, following HUD approval of the survey instrument.

B.5 Individuals Consulted on Statistical Aspects of the Design

The individuals shown in Exhibit B-5 assisted HUD in the statistical design of the evaluation.

Exhibit B-5. Individuals Consulted on the Study Design

Name	Telephone Number	Role in Study
Dr. Stephen Bell Abt Associates Inc.	301-634-1721	Co-Principal Investigator
Dr. Marybeth Shinn Vanderbilt University	615-322-8735	Co-Principal Investigator
Dr. Jill Khadduri Abt Associates Inc.	301-634-1745	Project Quality Advisor
Mr. Jacob Klerman Abt Associates Inc.	617-520-2613	Project Quality Advisor
Dr. Martha Burt Consultant to Abt Associates Inc.	202-261-5551	Project Advisor
Dr. Dennis Culhane University of Pennsylvania	215-746-3245	Project Advisor
Dr. Ellen Bassuk, Center for Social Innovation and National Center on Family Homelessness	617-467-6014	Project Advisor
Dr. Beth Weitzman New York University	212-998-7446	Project Advisor
Dr. Larry Orr Consultant to Abt Associates Inc.	301-467-1234	Project Advisor

Inquiries regarding the statistical aspects of the study's planned analysis should be directed to:

Dr. Stephen Bell	Co-Principal Investigator	Telephone: 301-634-1721
Dr. Marybeth Shinn	Co-Principal Investigator	Telephone: 615-322-8735

References

Burt, M.R. (2006). Characteristics of transitional housing for homeless families: Final report. Prepared for the US Department of Housing and Urban Development.

Cowal, K., Shinn, M., Weitzman, B.C., Stojanovic, D., & Labay, L. (2002). Mother-child separations among homeless and housed families receiving public assistance in New York City. *American Journal of Community Psychology, 30*, 711-730.

Culhane, D. P., Metraux, S., Park, J.M., Schretzman, M. & Valente, J. (2007). Testing a typology of family homelessness based on patterns of public shelter utilization in four U.S. jurisdictions: Implications for policy and program planning. *Housing Policy Debate, 18*(1), 1-28.

Culhane, D. P. (1992). The quandaries of shelter reform: An appraisal of efforts to “manage” homelessness. *Social Service Review, 66,* 428–440.

Duffer, Allen P. et al., "Effects of Incentive Payments on Response Rates and Field Costs in a Pretest of a National CAPI Survey" (Research Triangle Institute, May 1994).

Locke, G., Khadduri, J. & O’Hara, A. (2007). Housing models (Draft).

Rog, D.J. & Randolph, F.L. (2002). A multisite evaluation of supported housing: Lessons learned from cross-site collaboration. *New Directions for Evaluation, 94*, 61-72.

"National Adult Literacy Survey Addendum to Clearance Package, Volume II: Analyses of the NALS Field Test" (Educational Testing Service, September 1991), pp. 2-3.

Orr, L.L., Feins, J., Jacob, R., Beecroft, E., Sanbonmatsu, L., Katz, L., Liebman, J. & Kling, J. (2003). *Moving to Opportunity interim impacts evaluation: Final report*. Cambridge, MA: Abt Associates Inc. and National Bureau of Economic Research.

Shinn, M., Weitzman, B. C., Stojanovic, D., Knickman, J. R., Jiminez, L., Duchon, L., James, S., & Krantz, D.H. (1998). Predictors of homelessness from shelter request to housing stability among families in New York City. *American Journal of Public Health, 88* (10), 1651-1657.

Stretch, J. J. & Kreuger, L. W. (1993). A social-epidemiological five year cohort study of homeless families: A public/private venture policy analysis using applied computer technology. *Computers in Human Services, 9*(3-4), 209-230.

U.S. Department of Housing and Urban Development (2011). The Fifth Annual Homeless Assessment Report (AHAR) to Congress.