

# **Building Evidence on Employment Strategies for Low-Income Families**

**OMB Information Collection Request  
New Collection**

## **Supporting Statement**

### **Part B**

**September 2019**

Submitted By:

Office of Planning, Research, and Evaluation  
Administration for Children and Families  
U.S. Department of Health and Human Services

4<sup>th</sup> Floor, Mary E. Switzer Building  
330 C Street, SW  
Washington, D.C. 20201

Project Officer:

Tiffany McCormack

## **B1. Respondent Universe and Sampling Methods**

The Administration for Children and Families (ACF) at the U.S. Department of Health and Human Services (HHS) seeks approval to recruit evaluation sites and collect data from sites and study participants for the Building Evidence on Employment Strategies for Low-Income Families (BEES) study.

### ***Sampling and Target Population***

All BEES sites will focus on providing employment services to low-income populations, particularly individuals overcoming substance use disorder, mental health issues, and other complex barriers to employment. BEES will include up to 21 programs, approximately 14 of which will include an impact and implementation study, while the remaining programs – where an impact study is not feasible – will involve implementation-only studies.

### **Impact and Implementation Studies**

Among the programs with an impact study, we anticipate sample sizes of 1000 in each of the approximately 6 non-behavioral health programs and 800 in each of the approximately 8 behavioral health programs, which includes substance abuse and mental health programs.

Within each program participating in the impact and implementation analyses, participants who are eligible for the offered services will be enrolled in the BEES study, if willing to provide consent. Not participating in the study will not affect access to program services. Program staff will identify eligible applicants using their usual procedures. Then, a staff member will explain the study and obtain informed consent. We anticipate enrolling all eligible participants who provide consent over the enrollment period, with an estimated sample size of 12,400 across all BEES sites.

### **Implementation-Only Studies**

Although BEES is prioritizing mature programs ready for rigorous evaluation, such as randomized control trials, the research team has identified programs where an impact study is not possible at this time. New programs tend to be small and are still developing their program approach and services, making them unsuitable for rigorous evaluations. Designed to provide information to the field in a timely and accessible manner, implementation studies will complement the rigorous evaluations of larger, and perhaps more established, programs. Programs identified as unsuitable for random assignment at this time may be proposed for implementation-only studies. Within the set of SUD programs, we will select sites that represent a variety of approaches to designing and operating programs that provide employment services to individuals struggling with or in recovery from substance use disorder, particularly opioid use disorder. Programs using a whole family approach will be selected to represent different approaches to multi-generational work and in different locations. For these studies, site visitors will use a semi-structured interview protocol during two-day visits to each site; the protocol was developed for the BEES implementation study to conduct interviews with program managers, staff, and partners and tailored to these types of programs being visited. If it is not possible to do an in-person visit, interviews will be done virtually via phone calls or video conference.

### ***Research Design and Statistical Power***

#### **Impact Analysis**

This section briefly describes some principles underlying the impact analysis for each program in BEES.

- **Preference for random assignment.** Randomized control trials (RCTs) are the preferred method for estimating impacts in the BEES evaluations. However, RCTs might not be feasible, for example, if there are not enough eligible individuals to provide a control group or if a program serves such a high-risk population that it would be problematic to assign any of them to a control group. In that case, the team will propose strong quasi-experimental designs, such as regression discontinuity designs (Imbens and Lemieux, 2008), comparative interrupted time series (Somers, Zhu, Jacob, and Bloom, 2013), and single case designs (Horner and Spaulding, 2010). All of these designs can provide reliable estimates of impacts under the right circumstances, and we would use procedures that meet the best practices for these designs, such as those proposed by the What Works Clearinghouse.
- **Intent-to-treat impact estimates.** The starting point for the impact analysis is to compare outcomes for all program group members and control group members. In an RCT, random assignment ensures that this comparison provides an unbiased estimate of offering the program group access to the intervention. To increase precision, impact estimates are regression adjusted, controlling for baseline characteristics.

Following the recommendations of Schochet (2008), the impact analysis would reduce the likelihood of a false positive by focusing on a short list of confirmatory outcomes that would be specified before the analysis begins. To further reduce the chance of a false positive finding, results for any confirmatory outcomes would be adjusted for having multiple outcomes, for example, by using the method of Westfall and Young (1993).

Exploratory analyses might also be proposed, depending on what is found in the primary analysis. For example, if the primary impact analysis finds that an intervention increases cumulative earnings, secondary analyses could investigate whether earnings gains were sustained at the end of the follow-up period.

