

Evaluation of the Family Unification Program

OMB Information Collection Request *New Collection*

Supporting Statement

Part B

June 2018

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

4th Floor, Mary E. Switzer Building
330 C St, SW
Washington, DC 20201

Project Officer:
Kathleen Dwyer, Ph.D.

The evaluation of the Family Unification Program (FUP) will consist of an impact study and an implementation study. The goal of the evaluation is to determine whether FUP has significant impacts on the key outcomes of family preservation and reunification for families receiving vouchers and associated services in up to 10 sites, document how the program is implemented across the sites, and identify how the variations in implementation across sites might explain impact differences across sites. This evaluation will utilize both qualitative and quantitative data collection and analysis.

The impact study will employ a randomized controlled trial (RCT). The proposed FUP evaluation will randomly assign families to be referred either to receive a FUP voucher or to receive services as usual. Data collected for the impact study will come from existing administrative data. The data will be analyzed using an Intent to Treat model to determine the impact of the program on child welfare outcomes. In addition, we will explore the effects of the program on the mediating outcome of housing stability as measured by emergency shelter stays. The implementation study data collection will consist of interviews, focus groups, and forms collecting program data to be filled out by staff from organizations involved in implementing FUP as well as existing public housing authority and child welfare administrative data. In addition, data collection will include in-depth interviews with parents. Implementation study data will be analyzed using qualitative and quantitative methods to understand how the program is implemented across sites.

B1. Respondent Universe and Sampling Methods

Target Population

The target population for this study includes all families identified as eligible for FUP in the selected evaluation sites. The FUP program serves families involved in the child welfare system whose lack of adequate housing is a primary factor in the imminent removal of a child from the household (preservation families) or where housing is a barrier to reunification (reunification families). FUP also serves youth transitioning from foster care who do not have adequate housing; however, this population is not the focus of this evaluation. In addition, families must meet the minimum criteria for voucher eligibility, which are: the family has inadequate housing, has an income below 30 percent of the area median income, has no adult sex offender living in the household, and has no adult living in the household who has been convicted of producing methamphetamines in public housing. Individual public housing authorities may have additional eligibility criteria around criminal history or housing history (e.g., no felonies in the last 3 years or the family does not owe arrears to the housing authority).

Sampling Frame

This study will use family level randomization within a site, we will not select families from the population as a whole that could be referred to FUP. Instead, we will select sites that administer FUP and will include in the study sample any family identified as eligible for FUP in that site. There are approximately 242 sites, which have administered FUP vouchers at some point. However, many of these sites do

not have any vouchers currently available and therefore would not be appropriate candidates for random assignment.

In April 2018, HUD released the first Notice of Funding Availability (NOFA) for FUP since 2010.¹ According to the 2018 NOFA, HUD expects to make approximately 60 awards of up to 25, 50 or 100 vouchers. The number of vouchers awarded to each site will be based on the size of the public housing authority's (PHA) voucher program (i.e. number of Housing Choice Vouchers) and the identified need for FUP vouchers. Each PHA can choose to allocate the vouchers to families, youth, or both. With the infusion of new vouchers, sites receiving 50 or 100 vouchers should be able to support a random assignment study of FUP families. The evaluation team will select up to 10 sites from this group of 60 awardees. This group will likely be a selective set of sites, because vouchers will be awarded to sites based on the merits and quality of their funding proposals. The application review criteria listed in the NOFA suggest that the awardees will be PHAs that have developed a relationship with the public child welfare agency (PCWA) and the Continuum of Care (CoC), have high FUP utilization rates, and will be in good legal standing. In addition, the review criteria suggest that PHAs awarded FUP vouchers may allocate a greater proportion of their vouchers to youth relative to families than they have in the past. The criteria also suggest that families who receive a FUP voucher may be likely to get additional services such as housing search assistance, financial assistance, post-move counseling, case management, and be enrolled in Family Self-Sufficiency programs.

Sample Design

The study sample will be comprised of all families referred to the PHA for FUP vouchers in selected sites.

After the awardees are announced in the fall of 2018, the evaluation team will engage in activities to identify potential evaluation sites from among the 60 awardees. From the list of awardees the evaluation team will select sites for an initial list of eligible sites based on the number of vouchers the site will receive and the number they intend to allocate to families. Sites selected for this list will be those that:

- Receive 50 or 100 vouchers AND
- Plan to distribute at least 40 of these vouchers to families

For cost reasons, we will not include any sites that receive fewer than 50 vouchers. Sites that receive more than 50 but less than 100 vouchers were in a category that made them eligible for 100 vouchers, but did not demonstrate sufficient need for 100 vouchers. The fact that they did not receive 100 vouchers implies they do not have sufficient need to support a control group, so the team will not contact them about the evaluation.

The evaluation team will assess sites on these two criteria using information from HUD's award announcement and any available information from sites' applications. The potential evaluation sites will then be randomly selected from this initial eligible list, until there are enough sites to support the sample size requirements of 930 families total or 350 vouchers across all sites. We will have preliminary phone

¹ https://www.hud.gov/sites/dfiles/PIH/documents/FUPNOFA2017_2018FR-6100-N-41.pdf

conversations with each of these sites to determine whether they meet all of our criteria (Appendix A). If a site does not meet all criteria or chooses not to participate, we will randomly select a replacement site to reach out to.

In our selection of potential evaluation sites, we will consider two types of stratification, depending on the characteristics of the final list of 60 awardees: stratifying the initial eligible list based on rental vacancy rates (RVR) to ensure sufficient numbers of awardees in low versus high rental vacancy areas (a low vacancy rates implies a tight housing market) and/or based on the number of vouchers the site was awarded to ensure inclusion of sufficient numbers of large and medium size sites.

