Attachment F: HPOG Impact Analysis Plan: Estimating 15- and 30-36 Month Impacts

Alignment of Research Questions with Impact Study Analytic Techniques

Impact Research Questions	Analytic Techniques
1. What impacts do HPOG programs have on outcomes of interest?	Impact estimation: $I = Y_t - Y_c$ OR $y_i = \alpha + \delta T_i + \beta X_i + \varepsilon_i$
2. To what extent do these impacts vary by subgroups of interest?	Subgroup analyses: divide sample along selected baseline characteristics, estimate impacts, as in #1
3. To what extent does HPOG program participation (in particular components, with particular dosage) affect outcomes?	Individual participation analyses: (1) basic participation: impose conventional no-show correction (Bloom, 1984), if assumption of no effect on no-shows applies; (2) component/dosage analyses: predict component use/participation, create subgroups, divide sample, estimate impacts, as in #1 (Peck, 2003)
4. To what extent do various HPOG program models/ component have varying impacts?	Site-level component analyses: examine raw impacts across groups of common sites; use multi-level modeling to estimate the relative contributions (Bloom, Hill & Riccio, 2003)
5. To what extent do specific program enhancements have impacts, relative to the "standard" HPOG program?	Impact estimation: $y_i = \alpha + \delta_1 T_{1i} + \delta_2 T_{2i} + \beta X_i + \varepsilon_i$

1) What impacts do HPOG programs have on outcomes of interest?

The basic analytic approach involves computing the difference in mean outcomes between the treatment and control groups as: $I = \bar{Y}_t - \bar{Y}_c$. To control for other factors and increase the precision of the impact estimates, we estimate the following regression model:

$$y_i = \alpha + \delta T_i + \beta X_i + \varepsilon_i$$

where

y is the outcome;

T is the treatment indicator (treatment = 1; control = 0);

α is the intercept (interpreted as the control mean outcome);

δ is the impact of the treatment;

X is a vector of baseline characteristics;

β are the coefficients on the baseline characteristics;

e is the residual; and

the subscript i indexes individuals.

This impact estimate is referred to as the Intent-to-Treat (ITT) estimator. Those in the treatment group are offered *access* to the HPOG program, and not all of them will end up participating in the program itself. The difference in treatment-control outcomes represents the impact of the "intent" to treat, or making the program available to treatment group members.

2) To what extent do these impacts vary by subgroups of interest?

To conduct an analysis of the program's impact on selected subgroups, we will follow the analytic approach described above, dividing the sample along selected baseline characteristics already included in the PRS. We will explore variation in the magnitude of subgroup effects, testing also whether they are statistically significantly distinct.

3) To what extent does HPOG program participation (in particular components, with particular dosage) have an impact on outcomes of interest?

This research question focuses on the individuals' specific experience of HPOG program components, both overall and specifically across various packages of services that they might receive. The first research question considers the effect of *offering access* to the HPOG program; and here we analyze specifically the impacts on those individuals who actually participate. In its most basic form, this analysis produces what we refer to as the impact of the "treatment on the treated" (TOT). TOT estimates are computed by dividing the ITT estimate by the proportion of treatment group members who actually participate in program services. This results in an unbiased estimate of the program's impact on participants, under the assumption that there is no effect on no-shows (Bloom, 1984).

In addition, we will estimate the relative effects of participating in specific program components using a pathways analysis. This approach would follow the technique proposed in Peck (2003) to identify what Orr (1999) calls "endogenous subgroups" on the basis of baseline (exogenous) characteristics. Identifying the treatment path—or combinations of program components—that treatment group members follow as a function of their baseline characteristics, one can estimate impacts associated with having followed that path (or participated in that combination of program components) using control group members with the same predicted pathways, had they been assigned to treatment. The supplemental baseline survey items were added in large part to assist in the identification of control group members' predicted pathways. An advantage of this approach is that it retains the integrity of the experimental design while permitting estimation of impacts associated with variation in treatment characteristics.

In brief, the analytic approach involves the following steps:

- Step 1 – Use baseline characteristics to create program-related subgroups
- Step 2 – Split the sample into subgroups, analyze in the same way as #1
- Step 3 – Given distribution of predicted v. actual (correct subgroup placement), reweight (if desirable; otherwise, interpret results as likely subgroup members' impacts)

4) To what extent do various HPOG program models or components have varying impacts?

To address this question, we will exploit the cross-grantee or cross-site variation in what programs do, rather than the individual-level variation in program experiences used to address question #3. To the extent that we can identify sites that are common enough in their program characteristics to classify as using a common intervention "model," we will first estimate ITT

effects for sites using a common model. In addition, we will conduct a multi-level analysis, following the approach of Bloom, Hill & Riccio (2003). This approach involves creating program-specific variables and analyzing their relative effect. A two-level model would include experimental group status (treatment or control) on level one, with site/program characteristics, including economic conditions, on level two, effectively interacting these characteristics with treatment to measure how impacts vary with the measured factors.

This analytic approach, as slightly modified from Bloom, Hill & Riccio (2003) proposes the following two sets of equations to capture the individual and site-level effects of the intervention:

Level 1 (individuals):

(Eq. 1) $$y_{ji} = \alpha_j + \beta_j P_{ji} + \sum_k \delta_k C_{kji} + \sum_k \gamma_k C_{kji} T_{ji} + \varepsilon_{ji}$$

Level 2 (grantees/sites):

(Eq. 2) $$\beta_j = \beta_0 + \sum_k \phi_n P_{nj} + \psi E_j + \mu_j$$

(Eq. 3) $$\alpha_j = \alpha_0 + \lambda E_j + v_j$$

where

y_{ji} = the outcome measure for each sample member;

T_{ji} = a zero/one treatment group indicator for each sample member;

C_{kji} = client characteristic k for each sample member (grand mean centered);

α_j = the conditional control group mean outcome for each grantee/site;

β_j = the conditional program impact for each grantee/site;

δ_k = the effect of characteristic k on the control group mean outcome;

γ_k = the effect of characteristic k on the program impact; and

ε_{ji} = a random component of the outcome for each sample member;

and where

P_{nj} = program characteristic n for each grantee/site (grand mean centered);

E_j = economic environment for each grantee/site (grand mean centered);

β_0 = the grand mean impact;

ϕ_n = the effect of program characteristic n on program impacts;

ψ = the effect of economic environment on program impacts; and

μ_j = a random component of program impacts for each grantee/site;

and where

α_0 = the grand mean control group earnings;

λ = the effect of economic environment on control group earnings; and

v_j = a random component of control group earnings for each grantee/site.

5) To what extent do specific program enhancements have on impacts, relative to the "standard" HPOG program?

The research design involves two treatment arms and a control group in a subset of grantees/sites where they have agreed to implement an experimental test of a particular enhancement. The analysis that supports this research question is essentially the same as what is detailed under question #1, with the simple addition of a second treatment group indicator as follows:

$$y_i = \alpha + \delta_1 T_{1i} + \delta_2 T_{2i} + \beta X_i + \varepsilon_i$$

where

y is the outcome;

T_1 is the standard treatment group indicator (= 1 if standard treatment; = 0 otherwise);

T_2 is the enhanced treatment group indicator (= 1 if enhanced treatment; = 0 otherwise);

α is the intercept (interpreted as the control mean outcome);

δ_1 is the impact of the standard treatment;

δ_2 is the impact of the enhanced treatment;

X is a vector of baseline characteristics;

β are the coefficients on the baseline characteristics;

e is the residual; and

the subscript i indexes individuals.