- **Sample size and statistical power.** The ability of the study to find statistically significant impacts will depend in part on the number of families for which data are collected. Exhibit 1 presents some minimum detectable effect (MDE) estimates for several different scenarios (all assuming that 50 percent of study participants are assigned to the program group and 50 percent are assigned to the control group). An MDE is the smallest true effect that would be found in 80 percent of studies. Since our assumed sample sizes and data sources vary by type of intervention, results are shown separately for behavioral health and non-behavioral health interventions. Results are expressed both in effect sizes – that is, as a number of standard deviations of the outcome – and for illustrative outcomes. The sample sizes are as follows: (1) administrative data matched to 90 percent of the full sample (assumed to be 1000 in the non-behavioral health sites and 800 in the behavioral health sites) (2) a 6-month survey with 600 respondents in non-behavioral health sites, (3) a 12- to 18-month survey with 640 respondents in behavioral health sites. (Note that not all sites will be included in these follow-up surveys.)

Key points from Exhibit 1 include:

- With administrative data, the study would be able to detect impacts of 0.166 standard deviations in non-behavioral health sites and .185 in the other sites. This translates into impacts on employment of 7.6 and 8.5 percentage points, respectively, assuming 70 percent of the control group works.

- The 6-month survey would be able to detect impacts of 0.203 standard deviations, which translates into an impact on program participation of 6.1 percentage points, assuming 10 percent of the control group participates in services similar to those offered by the program.
- The 12- to 18-month survey would be able to detect impacts of 0.197 standard deviations, which would translate into a reduction in having moderate or worse depression of 9.6 percentage points (using results from the Rhode Island depression study).

**Exhibit 1: Minimum Detectable Effects by Data Source**

	<b>Control Group Level</b>	<b>Non-behavioral Health</b>	<b>Behavioral Health</b>
Administrative records		0.166	0.185
Employed (%)	70	7.6	8.5
Short-term survey		0.203	0.321
Participated in services (%)	10	6.1	9.6
12-month survey			0.197
Moderate or worse depression (%)	60		9.6

NOTES: Results are the smallest true impact that would generate statistically significant impact estimates in 80 percent of studies with a similar design using two-tailed t-tests with a 10 percent significance level. No adjustment for multiple comparisons is assumed.

- **Estimated effects for participants.** If a substantial portion of the program group receives no program services or a substantial portion of the control group receives similar services offered to those in the program group, the study may provide estimates of the effect of the intervention among those for whom being assigned to the program changed their receipt of program services. These so-called local average treatment effects (Angrist and Imbens, 1995) can be as simple as dividing intent-to-treat impacts by the difference in program participation between the program and control groups. Such estimates could be made more precise if there are substantial differences in use of program services across subgroups of participants or sites offering similar services (Reardon and Raudenbush, 2013).
- **Subgroup estimates.** The analysis would also investigate whether the interventions have larger effects for some groups of participants (Bloom and Michalopoulos, 2011). In the main subgroup analysis, subgroups will be chosen using baseline characteristics, based on each evaluation’s target population and any aspects of the theory of change that suggest impacts might be stronger for some groups. This type of subgroup analysis is standard in MDRC studies and has been used in analyzing welfare-to-work programs, including for subgroups at risk for mental health issues, such as depression (e.g., Michalopoulos, 2004; Michalopoulos and Schwartz, 2000).

## **B2. Procedures for Collection of Information**

### ***Impact Study Data Sources***

#### Data Collected at Study Enrollment.

BEES study enrollment will build upon each participating program's existing enrollment and data collection processes. The following describes the procedures for data collection at study enrollment and how study enrollment will be combined with program operations.

Before recruiting participants into the study, following the site's existing procedures, the program will collect information using their normal processes to determine whether the participant is eligible for the program's services. Information will be logged into their existing management information system (MIS) or data locating system.

For study enrollment, the program staff will then conduct the following procedures. Note that these procedures assume a randomized control trial design, which is the preference for BEES. If another design is used, step 4 (random assignment) would not be relevant, and the program could begin providing services after the individual has completed step 3 below.

- Step 1. Introduce the study to the participant. Provide a commitment to privacy, explain random assignment, and answer questions just prior to going through the consent form to ensure an understanding of the implications of the study.
- Step 2. Attempt to obtain informed consent to participate in BEES using the informed consent form (Attachment H). Staff members will be trained to explain the issues in the consent/assent form and to be able to answer questions. If the applicant would like to participate, they will sign the consent form – either electronically or on a paper form. The participant will also be given a copy of the consent form for their records.
- Step 3. A staff person will provide the Baseline Information Form (Attachment D) for the participant to complete on paper or electronically.
- Step 4. Indicate in the random assignment system that the participant is ready to be randomly assigned, if random assignment is occurring on an individual basis. The result of random assignment will be immediate. Random assignment may also occur on a cohort-basis, depending on the site's usual enrollment processes.
- Step 5. Inform the participant of assignment to receive the services available to the program group or the alternative services that will be available to the control group.