If we stratify by RVR, we will be able to examine differences in the potential effectiveness of FUP due to housing market differences, albeit across a small number of sites. Rental vacancy rates will be based on the latest estimates from the Census Bureau's American Community Survey. We would use two strata, one stratum including sites above the median RVR and the other stratum including sites below the median RVR, based on the RVR distribution for all 2018 FUP awardees receiving 50 or 100 vouchers. If we stratify by the size of the voucher award we can examine potential differences by site size. We would use two strata, one including sites receiving 50 vouchers and the other including sites receiving 100 vouchers. The decision of whether to stratify and which type or combination of stratification to use will be based on how well the initial list of eligible sites can support stratification. For example, if the awardees distribute evenly across the strata, then stratification would be warranted to avoid having too many sites within one stratum by chance. Conversely, if the sites are most present in one stratum, then stratification would not be necessary.

If we stratify, the awardees on the initial list of eligible sites will be grouped according to their stratum and we will randomly select from each stratum. If we stratify by the size of the voucher award, the number of sites per stratum will be based on the share of the total 2018 FUP awardees receiving 50 or 100 vouchers in each stratum.² Using these proportions, if all selected sites combined do not achieve the desired sample size, we will need to select one additional site. We will select at random from the stratum that comes closest to providing the number of vouchers needed to reach the desired sample size. If we stratify by RVR and do not meet the necessary sample size, we will randomly select one stratum, then randomly select one site from that stratum.

Once potential sites have been selected, the evaluation team will hold site-specific follow-up conversations with the PHA and PCWA agency heads, or their designates to get information about their FUP programs. These conversations will use the guide for recruitment with PHA and PCWA administrators (appendix A). In particular, the information gained from these conversations will be used to assess the site's ability and willingness to participate in an RCT. We will use two primary criteria to evaluate whether the site can support an RCT:

- *Size of the eligible population.* Based on a site's estimate of their eligible population, a site must have enough FUP eligible families expected within

² By definition, the share of sites in each RVR stratum will be ½ since we would use the median RVR as the cutoff. The share of sites in each voucher size group will depend on HUD's awards.

twelve months of the project implementation (currently estimated as 10/24/2018) to provide a reasonable size control group while utilizing all their vouchers. We anticipate randomizing at a 1:1 ratio so that each site will provide equal sized treatment and control groups, though we will consider other ratios if necessary.

- *Referral process.* The site must have a referral process, or be willing to adopt a referral process, that allows for appropriate randomization of families to treatment or control groups. If the PCWA has an existing waitlist, they must be willing to reassess the family's housing status and randomly assign those on the waitlist. A family should be pre-screened for eligibility by the PCWA before randomization occurs so that ineligible families are not randomized into the study.³

To summarize, we will attempt to represent the population of sites receiving either 50 or 100 vouchers under the 2018 NOFA by randomly selecting sites. Given the small number of sites to be selected, we may stratify the selection to avoid selecting mostly similar sites by chance, though the decision to stratify will be determined after examining the diversity in the eligible sites. If a selected site does not have a sufficient eligible population size or does not have a referral process conducive to randomization, then a new site will be randomly selected (from the same stratum if stratification is used) and the evaluation team will undergo the same process of outreach and screening. By randomly selecting a replacement, representativeness is maintained. We will reach out to sites using an outreach email (attachment 5) and the FUP evaluation information sheet (attachment 7).

Within each evaluation site there will be a sample of families for the impact study and a sample of other stakeholders for the implementation study.

Impact study

Within each site, families will be identified as potentially eligible by either the PCWA, the CoC, or a referral partner. The PCWA must certify that all families meet the child welfare eligibility criteria, i.e. have an open child welfare case and housing is a primary factor in either the imminent removal of a child into out of home care or a barrier to reunification with a child in out of home care. The PCWA will also conduct a pre-screen of families for voucher eligibility, i.e. eligible income level, criminal background, and other requirements of the partner PHA. However, final voucher eligibility determination will be made by the PHA. Families identified by the PCWA as eligible for FUP will be randomized into two groups. Families randomized into the treatment group will be referred to FUP and families randomized into the control group will receive services as usual. All randomized families will be included in the study.

Implementation study

For a majority of the semi-structured interviews with program administrators, there will be only one potential respondent per site. For instance, based on a past implementation study of FUP, there will likely only be one PCWA FUP program manager per site (Cunningham et al, 2015). For the focus groups with frontline workers, we will reach out to the FUP program manager at each organization to identify and recruit focus group participants. For the in-depth interviews with

³The final determination of eligibility will be made by the PHA.

parents, families will be randomly selected from the pool of families that were issued vouchers, signed a lease, and the parent consented to be contacted for the interview.

Sample Size and the Precision Needed for Key Impact Study Estimates

As discussed in detail below, we will conduct Intent to Treat (ITT) and Treatment on the Treated (TOT) analyses of all study outcomes. The ITT estimate is defined as the difference between the average outcomes for those randomized to the treatment group and those randomized to the control group, adjusting for pre-randomization covariates. All eligible families randomized to the treatment group will be counted in the treatment group, regardless of whether they sign a lease with FUP. All eligible families randomized to the control group will be counted in the control group, even if they inadvertently are enrolled in FUP. To estimate the effect of FUP for families who actually sign a lease, we will estimate the TOT estimate using an "instrumental variable" estimation procedure (IV) (Angrist, Imbens, & Rubins, 1996). The IV estimate is per family served, which accounts for the fact that some families referred to FUP may not sign a lease and that some families in the control group may end up sign a lease through FUP (though this is unlikely).

