



Technical Appendix: Bringing Families Home Program Evaluation

KRISTA RUFFINI, EMILY PUTNAM-HORNSTEIN, JANE MAULDON AND HIMAL SUTHAR



MAY 2024

Bringing Families Home Program Evaluation

Technical Appendix

Identifying the effect of BFH

This study aims to examine the effect of the Bringing Families Home (BFH) program using a matched control research design. Matched control designs are common to estimate the causal effect of a program, particularly in settings where a Randomized Control Trial (RCT) is not feasible.

To provide intuition behind the matched control research design, it is helpful to first outline the logic behind RCTs. An RCT would randomly assign “treatment” — in this case, an offer to participate in BFH to a subset of families that met the eligibility criteria. Since participation would be randomly assigned, families that did not participate would be otherwise similar to participating families on average by design, both on observed and unobserved characteristics. With an RCT, researchers can compare outcomes between the treatment and control group to identify the causal effect of the program, and this approach is often referred to as the “gold standard” of program evaluation.

In many settings, including BFH, conducting an RCT is not feasible. One simple approach in these instances might be to compare participating families with those that did not participate in the program. However, as shown in Tables 2-3 of the main text, families that participated in BFH were different from the broader population of families that were involved with the child welfare system between 2017 and 2019 on multiple dimensions. For example, children in BFH-participating families were more likely to be non-Hispanic White and less likely to be Hispanic than the general child welfare caseload. Participants also tended to be slightly younger when they first entered the child welfare system, and had greater involvement with the child welfare system, as measured by the number of previous referrals and the likelihood of a previous out-of-home placement.

If these differences in demographic characteristics or past child welfare involvement are correlated with housing and family outcomes, simple comparisons between BFH participants and non-participants would conflate the causal effect of BFH with differences in outcomes that arise because participants and non-participants have differing characteristics (“selection bias”).

An alternative approach for identifying the effect of the program is to use quasi-experimental or “as-if” random variation in program variation and compare similar populations, some of which received the program and others that did not. Matched control designs are one type of such quasi-experiment that attempt to overcome selection bias by creating a control group that has similar characteristics to program participants.

We implement a matched control design by leveraging the fact that BFH was a new program, first implemented in summer of 2017. Families who were involved in the child welfare and homelessness systems prior to 2017 could not participate because the program did not exist. Therefore, the timing of

BFH implementation provides the key source of “as-if” random variation in whether families participated in BFH.

In the coarsened exact matching approach used in this study, each BFH participant is only compared to individuals in families that have similar demographic characteristics and similar involvement with the child welfare or housing systems. This approach allows us to estimate the effect of the program if, conditional on the characteristics included in the matching function, there are no other factors affecting both BFH participation and future child welfare or housing outcomes. While we construct the match from a rich set of covariates, one shortcoming to this approach is that we can only condition on factors we observe in the data. Therefore, there may be remaining *unobserved* (to the analyst) differences between participating and non-participating families.

The following sections provide greater detail on the matching process and control group construction for child welfare and housing outcomes.

Child welfare outcomes

All child welfare analyses are conducted at the child level. There are at least two advantages to this approach. First, as families may consist of multiple children that have different interactions with the child welfare system, focusing on the child level allows us to capture all of these types of involvement. Second, the child-level focus considerably boosts sample sizes and increases statistical power, allowing us to detect small-to-moderate changes in outcomes.

For each child whose families received BFH, we use historical CWS-CMS data to identify potential control group observations as the group of children whose families did not receive BFH, but who have the same child welfare case type (Family Maintenance or Family Reunification), race, ethnicity, gender, and county of residence as the “focal” treatment observation. We additionally limit the pool of potential control group observations who were within a 3-year age band when their first child welfare case was opened.

The next step in constructing the control group is to define a placebo “treatment” date for each control group observation – that is, the hypothetical date at which control group member’s family *would have* begun receiving BFH service had they participated in the program.¹ We define the placebo treatment date to ensure balance between the treatment and the control group on the length of time with an open child welfare case before they began receiving BFH services. Specifically, we determine the number of days each BFH participant had an open child welfare case, $prev_i$. We then assign all potential control group observations a placebo start date as the date the child is first observed in the child welfare data, plus the number of days $prev_i$ for which they had an open case. An example calculating the placebo treatment is provided in Table A1 below.

¹ The need to construct a placebo treatment date for each control group observation prevents using alternative matching approaches, such as estimators that rely on propensity score matching.