#### Data Collected After Study Enrollment

This section describes procedures related to data collection after study enrollment.

**Interviewer Staffing.** An experienced, trained staff of interviewers will conduct the 6- and 12- to 18-month participant surveys. The training includes didactic presentations, numerous hands-on practice exercises, and role-play interviews. The evaluator's training materials will place special emphasis on project knowledge and sample complexity, gaining cooperation, refusal aversion, and refusal conversion. The study team maintains a roster of approximately 1,700 experienced interviewers across the country. To the extent possible, the BEES study will recruit interviewers who worked successfully on prior career pathways studies for ACF (such as first round Health Profession Opportunity Grants (HPOG) Impact and Pathways for Advancing Careers and Education (Pace) 15-month, 36-month, and 72-month surveys under

OMB control numbers 0970-0394 and 0979-0397 respectively). These interviewers are familiar with employment programs, and they have valuable experience locating difficult-to-reach respondents. We will also recruit other interviewers with expert locating skills to ensure that the needs of all studies are met successfully.

All potential interviewers will be carefully screened for their overall suitability and ability to work within the project's schedule, as well as the specific skill sets and experience needed for this study (e.g., previous data collection experience, strong test scores for accurately recording information, attention to detail, reliability, and self-discipline).

**Participant Contact Update Request and Survey Reminders.** A variety of tools will be utilized to maintain current contact information for participants and remind them of the opportunity for study participation, in the interest of maximizing response rates.

#### *Impact Study Instruments*

- **Contact Update Request Letter and Form (Attachment E).** All study participants will be included in the effort to maintain updated contact information.<sup>1</sup> The participant contact update form will be self-administered. The letter and accompany form (Attachment E) will be mailed to sample members, beginning three months after random assignment. Participants will be encouraged to respond by returning the form by mail, through a secure online portal, or they can update their contact information by phone. Participants can indicate that the information is correct, or they can make any necessary changes in contact information.

**Locating Materials.** The following materials will be used throughout the study to maintain contact with participants, remind them of their participation in the study, and ensure the correct contact information for successful survey completion.

The survey-related materials will only be used among participants who are assigned to the respective survey: 6-month survey for participants at sites serving individuals *without* mental health or substance use challenges, and 12- to 18-month survey for participants at sites serving populations *with* those challenges. As noted previously, these surveys will not occur in all sites.

#### *Proposed Phase 1 Supplementary Materials:*

- **Welcome Letter (Attachment I).** The welcome letter will be mailed to all participants soon after enrollment in the study. Its intention is to remind the participant of their agreement to participate, what it means to participate, and what future contacts to expect from the research team.

#### *Proposed Phase II Supplementary Materials:*

- **6- and 12- to 18-Month Survey Advance Letters (Attachment P).** To further support the data collection effort, an advance letter will be mailed to study participants selected to participate in the interview approximately one and a half weeks before interviewers begin the data collection. The advance letter serves as a way to re-engage the participant in the study and alert them to the upcoming effort so that they are more prepared for the interviewer's call. The evaluators will update the sample database prior to mailing advance letters to help ensure that the letter is mailed to the most up to date address. The advance letter will remind each participant about the timing of the upcoming data collection effort, the voluntary nature of participation, and that researchers will keep all answers private. The letter provides each selected study participant with a toll-free

---

<sup>1</sup> Note that the burden estimate assumes that a subsample of study participants will respond among the sites where these follow-up interviews are planned. In addition, we anticipate that on average participants will respond an average of one time to any of these requests.

number that he/she can call to set-up an interview. Interviewers may send an email version of the advance letter to participants for whom we have email addresses.

- **6- and 12- to 18-Month Survey Email Reminders (Attachment Q).** Interviewers will attempt to contact participants by telephone first. If initial phone attempts are unsuccessful, interviewers can use their project-specific email accounts to follow up and attempt to set up a survey. They send this email, along with the advance letter, about halfway through the time during which they are working the cases.
- **6- and 12- to 18-Month Survey Flyers (Attachment R).** This flyer, to be sent by mail, is another attempt to contact participants who have not responded to 6- or 12- to 18-month survey attempts. Interviewers may leave specially designed project flyers with family members or friends.