We are seeking a sample size commensurate with detecting at least an effect size of 0.2, considered "small, but meaningful" (Cohen 1988), among those who receive the treatment, the TOT estimate. To determine the appropriate sample size, we estimated Minimum Detectable Effect (MDE) sizes for both the ITT and TOT estimates. The effect size scales the effect of the intervention on an outcome by the control group's standard deviation on that outcome, normalizing the size of the impacts across outcomes. For example, if the control group has a reunification rate of 65 percent (SD=0.48) and the treatment group has a reunification rate of 75 percent (SD=0.43), then the effect size would be the difference in the rates, (10 percentage points), divided by the control group standard deviation (0.48), which equals 0.21.

We made initial assumptions to allow us to calculate the MDEs.⁴ These include:

- Administrative data will be available for all families in the evaluation
- FUP awardees in the evaluation will have either 50 vouchers (medium sites) or 100 vouchers (large sites)
- 20 percent of each site's vouchers will be allocated to youth yielding the number of vouchers per site for families to be:
 - o Medium sites: 40 vouchers for families
 - o Large sites: 80 vouchers for families
- 13 percent of family vouchers will turnover within the year and will continue to be used for families, yielding the number of family vouchers per site of:
 - o Medium sites: an additional 5 vouchers available per year, total of 45 vouchers for families
 - o Large sites: an additional 10 vouchers available per year, total of 90 vouchers for families
- 85 percent of treatment families will complete the housing process and sign a lease; the remainder of vouchers will be re-issued to families, yielding the following total number of treatment group families per site:

⁴ These assumptions are primarily based on Pergamit, Cunningham and Hanson (2017).

- o Medium sites: 53 treatment families to use the 45 voucher slots
- o Large sites: 105 treatment families to use the 90 voucher slots
- 1 treatment to 1 control randomization ratio
 - o Medium sites: 106 families randomized to get 53 treatment families
 - o Large sites: 210 families randomized to get 105 treatment families.

To be able to detect meaningful effect sizes, that is effect sizes of 0.20 or greater in the TOT, we need a sample of 930 families based on the assumptions above. This would require 9 medium sites. To account for the possibility that our assumptions could be incorrect, our plan accounts for up to 10 sites.

Analysis Plan

Impact study

We will conduct ITT and TOT analyses of the outcomes. The ITT estimate is defined as the difference between the average outcomes for those randomized to FUP, the treatment group, and those randomized to the control group, adjusting for pre-randomization covariates. All eligible families randomized to the treatment population will be counted in the treatment population, regardless of whether they engage with FUP. All eligible families randomized to the control population will be counted in the control population, even if they inadvertently are enrolled in FUP.

One key issue when estimating the effects of FUP on child welfare involvement is the level of analysis. The program provides vouchers and services to families; however, the outcomes are at the child level (e.g. removed, reunified). We intend to estimate the impacts at both the family level and the child level as a robustness check but to primarily report outcomes at the child level. Child level because it is more intuitive to model the outcomes at this level.

The ITT estimate is measured as the average child outcomes for the treatment population less the average child outcomes for the control population. Specifically, the ITT estimate would be measured using the regression equation below:

$$Y_i = \alpha + \beta^T T_i + \sum_{n=1}^N \beta^n X_i^n + \varepsilon_i$$

Where Y_i is the outcome for each child, i , that was randomly assigned; T_i is an indicator equal to 1 for children in families who were assigned to the treatment group and 0 for children in families assigned to the control group; β^T is the parameter of the ITT effect on the outcome (Y_i); X^n is a vector of pre-randomization covariates; β^n is the vector of coefficients on the covariates, X^n ; and ε is the regression error term. For continuous outcomes, we will estimate an OLS regression model and for binary outcomes, we will estimate a logit or probit model. The inclusion of the pre-randomization covariates is intended to improve the precision of the estimates. The exact covariates will be finalized after reviewing the data for data quality and completeness. This regression will be estimated with clustered standard errors at the family level.

The sample will be evaluated for equivalence between the treatment and control groups on observable pre-randomization variables. Although random assignment is intended to create two equivalent groups, small samples can result in some differences between the groups by chance. Variables that show differences between the two groups at $p \leq .05$, that is, with at least 95 percent confidence they are different, will be included as covariates in the regressions.

As discussed above, not all families referred for FUP vouchers will obtain a lease. These families are in the treatment group, but do not receive the treatment. Many program and practice stakeholders will want to know whether the program helped those who received vouchers. To estimate the effect of FUP for families who actually sign a lease we will also estimate the TOT estimate using an "instrumental variable" estimation procedure (IV) (Angrist, Imbens, & Rubins, 1996). The IV estimate is per child served, among those who comply with their referral assignment, which accounts for the fact that some families referred to FUP may not sign a lease and that some people in the control group may end up leasing up through FUP. For example, all study participants can be divided into three types of individuals: (1) those who will always sign a lease with FUP regardless of whether they are referred to it or not; (2) those who will never sign a lease with FUP even if they are referred to it; and (3) those who comply with whatever referral assignment they are given, whether it is to sign a lease with FUP or to remain in the control group. The IV estimate represents the effect of signing a lease with FUP on study outcomes among this third group, the compliers. In the special circumstance where decisions to comply are independent of the study outcomes, the IV estimate also represents the average treatment effect.