Appendix Table A1: Constructing placebo BFH start dates for the comparison group

Step 1: Calculate cumulative days in child welfare system for BFH participants

ID	BFH	CW_start 1	CW_end 1	CW_star t2	CW_en d2	BFH_start	Cumul_days_C W
A	1	10/31/2017	11/15/2017	12/16/2018		12/31/2018	30
B	0	1/1/2017	2/1/2017	3/1/2017	5/1/2017		

(11/15/2017 - 10/31/2017) + (12/31/2018-12/16/2018) = 30 days of child welfare system involvement before the BFH start date.

Step 2: Assign BFH cumulative days in child welfare system for non-BFH participants

ID	BFH	CW_start 1	CW_end 1	CW_star t2	CW_en d2	BFH_start	Cumul_days_C W
A	1	10/31/2017	11/15/2017	12/16/2018		12/31/2018	30
B	0	1/1/2017	2/1/2017	3/1/2017	5/1/2017	1/31/2017	30

1/31/2017 - 1/1/2017 = 30 days of child welfare system involvement before the placebo BFH start date.

Once the placebo treatment date is assigned for each control group member, we determine the age at BFH entry for the treated observation and drop control group observations who are not within 3-years of this age restriction. We additionally drop observations who would have been older than 18 when they entered BFH. Finally, for computational ease, we limit the control group to up to 200 observations that have a placebo start date that is closest to the treated individual's actual start date. We repeat this process for each treatment group member, sampling with replacement among children whose families did not receive BFH.

Housing outcomes

The matched comparison group is constructed in a similar fashion for housing outcomes, with several modifications. First, since HMIS participation is more comprehensively recorded at the adult level, all housing outcomes are examined at the adult or guardian level. Second, since the set of observable characteristics reported in HMIS is different from those reported in the child welfare data, the information used to construct the match is slightly different. Specifically, we perform an exact match on county of residence, household size at HMIS entry, race, ethnicity, and gender, and limit to comparisons that are within a 10-year age band of the BFH adult participant. We determine the time between the initial HMIS placement and BFH entry for each BFH participant, and assign the placebo BFH start date for the comparison group as the same number of days relative to each individual's first HMIS involvement, similar

to the approach in Table A1. This final difference in the matching frameworks accounts for the fact that individuals have fewer unique time periods receiving services that tend to be briefer in duration. As with the child welfare sample, we limit the housing control group to up to 200 observations that have a placebo start date that is closest to the treated individual’s actual start date. We repeat this process for each treatment group member, sampling with replacement among adults whose families did not receive BFH.

Regression framework

All matches for both the housing and child welfare outcomes are conducted by sampling with replacement so a single non-BFH participant can exist in the control group for multiple BFH participants, albeit with a different BFH “entry” date. Over the full period for which we have child welfare and homelessness data, economic conditions — including the labor and housing market — changed in ways that could affect housing opportunities and family stability. Many of these dimensions are observed in the data; however, others, such as the local homeless service environment are not. Therefore, in order to compare groups that face similar policy and area conditions, we limit the control group for each BFH participant to the 200 observations whose placebo start dates are closest to the participant’s actual start date. Since the size of the matched control group varies across BFH participants, we weigh each control observation by $(1/(N-1))$ where N is the number of observations in each demographic group. Each BFH participant receives a weight of 1.

With the matched control group, we then estimate the effect of BFH as the difference in outcomes between children and adults whose families participated in BFH and control group members who have the same demographic characteristics and past program participation, denoted a . Conditional on the variables used to construct the match, the only observed dissimilarity is BFH participation. The estimating equation then takes the form:

$$y_{i(a)} = \beta BFH_{i(a)} + X_i' \gamma + \theta_a + \varepsilon_{i(a)}$$

Where $BFH_{i(a)}$ equals one for BFH participants and zero for non-participants, and θ_a is a matched-group fixed effect to capture demographic characteristics that are associated with participation in BFH and eventual housing and child welfare outcomes (*confounding variables*). We additionally include controls for the number of past referrals and time since the current child welfare case opening in the vector X_i' (child welfare outcomes) in order to account for remaining differences in the treatment and control group that were not captured by the matching algorithm. $\varepsilon_{i(a)}$ is an iid error term.

In this equation, the reported coefficient β reports the percentage point change in outcome y due to the BFH program. In order to provide a sense of the magnitude of the effects, we also report the average value among the control group (“CG mean” in Tables 5-7). Accordingly, the percentage change in each outcome is given by $\beta/CGmean$. Finally, all tables report cluster-robust standard errors clustered by county in order to account for differences in program design across counties.