### ***Implementation Study Data Sources***

An implementation study will be conducted in all the sites in BEES using a random assignment research design. The instruments for the implementation study include the Program, Managers, and Staff Interview Guide (Attachment L), the In-depth Case Study of Participant-Staff Perspectives Interview Guides (Attachments M and N), and the Program Staff Survey (Attachment O). Staff interviews will be conducted during two rounds of implementation study visits for each impact study site while, in order to tailor instruments based on data collected in the first round, the in-depth case studies and program staff survey will be conducted during the second round of site visits. If it is not possible to do an in-person site visit, interviews will be done virtually via phone calls or video conference.

Programs recommended for implementation-only studies under Phase I were identified through phone interviews conducted over the last nine months as part of the scan to identify BEES sites conducted under OMB #0970-0356. These sites were determined to be inappropriate for more rigorous study due to a range of factors, including limited scale or lack of excess demand to create a control group. Based on this work, we identified nine potential programs for implementation-only studies. Upon OMB approval of Phase I, we will contact each program to ensure it is still a good fit for BEES, assess their interest in participating in BEES, and develop plans to conduct a site visit if they want to move forward. If a site does not want to participate, we have identified alternate sites that could be candidates for inclusion.

Data collection procedures for the implementation studies are outlined below.

#### ***Proposed Phase I Instruments:***

- **Program Managers, Staff, and Partners Interview Guide – Substance Use Disorder Treatment Programs (Attachment F).** Staff interviews will be conducted during a site visit (done either in-person, or if not possible, virtually) to each of the 6 separate sites selected for the implementation study of SUD Treatment/Employment programs. During each visit, 90-minute semi-structured interviews of program staff and partners will explore implementation dimensions, including program context and environment; program goals and structure; partnerships; recruitment, target populations, and program eligibility; program services; and lessons learned. The number of interviewees will vary by site depending on the organization staffing structure, however, the goal is to include all program managers, program staff (i.e. case managers), and key partners. If the program is large and interviews with all staff cannot be completed in the time available, we will interview at least one (and likely more) in each job category. We estimate approximately 10 staff total will be interviewed at each. Topics for the interviews include: the choice of target groups; participant outreach strategies; employment, training, and support service

provided; SUD treatment and recovery services, development or refinement of existing employment-related activities and curricula to serve the target group; how and why partnerships were established; strategies for engaging employers with a SUD population; and promising practices and challenges. All protocols begin with a brief introductory script that summarizes the overall evaluation, the focus of each interview, how respondent privacy will be protected, and how data will be aggregated.

- **Program Managers, Staff, and Partners Interview Guide – Whole Family Approach Programs (Attachment G).** Staff interviews will be conducted over two rounds of visits to two locations within these sites (done either in-person, or if not possible, virtually). The number of interviewees will vary by location depending on the organization staffing structure, however, the goal is to include all program managers, program staff (i.e. case managers), and key partners. If the program is large and interviews with all staff cannot be completed in the time available, we will interview at least one (and likely more) in each job category. We estimate approximately 10 staff total will be interviewed at each program. Topics for the interviews include: program model and structure, program start up, staffing, program implementation, program components strategies and staff experiences, participant knowledge, awareness, participation, and views of program, interactions with partners, and eligibility criteria for participants.

*Proposed Phase II Instruments:*

- **Program Managers, Staff, and Partners Interview Guide (Attachment L).** Site visits (done either in-person, or if not possible, virtually) will be led by senior evaluation team members with expertise in employment program and in implementation research. All protocols (Attachment I) begin with a brief introductory script that summarizes the overall evaluation, the focus of each interview, how respondent privacy will be protected, and how data will be aggregated. During each visit, 90-minute semi-structured interviews of program staff and partners will explore staff roles and responsibilities, program services, and implementation experiences. The number of interviewees will vary by site depending on the organization staffing structure, however, the goal is to include all program managers, program staff (i.e. case managers), and key partners. If the program is large and interviews with all staff cannot be completed in the time available, we will interview at least one (and likely more) in each job category. We estimate approximately 10 staff total will be interviewed at each program. Topics for the interviews include: program model and structure, staffing, program implementation, program components strategies and staff experiences, participant knowledge, awareness, participation, and views of program, use of services and incentives, and counterfactual environment.
- **In-depth Case Study of Participant-Staff Perspective Program Staff Interview Guides (Attachments M and N).** In-depth case studies will be conducted at up to 21 evaluation sites examining selected participants and their corresponding case manager to understand how program staff addressed a specific case, how the participant viewed the specific services and assistance received, and the extent to which program services addressed participant needs and circumstances. For each site, one-on-one interviews will be conducted with six participants (each 90 minutes in length), with separate one-on-one interviews with their respective case managers (60 minutes in length). We will work with case managers who express an interest in this study and select participants and to represent a range of situations, such as different barriers to employment and different completion and drop-out experiences, and different background characteristics. If a participant declines to participate, we will select another participant with similar experiences, to ensure a sample of six. Staff interviews differ from those described above because they will focus on how the staff member handled a specific case, in contrast to how the program works overall. The case studies will provide examples of how the program worked for

specific cases, and will enhance the overall understanding of program operations, successes, and challenges.