The IV estimate scales up the ITT estimate by the difference between the treatment and control groups' fractions enrolled in FUP. Conceptually, we will estimate the effect of referring a family to FUP on leasing up with FUP in the same manner as calculating the ITT above, except that the dependent variable in the model will be enrollment:

$$P_i = \alpha + \delta^T T_i + \sum_{n=1}^N \delta^n X_i^n + \varepsilon_i$$

where P_i is 1 if the child, i , enrolled in the program, regardless of whether they were in the treatment group or the control group. Enrollment will be defined as the participant having an initial housing lease-up date through FUP. T_i is an indicator equal to 1 for children in families assigned to the treatment group and 0 for children in families assigned to the control group. δ^T is the parameter of the effect of getting randomly assigned into treatment on actual enrollment (P_i). X_i^n is a vector of prerandomization covariates, and δ^n is the vector of coefficients on the covariates, X_i^n . ε is the regression error term. The IV estimate is the ratio of the two estimates:

$$\text{TOT estimate} = x = \frac{\beta^T}{\delta^T}$$

In practice, the two equations are estimated simultaneously using a two-stage least squares estimation procedure. In the first stage, the dependent variable (enrolling in the program) is regressed on the exogenous covariates plus the instrument (randomization into treatment). In the second stage, fitted values from the first-stage regression are plugged directly into the structural equation in place of the endogenous regressor (enrolling in the program). We will include the same covariates as used in the ITT regression.

In addition to our main analysis, we plan to conduct subgroup analysis by family type and site. Specifically, we would like to see how the program effects vary for preservation families and reunification families. There are many reasons that FUP may affect these families differently. Preservation and reunification families may be very different. Reunification cases are likely to be more severe since the caseworker decided to remove the child. In addition, there are different mechanisms for preservation families to remain intact than for reunification families to return to being intact. For a preservation family to remain intact implies preventing a removal, which typically will be largely based on caseworker judgement, while for a reunification family to become intact through a child returning home involves a court decision. We will run regressions separately for preservation and reunification families using the same methodologies described above.

There are many reasons that FUP may affect families differently across sites. One is that program implementation could vary widely. For instance, one site could provide many intensive support services, whereas another site could provide few support services. Child welfare practices may differ across sites (e.g. when a case is opened, when a child is removed), leading to differences in the child welfare population from which families are identified. Subjective interpretations of inadequate housing could systematically differ by sites leading to differences in families deemed eligible. On the housing side, PHAs may differ in their voucher eligibility criteria and application processes. Furthermore, one site could have a very tight housing market leading to longer waits and/or lower rates of signing a lease than other sites. We will run regressions separately for each site using the same methodologies described above to explore potential differential impacts across sites. The implementation study will document site program differences and help explain why we might see different impacts across sites.

Implementation study

The implementation study will use a combination of qualitative and quantitative data analysis. Qualitative data analysis will combine information from the various data sources. The semi-structured interview guides and focus group protocols we have developed to guide qualitative data collection include discussion topics and questions that reflect key implementation study research questions, as will the tools used for extracting information from program documents. The evaluation team will take detailed notes during qualitative data collection. We will develop a coding scheme to organize the data into themes or topic areas. Notes will be coded (tagged based on the theme or topic for which they are relevant) and analyzed using a qualitative analysis software package, such as NVivo.

Although analysis of data for the implementation study will primarily draw on qualitative methods, the evaluation team will also produce descriptive statistics based on program data and measures of fidelity. Specifically, the evaluation team will look at how families progress through the leasing process using the quantitative data collected through the dashboard. This analysis will present the share of families randomized that complete a housing application, receive a voucher, are deemed eligible, sign a lease, lose their voucher, and exit housing. In addition, it will include descriptive statistics on the reasons for voucher denial, voucher loss, and housing exit. If there is sufficient variation, we will run regression analyses to determine what factors were correlated with obtaining a voucher, leasing up, and exiting housing. Additional descriptive analyses will look at the service receipt reported in the housing assistance questionnaire and ongoing services questionnaire.

Expected Response Rate

The sample is constructed by first selecting sites, then gathering administrative data on families at each site (impact study) and interviewing individuals involved in some capacity in the FUP program at each site (implementation study). We first discuss expected response rates in our site selection and recruitment followed by expected response rates for data and interviews collected within each site.

Site level:

There is no existing evidence on the likely response rate, or share of eligible sites that will agree to participate in the evaluation. For planning purposes, we are assuming 66%. The 2018 NOFA specifically states that sites are expected to participate in an evaluation and abide by and facilitate the random assignment procedures. The NOFA also requires that the Memorandum of Understanding between the PHA and PCWA explicitly state that both the PHA and PCWA agree to cooperate with any program evaluation efforts undertaken by HUD, or a HUD-approved contractor. However, we recognize that some sites may be reluctant to participate in the evaluation, because they have concerns about randomization, worry about potential staff burden associated with participating in an evaluation, or worry about having their operations scrutinized, especially by a federal evaluation. During our conversations with sites, we will work to gain sites' participation. We will explain how the random assignment process would work and clarify any questions they may have. We will explain that an RCT is like a lottery and is a fair way of allocating a scarce resource like housing vouchers, when there are more families who need the vouchers than there are vouchers available. We will explain how we will work to reduce as much as possible the burden an evaluation poses on their staff. We will also talk through how their participation helps the federal government understand the impact of the program. OPRE and the evaluation team will work with federal partners, including HUD and the Children's Bureau, to promote grantee engagement in the study and data sharing.

With-in site level:

Impact study

At the individual level, we expect a 100 percent response rate for the impact study. The impact study will exclusively use administrative data. We assume that we will have administrative data for all families in the study. Past studies conducting

impact studies on housing programs for child welfare involved families using administrative data (Pergamit et al. 2017 and Pergamit et al. 2016) have had 99 percent response rates on the impact study data collection.

Implementation study

We also expect a 100 percent response rate for the implementation study data collected from program staff. Past evaluations conducting implementation studies on housing programs for child welfare involved families (Cunningham et al. 2015 and Cunningham et al. 2014) have had 100 percent response rates on implementation study data collection. As discussed in Part A (A9), a past study with a similar population (Holcomb et al. 2015) achieved an average of a 75 percent response rate for in-depth parent interviews. As with that study, by including an incentive, we assume a similar response rate for this study.

Expected Item Non-Response Rate for Critical Questions

Site level:

Impact study

At the site level, we expect a high response rate for the critical questions in the impact study. The critical questions for the impact study are whether the family was preserved, was reunified, and had a subsequent report of abuse and neglect. These data will come from the public child welfare agencies. These data already exist in a form required for submission to the federal government for the Administrative Foster Care and Adoption Reporting System (AFCARS) and the National Child Abuse and Neglect Data System (NCANDS). For this reason, we believe the data should be available and complete. Past evaluations (Pergamit et al. 2017 and Pergamit et al. 2016), have been able to obtain measures of family preservation and reunification through administrative data for all eight sites included in those evaluations. However, these past evaluations were not able to obtain reports of abuse and neglect from one of the eight sites. For this reason, we expect a 100 percent response rate on preservation and reunification at the site level, and an 88 percent response rate on reports of abuse and neglect.

Implementation study

When a site agrees to participate in the evaluation, they will agree to all of the data collection activities involved including collecting the program data and participating in the site visits.

With-in site level:

Impact study

At the individual level, we expect a high response rate. In past studies, we had been able to obtain administrative data on about 99 percent of individuals in child welfare administrative data sets (Pergamit et al. 2017, Pergamit et al. 2016).

Implementation study

The implementation study data collection questions for program staff do not include sensitive topics and are designed to ask questions appropriate questions to each respondent. We therefore expect a 100 percent response rate for those implementation study questions. Although the in-depth interviews with parents

include some sensitive questions, similar questions have successfully been asked of similar respondents in other data collection efforts, such as in the in-depth parent interviews conducted for the supportive housing study of child welfare involved families (Cunningham et al. 2014). We do not expect significant item non-response for these interviews.

B2. Procedures for Collection of Information

Data collection will take place in the form of phone interviews, three site visits, program data collection, and administrative data collection. Table A3 in Supporting Statement A summarizes all of the data collection that will be conducted for the evaluation.

Preliminary calls

Once awards are made we will conduct phone interviews with PHA and PCWA agency heads (or their designates) to collect information relevant for site selection and recruitment (see appendix A) and evaluation design (see appendices B and C). We will first reach out to sites using an outreach email (attachment 5) and the FUP evaluation information sheet (attachment 7). At this stage, we plan to collect data through phone interviews—rather than in-person visits—to expedite the collection of this information as this information will need to be collected from 10 or more sites within two weeks of award.

First site visit

The first site visit will focus on setting up the evaluation and will include one semi-structured interview with the PCWA agency head. We begin with this site visit and interview because referrals all must go through the PCWA. We will gather information on the child welfare system in which the FUP program operates; details on the local FUP program including eligibility, referrals, and screening; the structure of the FUP partnership; services offered; and the community context. At least one senior researcher from the evaluation team will attend each site visit, and visits will generally last one day. The discussion guide questions for this interview are designed to elicit nuanced responses, and the evaluation team will need to probe when answers are vague, ambiguous, or when we want to obtain more specific or in-depth information. A phone interview will be used if the PCWA agency management is unavailable during the site visit. At the start of the interview, the evaluation team will ask the respondents for verbal consent to participate using the informed consent for staff (attachment 4). The team will cover the following: the study's purpose and funder, the nature of the information that will be collected, how the information will be used, the potential benefits and risks of participating, and assurance that participation in the study is voluntary. The team will also inform study participants that they may choose to skip any questions or stop participating in the interview at any time.

Program data collection at referral

The housing status form (appendix O) will be completed by child welfare caseworkers for all families on the workers' caseloads when the site first implements FUP using the vouchers received in 2018 as well as for all families entering the child welfare system while those vouchers are available during the first

year of implementation. We only propose this one data point to minimize the burden a more regular assessment would impose on caseworkers fearing that greater reporting burden could reduce overall compliance (though caseworkers may find this form a convenient way to regularly review their caseloads). The housing status forms will be gathered and sent to the evaluation team every month in batches by the PCWA.

The referral form (appendix P) will be filled out by caseworkers only for families referred to FUP. All FUP programs use some sort of referral form. We propose to have sites use a common referral form to ensure uniform data collection across sites. For some FUP sites, this standard form may include items the site would not have otherwise collected. Sites will be able to modify the format or create an electronic version of the form; they may also add, but not delete, items, as long as it is clear which items are part of the OMB-approved information collection. To minimize burden, the information on the housing status form is mimicked on the referral form, so the information can be transferred directly to the referral form.

The randomization tool (appendix Q) is an online system used when a family is referred to FUP. The randomization tool collects the family's identification number, whether they are preservation or reunification, and a copy of the referral form (appendix P). This format allows PCWA staff to easily randomize families and receive their status in real time. The randomization tool will allow the evaluation team to monitor the randomization in real time to ensure that enough families are being referred to FUP.