- **Program Staff Survey (Attachment O).** Online staff surveys will be fielded to all line staff (will not include supervisors or managers) at each site (we estimate 20 per site) and will cover background and demographics, staff responsibilities, types of services provided by the organization, barriers to employment, program participation, and organizational and program performance. Each survey will take 30 minutes to complete. The online staff survey (Attachment K) will be accessed via a live secure web-link. This approach is particularly well-suited to the needs of these surveys in that respondents can easily start and stop if they are interrupted or cannot complete the entire survey in one sitting and review and/or modify responses in a previous section. Respondents will be emailed a link to the web-based survey.

### **B3. Methods to Maximize Response Rates and Deal with Nonresponse**

#### *Expected Response Rates*

Study participants are required to complete the consent form (Attachment H) and baseline information form (Attachment D) as part of study enrollment. As such, we expect nearly 100 percent participation for these data collections.

For the 6-month follow-up participant surveys (Attachment J), we expect a minimum response rate of 60 percent since it is a phone-only effort. <sup>2</sup>A similar follow-up interview effort focused on service receipt for the WorkAdvance evaluation reached a 69 percent response rate, on average (Hendra et al., 2016) and the STED evaluation's phone efforts for the in-program survey averaged just over 60 percent across the 6 sites (Williams and Hendra, 2018). One of the primary goals of this data collection is to understand and detect impacts of BEES sites on service receipt. Because we expect there to be a large difference in service receipt between the treatment and control groups, a 60 percent response rate will be adequate to detect policy relevant impacts. Within an RCT framework, a primary data quality concern is the internal validity of impact estimates. If the response rate for any major analytic subgroup is lower than 80%, findings will be appropriately caveated to aid readers' interpretation. The project's plan for addressing non-response bias, both while in the field and after the completion of data collection, is detailed in the following section. In addition, while this data collection will provide an early look at service receipt and treatment contrast in a subset of sites, there will be other data sources, such as administrative records, that will ultimately answer key impact analysis questions.

For the 12- to 18-month follow-up participant surveys (Attachment K), we expect a response rate of 80 percent. This response rate provides the sample size used in the MDE calculations (shown in Exhibit 1), which will allow us to detect differences across research groups on key outcomes. Numerous MDRC studies with similar populations have achieved response rates of at least 80 percent. For example, the Work Advancement and Support Center demonstration achieved an 81 percent response rate for the 12-month follow-up interview for a sample which included ex-offenders (Miller et al., 2012). The Parents' Fair Share study, which included non-custodial parents, achieved a response rate of 78 percent (Miller

---

<sup>2</sup> The survey effort will be monitored closely, especially with respect to the differential in response rates between the research groups. The recommendations of the What Works Clearinghouse and AAPOR (noted below) provide clear guidance regarding acceptable rates for differential attrition with respect to overall attrition of response rates.

and Knox, 2001). The Philadelphia Hard-to-Employ study (a transitional jobs program for TANF recipients) achieved a 79 percent response rate (Jacobs and Bloom, 2011). Several sites in the Employment Retention and Advancement evaluation achieved 80 percent response rates as well (Hendra et al., 2010).

Program staff will be asked to complete a survey (Attachment O). Based on the response rates for similarly fielded surveys in recent MDRC projects, we expect at least 80 percent of staff to complete the survey.

A subset of staff members will also be asked to participate in semi-structured interviews (Attachment L). Based on past experience, we expect response rates to be nearly 100 percent for these interviews, with the only nonrespondents being those staff members who are not available on the days the interviews are occurring. Similarly, we expect response rates to be high for the case study interviews (Attachments M and N). These are not intended to be representative, so program managers will select participants and take interest in participating into account.

### ***Dealing with Nonresponse***

All efforts will be made to obtain information on a high proportion of study participants through the methods discussed in the section below on maximizing response rates as well as elsewhere in the Supporting Statement. Further, the study team will monitor response rates for the program and control groups throughout data collection. Per the American Association for Public Opinion Research (AAPOR) guidelines, specific attention will be paid to minimizing differences in response rates across the research groups.