Program data collection after referral

FUP managers at the PCWA and PHA will be asked to enter information into a dashboard (appendix T) to track progress of referred families through the FUP application, search, and leasing process. The FUP program manager at the PCWA will fill out the elements in the dashboard prior to the date of application submission. The PHA FUP program manager will fill out the elements in the dashboard at and after the date of application submission. Each element in the dashboard will be filled out once per family by the appropriate staff person. The dashboard will be sent by the PCWA and PHA program managers to the evaluation team every other week during the first two years of implementation. The dashboard will allow the evaluation team to track on a regular basis how families that are randomized move through the referral pathway and into housing. In combination with the randomization tool, the information collected in the dashboard will allow the evaluation team to check that all families randomized to treatment were referred to housing and that no families randomized to control were referred. It will also allow the team to monitor whether families are getting housed and how long it is taking them to get housed. At the end of each year, the dashboard information will be verified against PHA administrative data.

The housing assistance questionnaire (appendix R) will be filled out by staff providing families with services to families once for each family immediately after a family is leased up into housing or upon voucher denial. The on-going services questionnaire (appendix S) will be filled out by staff providing families with services to families once for each family six months after they have signed a lease or upon exiting housing if the family exits before six months in housing. Together these

forms capture what assistance and services FUP families receive, which can be compared with the site's FUP logic model for fidelity. Both forms will be sent to the evaluation team every month in batches by whichever staff is providing services to families. The evaluation team will monitor the timely submission of these forms using the dashboard. These forms are necessary as there is unlikely a consistent and synthesized recording of these types of assistance and services for FUP families.

Second site visit

The second site visit will focus on collecting information about the program structure and housing assistance. This site visit will include one-on-one interviews with PHA and CoC agency heads, FUP program management (PCWA and PHA liaisons), focus groups with frontline workers (child welfare caseworkers, PHA intake workers, CoC frontline workers, and frontline workers from other referral partners), and in-depth interviews with parents to learn how families experience the FUP program. As with the first site visit, at least one senior researcher from the evaluation team will attend each site visit, and visits will last between one and three days. Information will be collected through semi-structured interviews and focus groups. We will conduct one-on-one interviews, rather than focus groups, with staff at different levels of an organization or from different organizations because staff may not be as forthright in front of their superiors or individuals from outside their organization. The evaluation team will conduct focus groups with individuals from the same level and organization. Phone interviews will be used when specific staff are unavailable during the site visit. At the start of each interview or focus group, the evaluation team will ask respondents for their verbal consent to participate using the informed consent for staff (attachment 4).

During this site visit, the evaluation team will also conduct one-on-one in-depth interviews with parents in person in their homes or at a location that they prefer. During the voucher issuance process, all parents who are issued a voucher through FUP by the PHA will be administered an informed consent (attachment 1) allowing the PHA to share the parent's contact information with the Urban Institute. The PHA frontline worker will go through the form with the parent and, if the parent consents to share their contact information, will provide that information to the evaluation team. The evaluation team will randomly select six parents at each site from among those who signed a lease and gave consent to be contacted. We will recruit parents for participation through phone calls using the outreach phone call script for parents (attachment 6). Prior to the interview the evaluation team will read the informed consent for parents (attachment 3) and if the parent agrees to the interview they will sign the informed consent. If the parent agrees to be recorded they will sign again on the appropriate line. A copy of this consent form will be provided to parents.

Third site visit

The third site visit will focus on collecting information about the services provided to families. This site visit will include one-on-one interviews with agency management, FUP program management, focus groups with frontline workers, and in-depth

interviews with parents. A new random selection of parents will be used to recruit for participation in the in-depth interviews. The same procedures for data collection used in the second site visit will be used in the third site visit.

Administrative data

The elements to be collected from administrative records are outlined in the administrative data list (appendix U). For each agency, this activity will consist of four components. The first will be having site-specific conversations with the staff most familiar with the data to understand what data are available and its structure and quality. During this conversation, we will establish a timeline and procedures for transferring the data to the evaluation team. The second component will be the first data pull, which will occur at one-year after the last family is randomized. This first pull will require the data administrator at the agency to identify families randomized in the study and extract the relevant data elements. The PCWA should have a list of families randomized in the study. However, the PHA and the CoC may need to have the PCWA share the list of families to identify them in their data. The third component will be a follow-up conversation with the data staff to answer any questions or address any concerns that have come up around the first pull. The fourth component will be the second and final data pull, which will occur at two years after the last family is randomized. Since each agency already collects the information that we are requesting, pulling administrative data is the lowest burden method of collecting the necessary information. The evaluation team will develop data sharing agreements with each agency, PCWA, PHA and CoC, as necessary for data access and will also facilitate data sharing agreements across agencies as necessary.

B3. Methods to Maximize Response Rates and Deal with Nonresponse

Expected Response Rates

The implementation study relies on information collected via one-on-one interviews, focus groups, and program data. Although interviews and focus groups may face nonresponse, program data should be accessible for all families referred to FUP. Furthermore, the impact study relies exclusively on administrative data for the primary outcomes of interest and we therefore do not expect response rates and nonresponse bias to be an issue.

Interviews and focus groups

Program Staff

Previous implementation studies of FUP (Cunningham et al. 2015) and of a supportive housing program for child welfare involved families (Cunningham et al. 2014) have had 100 percent response rates for all interviews and focus groups. ACF anticipates that once management of a program agree to participate in the evaluation, they will encourage program staff to participate in all study activities. The evaluation team will conduct interviews and focus groups on site and will be

scheduled with adequate notice to facilitate matching interview and focus group times with staff's schedules.

Parents

We expect a 75 percent response rate for in-depth family interviews based on past studies (Holcomb et al. 2015). Additionally, the evaluation team will select locations, dates and times that are most convenient for parents, and propose to offer an incentive to offset incidental costs to maximize response rates.

Program data

ACF expects a 100 percent rate to all data collection activities on the referral form and other program data. The referral form will be required to be filled out to refer a family to FUP, therefore we expect that this form will be filled out for all families. As this form will be collected at the time of randomization, we will check when each form is submitted to ensure that this form is complete. The dashboard will be filled out twice per month. We expect that when sites agree to participate in the evaluation, they will agree to fill out this dashboard. The information collected in the dashboard is information that sites must track, though they may not do so in the form of the dashboard. To minimize burden, we will allow sites to use their own forms or systems to collect these data if their forms contain the same information in a different format. Finally, we will closely monitor the submission of the services questionnaires which will be submitted on an ongoing basis. We will check against the other information we collect to ensure that we have received all appropriate forms.

Administrative data

ACF expects a high response rate on the administrative data. A majority of the administrative data we intend to collect already exist in a form required for submission to the federal government including the Administrative Foster Care and Adoption Reporting System (AFCARS), National Child Abuse and Neglect Data System (NCANDS), Homeless Management Information System (HMIS) and required submissions to U.S. Department of Housing and Urban Development. For this reason, we believe the data will generally be available and complete. From past studies (Pergamit et al. 2017, and Pergamit et al. 2016), we expect most sites to have the data necessary.

At the individual level, we expect a high response rate. In similar, prior evaluations we have been able to obtain administrative data on about 99 percent of individuals in child welfare administrative data sets (Pergamit et al. 2017, Pergamit et al. 2016). However, there are a few data elements that may not be available. While we expect our primary outcomes of reunification, preservation and new reports of abuse and neglect to be available, our secondary outcomes of case opening, case closure, and emergency shelter stays may have lower rates of availability.

One open question is whether FUP leads to cases closing faster or whether cases are held open to facilitate service provision under FUP. However, not all child welfare agencies maintain information on case opening and closing dates or they may define a case in a way that makes it difficult to assess when a case is open. In similar evaluations (Pergamit et al. 2016 and Pergamit et al. 2017) we were not able to obtain these case opening and case closing data for three out of eight sites. If

case opening and case closing dates do not exist in electronic form, we will work with the site to identify if there is an alternative method of acquiring the data. Should we not be able to acquire these dates for one or more sites, we will estimate impacts using only sites that supplied these items.

We also expect that not all sites will have complete HMIS data. Coverage of shelters in the HMIS data can vary widely from site to site, therefore the HMIS data in some sites may not cover all shelter stays. In addition, CoC's will need consent from individuals to share data. Typically a shelter will collect a standard consent form for every person who enters the shelter. If a person has had a shelter stay and has not consented to have their data shared, we will not observe their stays. We will assess both the coverage of the HMIS data and the consent rates in our conversations with the CoC to determine the quality of the data. Unfortunately, for emergency shelter stays, if some shelters do not report into the HMIS, or if the family has not given consent to share their information with researchers, we have no other way to acquire these data. For emergency shelter stays, we will estimate impacts using data from sites for which we have the data items.

Dealing with Nonresponse and Nonresponse Bias

Interviews with agency and program managers, focus groups with front line workers, interviews with parents, and program data all support the implementation study. Although each respondent brings their own perspective on the program, it is the sum of all the information that is used to describe the program. Thus nonresponse and nonresponse bias are less important for the implementation study than for the impact study. The impact study derives its findings exclusively from administrative data, most of which exists in the child welfare administrative data system and can be expected to be fairly complete for the primary outcomes of interest. Below we discuss any issues of nonresponse and nonresponse bias associated with each type of data collection.

Interviews and focus groups

Program Staff

If staff are not available to participate in interviews, we will work closely with program leaders to either identify other staff with similar knowledge, or ways to schedule telephone interviews or follow-up conversations. Because the evaluation is voluntary, any member of the program may choose not to participate. This may lead to nonresponse bias in the results if those who do not respond are somehow different from those who respond. Any substantial nonresponse from members of a program, essentially if agencies at a site do not make their staff available, will be reported as a study limitation and may result in excluding the program from the implementation study analysis.

Parents

The sample for the parent interviews will be selected randomly from those parents who receive a FUP voucher and agree to be contacted by the Urban Institute for an interview. If parents decline to participate or are not available to participate in interviews, we will randomly select additional parents to interview. Because the evaluation is voluntary, any parent may choose not to participate. This may lead to

nonresponse bias in the results if those who do not respond are somehow different from those who respond. We will examine demographic and other data from the referral form to assess the representativeness of the respondents. Substantial nonresponse from parents will be reported as a study limitation.

Program data

If program data have incomplete or missing information, the evaluation team will note the data elements that are missing and work with program staff to understand what factors drive missingness and develop procedures for collecting the information more consistently. For example, a site might send back referral forms which are incomplete before referring to FUP. In the analysis, we will consider whether imputation techniques to account for any missing data will be useful and feasible based on how extensive the missingness is and the availability of variables useful for imputation. If we decide to impute any missing variables, we will use Stata to run multiple imputation (Little and Rubin, 2002).