In addition to monitoring response rates across research groups, the study team will minimize nonresponse levels and the risk of nonresponse bias through strong sample control protocols, implemented by:

- Using trained interviewers with experience working on studies with similar populations and who are skilled in building and maintaining rapport with respondents, to minimize the number of break-offs and incidence of nonresponse bias.
- If appropriate, providing a Spanish language version of the instrument to help achieve a high response rate among study participants for whom Spanish is their first language.
- Sending email reminders to non-respondents (for whom we have an email address) informing them of the study and allowing them the opportunity to schedule an interview.
- Providing a toll-free study hotline number—which will be included in all communications to study participants—to help them ask questions about the interview, update their contact information, and indicate a preferred time to be called for the interview.
- For the mixed mode efforts, taking additional locating steps in the field, as needed, when the interviewers do not find sample members at the phone numbers previously collected.
- Reallocating cases to different interviewers or more senior interviewers to conduct soft refusal conversions.
- Requiring the interview supervisors to manage the sample to ensure that a relatively equal response rate for treatment and control groups is achieved.

- Using a hybrid approach for tracking respondents. First, our telephone interviewers look for confirmation of the correct number (e.g., a call back later, respondent not at home, or a voicemail message that confirms name of respondent). With contact confirmation, interviewers will continue with up to 10 attempts by phone. If interviewers are not able to confirm the correct telephone number, they will make fewer attempts before sending the case to locating. Interviewers will send cases incorrect or incomplete numbers to batch locating with a vendor such as Accurint/LexisNexis. Following locating, interviewers will attempt new phone attempts if a number is supplied or, will begin field locating, if appropriate. Once a respondent is located, “Sorry I Missed You” cards are left with field interviewers’ phone numbers so the cases can be completed by phone.

Through these methods, the research team anticipates being able to achieve the targeted response rate.

To assess non-response bias, several tests will be conducted:

- The proportion of program group and control group respondents will be compared throughout data collection to make sure the response rate is not significantly higher for one research group.
- A logistic regression will be conducted among respondents. The “left hand side” variable will be their assignment (program group or control group) while the explanatory variables will include a range of baseline characteristics. An omnibus test such as a log-likelihood test will be used to test the hypothesis that the set of baseline characteristics are not significantly related to whether a respondent is in the program group. Not rejecting this null hypothesis will provide evidence that program group and control group respondents are similar.
- Baseline characteristics of respondents will be compared to baseline characteristics of non-respondents. This will be done using a logistic regression where the outcome variable is whether someone is a respondent and the explanatory variables are baseline characteristics. An omnibus test such as a log-likelihood test will be used to test the hypothesis that the set of baseline characteristics are not significantly related to whether someone responded to the follow-up interview. Not rejecting this null hypothesis will provide evidence that non-respondents and respondents are similar.
- Impacts from administrative records sources – which are available for the full sample – will be compared for the full sample and for respondents to determine whether there are substantial differences between the two.

The main analysis will report results for individuals who respond to data collection efforts. If any of these tests indicate that non-response is providing biased impact estimates, the main analysis will be supplemented by a standard technique for dealing with nonresponse to determine the sensitivity of impact estimates to non-response. Two typical methods – inverse probability weighting and multiple imputation – are described below:

- **Inverse probability weighting:** A commonly used method for adjusting for survey response bias is to weight survey responses so they reflect the composition of the fielded survey sample. A method such as logistic regression results in predicted probability of response for each sample member. The outcomes for respondents are weighted inverse to their predicted probability of results so that the weighted results reflect the observed characteristics of the original study sample.
- **Multiple imputation:** With multiple imputation, information for respondents is used to develop a model that predicts outcome levels.<sup>3</sup> The model is used to generate outcome data for

---

<sup>3</sup> For more discussions on imputation model specification, see Little and Rubin (2002), and Allison (2002).

nonrespondents. This is done multiple times, with each iteration generating one dataset with imputed values that vary randomly from iteration to iteration. Multiple imputation accounts for missing data by restoring the natural variability in the missing data as well as the uncertainty caused by imputing missing data.

As discussed in Puma, Olsen, Bell, and Price, both methods reduce the bias of the estimated effects if survey response are due to observable factors. However, neither is guaranteed to reduce bias if nonresponse is associated with unobserved factors, which is why they methods would be used as a sensitivity check rather than for the main analysis.

### ***Maximizing Response Rates***

#### **Impact Study**

For the 6-month and 12- to 18-month surveys, the study team’s approach couples incentives with active locating to be conducted by letter or online as well as passive locating (access to public-use databases). This combination of strategies is a powerful tool for maintaining low attrition rates in longitudinal studies, especially for the control group who is not receiving the program services. The 3-month interval and incentive strategy is based on the study team’s experience as well as the literature on the high rates of mobility among low-income individuals and families (Phinney, 2013), and the effectiveness of pre-payments for increasing response rates (Cantor, O’Hare, and O’Connor, 2008).