Administrative data

Similarly, if administrative data have incomplete or missing information, the evaluation team will note the data elements that are missing and work with program staff to understand what factors drive missingness and develop procedures for collecting the information more consistently. For example, if the HMIS is attempting to pull data on those who participated in the program the evaluation team may suggest they try matching on different data elements when trying to pull data on individuals in the study. In the analysis, we will consider whether imputation techniques to account for any missing data will be useful and feasible based on how extensive the missingness is and the availability of variables useful for imputation. If we decide to impute any missing variables, we will use Stata to run multiple imputation (Little and Rubin, 2002).

Maximizing Response Rates

Interviews and focus groups

Program Staff

To maximize response rates the evaluation team will work closely with the identified programs to ensure program leaders and staff fully understand what to expect from participating in the interviews and that they are interested in participating. All site visits will be scheduled with sufficient advance notice and with input from the programs to ensure the dates are convenient for most staff. Experience with similar evaluations suggest participation can be maximized by sufficiently informing staff about the process. The team will also seek buy-in from program leadership so that leaders encourage their staff and other partners and stakeholders to participate.

Parents

To maximize response rates and minimize nonresponse bias in the parent interviews, the evaluation team proposes to provide each parent a \$35 gift card to offset the costs of taking part in the interview and will work with parents to schedule the interview during a time that is convenient. The incentive is intended to assist with transportation costs, child care, or other expenses that might prevent

some in our target population from participating – i.e., those with the greatest financial challenges or other barriers, and whose absence could contribute to nonresponse bias.

Program data

We will collect program data on an ongoing basis. To ensure we have complete data, we will monitor data collection closely and work with FUP program staff to increase response rates where we find these falling short.

Administrative data

We will collect the administrative data at two points in time, the first time helping ensure that the data are complete before the final data is sent to us.

B4. Tests of Procedures or Methods to be Undertaken

As discussed in part A, the team has based most of the guides and program data collection forms on prior studies (Cunningham et al. 2015, Cunningham et al. 2014, Cunningham et al. 2016) or on program requirements and have found them to be effective. The evaluation team does not plan to collect any survey data. We will explore piloting the housing status, referral, housing assistance, and ongoing services form with up to nine existing FUP sites, if we can find sites with meaningful numbers of vouchers turning over within a month (i.e. five or more).

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

The information for this study is being collected by the Urban Institute on behalf of ACF. Co-Principal Investigators Michael Pergamit and Mark Courtney led development of the study design plan and data collection protocols, and will oversee collection and analysis of data gathered through on-site interviews and telephone interviews.

The agency responsible for receiving and approving contract deliverables is:
The Office of Planning, Research, and Evaluation (OPRE),
Administration for Children and Families (ACF)
U.S. Department of Health and Human Services

The Federal project officer for this project is Kathleen Dwyer.

References

- Angrist, Joshua, Guido W. Imbens and Donald Rubin. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association*. 91.434 (1996): 444-455.
- Cohen, Jacob. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed., L. Erlbaum Associates, 1988.
- Courtney, Mark E., Michael Pergamit, Maria Woolverton, and Marla McDaniel. 2014. "Challenges to learning from experiments: Lessons from evaluating independent living services." In *From Evidence to Outcomes in Child Welfare: An International Reader*, Aron Shlonsky and Rami Benbenishty, eds. New York: Oxford University Press.
- Cunningham, Mary, Michael Pergamit, Maeve Gearing, Simone Zhang, Brent Howell. 2014. *Supportive Housing for High-Need Families in the Child Welfare System*. Urban Institute. <https://www.urban.org/research/publication/supportive-housing-high-need-families-child-welfare-system>
- Cunningham, Mary, Michael Pergamit, Abigail Baum, Jessica Luna. 2015. *Helping Families Involved in the Child Welfare System Achieve Housing Stability: Implementation of the Family Unification Program in Eight Sites*. Urban Institute. <https://www.urban.org/sites/default/files/publication/41621/2000105-Helping-Families-Involved-in-the-Child-Welfare-System-Achieve-Housing-Stability.pdf>
- Cunningham, Mary, Michael Pergamit, Sarah Gillespie, Devlin Hanson, Shiva Kooragayala. 2016. *Denver Supportive Housing Social Impact Bond Initiative: Evaluation and Research Design*. Urban Institute. <https://www.urban.org/sites/default/files/publication/79041/2000690-Denver-Supportive-Housing-Social-Impact-Bond-Initiative-Evaluation-and-Research-Design.pdf>
- Holcomb, Pamela, Kathryn Edin, Jeffrey Max, Alford Young, Jr., Angela Valdovinos D'Angelo, Daniel Friend, Elizabeth Clary, Waldo E. Johnson, Jr. 2015. *In Their Own Voices: The Hopes and Struggles of Responsible Fatherhood Program Participants in the Parents and Children Together Evaluation*. OPRE Report Number 2015-67. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Little, Roderick J.A., Donald B. Rubin. 2002 *Statistical analysis with missing data (second edition)*. Wiley, NJ.
- Pergamit, Michael, Mary Cunningham, Julia Gelatt, Devlin Hanson. 2016. *Analysis Plan for Interim Impact Study: Supportive Housing for Child Welfare Families Research Partnership*. Urban Institute. <https://www.urban.org/sites/default/files/publication/80986/2000802-Analysis-Plan-for-Interim-Impact-Study-Supportive-Housing-for-Child-Welfare-Families-Research-Partnership.pdf>
- Pergamit, Michael, Mary Cunningham, and Devlin Hanson. "The impact of family unification housing vouchers on child welfare outcomes." *American Journal of Community Psychology* 60.1-2 (2017): 103-113.