Another important objective of active locating is to build a connection between the study and the respondents. Written materials remind the respondents that they are part of an important study, and that their participation is valuable because their experiences are unique. Written materials also stress the importance of participation for those who may be part of control group, if random assignment is used. At the same time, locating methods must minimize intrusiveness.

1. Active locating will start at baseline, with the collection of the participant’s address, telephone number, email address, and contact information of two people who do not live with the respondent but who will know how to reach him or her (Attachment D).
2. Following study enrollment, participants will receive a “welcome letter” which reminds the participant that they are part of a study, what participation in the study entails, and why their experiences are important. (Attachment I).
3. After that, all locating letters will include a contact information update form (Attachment E), and a promise of a \$5 gift card for all sample members who respond to the letter (either by updating contact information or letting the study team know that there have been no changes).
4. The passive methods we use will require no contact with the sample member. For example, the study team will automatically run the last known address for each respondent through the National Change of Address (NCOA) database, as well as LexisNexis.

#### **Implementation Study**

Maximizing response rates for the data collection efforts targeted towards staff members is also important. When a site enters the study, the research team will explain the importance of the data collection efforts for advancing the evidence base. In addition:

- For the Program Managers, Staff, and Partners Interview Guide (Attachments F, G, and L), it is important to plan the visits well in advance with the assistance of program management and schedule interviews at the most convenient times for staff.
- For the In-Depth Participant/Program Staff Case Study (Attachments M and N), we will work to the gain cooperation of six participants and their corresponding case manager. We will select staff who want to participate, and then participants from their caseloads. If a participant does not want to participate or does not show up, we will identify another participant to include in this study. As described above, we are seeking participants who provide examples of a range of experiences in the program.
- For web-based, Program Staff Survey (Attachment O), we maximize response rates primarily through good design and monitoring of completion reports. It is important to 1) keep the survey invitation attractive, short, and easy to read, 2) make accessing the survey clear and easy, and 3) communicate to the respondent that the completed survey is saved and thank them for completing the survey. Research staff will closely monitor data completion reports for the survey. If a site's surveys are not completed within one week of the targeted timeframe, the site liaison will follow up with the site point of contact to remind their staff that survey responses are due and send out reminder e-mails to staff.

## **B4. Tests of Procedures or Methods to be Undertaken**

Where possible, the survey instruments contain measures from published, validated tools. The study team will conduct pretests for both the 6- and 12- to 18-month surveys to test the instrument wording and timing. The study team will recruit 9 diverse individuals for these pretests. For the 6-month survey, most of the questions are focused on program participation and are similar across sites (the “basic” instrument), but references to particular services may differ across program domains and localities. The pretest will include a debriefing after the interview is completed, through which we can explore individuals’ understanding of questions, ease or difficulty of responding, and any concerns. Since the 12- to 18-month survey will be targeted to the evaluations focused on behavioral health, the basic instrument will be similar across these sites, with minor adaptations to local terminology. Through the pretests, the same question will not be asked of more than 9 people.

## **B5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data**

The following is a list of individuals involved in the design of the BEES project, the plans for data collection, and the analysis.

Peter Baird	Senior Associate, MDRC
Dan Bloom	Project Director, Vice President and Director, MDRC
Lauren Cates	Senior Associate, MDRC
Clare DiSalvo	Contract Social Science Research Analyst, OPRE/ACF
Mary Farrell	Subcontractor, MEF Associates

Mike Fishman	Subcontractor, MEF Associates
Robin Koralek	Subcontractor, Abt Associates
Caroline Mage	Research Associate, MDRC
Karin Martinson	Co-Principal Investigator, Subcontractor, Abt Associates
Tiffany McCormack	Senior Social Science Research Analyst, OPRE/ACF
Doug McDonald	Subcontractor, Abt Associates
Charles Michalopoulos	Co-Principal Investigator, Chief Economist, MDRC
Megan Millenky	Senior Associate, MDRC
Megan Reid	Social Science Research Analyst, OPRE/ACF
Sue Scrivener	Senior Associate, MDRC
Johanna Walter	Senior Associate, MDRC

## References

- Allison, P.D. (2002). *Missing Data*. Thousand Oaks, CA: Sage University Paper No. 136.
- Angrist, Joshua D., and Guido W. Imbens. 1995. *Identification and estimation of local average treatment effects*. NBER Technical Working Paper No. 118.
- Bloom, Howard S., and Charles Michalopoulos. 2011. "When is the Story in the Subgroups? Strategies for Interpreting and Reporting Intervention Effects on Subgroups" *Prevention Science*, 14, 2: 179-188.
- Bloom, Dan, and Charles Michalopoulos. 2001. *How Welfare and Work Policies Affect Employment and Income: A Synthesis of Research*. New York: Manpower Demonstration Research Corporation.
- Cantor, David, Barbara C. O'Hare, and Kathleen S. O'Connor. 2008. "The use of monetary incentives to reduce nonresponse in random digit dial telephone surveys." *Advances in telephone survey methodology*, 471-498.
- Gennetian, L. A., Morris, P. A., Bos, J. M., and Bloom, H. S. 2005. *Constructing Instrumental Variables from Experimental Data to Explore How Treatments Produce Effects*. New York: MDRC.
- Hendra, Richard, Keri-Nicole Dillman, Gayle Hamilton, Erika Lundquist, Karin Martinson, and Melissa Wavelet. 2010. *The Employment Retention and Advancement Project: How Effective Are Different Approaches Aiming to Increase Employment Retention and Advancement? Final Impacts for Twelve Models*. New York, NY: MDRC.
- Hendra, Richard, David H. Greenberg, Gayle Hamilton, Ari Oppenheim, Alexandra Pennington, Kelsey Schaberg, and Betsy L. Tessler. 2016. *Encouraging Evidence on a Sector-Focused Advancement Strategy: Two-Year Impacts from the WorkAdvance Demonstration*. New York, NY: MDRC.
- Horner, R., and Spaulding, S. 2010. "Single-case research designs" (pp. 1386–1394). In N. J. Salkind (Ed.), *Encyclopedia of Research Design*. Thousand Oaks, CA: Sage Publications.
- Imbens, Guido W., and Thomas Lemieux. 2008. "Regression discontinuity designs: A guide to practice." *Journal of econometrics* 142.2: 615-635.
- Jacobs, Erin, and Dan Bloom. 2011. *Alternative Employment Strategies for Hard-to-Employ TANF Recipients: Final Results from a Test of Transitional Jobs and Preemployment Services in Philadelphia*. New York: MDRC.

- Little, R.J.A., and D.B. Rubin (2002). *Statistical Analysis with Missing Data*. New York: John Wiley and Sons.
- Michalopoulos, Charles. 2004. *What Works Best for Whom: Effects of Welfare and Work Policies by Subgroup*. MDRC: New York.
- Michalopoulos, Charles and Christine Schwartz. 2000. *What Works Best for Whom: Impacts of 20 Welfare-to-Work Programs by Subgroup*. Washington: U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation and Administration for Children and Families, and U.S. Department of Education.
- Miller, Cynthia, and Virginia Knox. 2001. *The Challenge of Helping Low-Income Fathers Support Their Children: Final Lessons from Parents' Fair Share*. New York, NY: MDRC.
- Miller, Cynthia, Mark Van Dok, Betsy Tessler, and Alex Pennington. 2012. *Strategies to Help Low-Wage Workers Advance: Implementation and Final Impacts of the Work Advancement and Support Center (WASC) Demonstration*. New York, NY: MDRC.
- Phinney, Robin. 2013. "Exploring residential mobility among low-income families." *Social Service Review*, 87(4), 780-815.
- Puma, Michael J., Robert B. Olsen, Stephen H. Bell, and Cristofer Price. 2009. *What to Do When Data Are Missing in Group Randomized Controlled Trials*. (NCEE 2009-0049). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Reardon, Sean F., and Stephen W. Raudenbush. 2013. "Under what assumptions do site-by-treatment instruments identify average causal effects?" *Sociological Methods & Research* 42.2: 143-163.
- Schochet, Peter Z. 2008. *Guidelines for multiple testing in impact evaluations of educational interventions. Final report*. Princeton, NJ: Mathematica Policy Research, Inc. Retrieved from <http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED502199>
- Somers, Marie-Andree, Pei Zhu, Robin Jacob, and Howard Bloom. 2013. *The validity and precision of the comparative interrupted time series design and the difference-in-difference design in educational evaluation*. New York, NY: MDRC working paper in research methodology.
- Westfall, Peter H., and S. Stanley Young. 1993. *Resampling-based multiple testing: Examples and methods for p-value adjustment*. Vol. 279. Hoboken, NJ: John Wiley & Sons.

Williams, Sonya and Richard Hendra. 2018. *The Effects of Subsidized and Transitional Employment Programs on Noneconomic Well-Being*. New York, NY: MDRC.