Analysis Plan

We will estimate impacts using both Intent-to-Treat (ITT) and Treatment-on-the-Treated (TOT) analyses. The ITT estimate is defined as the difference between the average outcomes for those randomized to GI (i.e., the treatment group) and those randomized to the control group, adjusting for pre-randomization covariates. In our ITT analysis, all eligible participants randomized to the treatment population will be counted in the treatment population, regardless of whether they receive GI. All eligible participants randomized to the control population will be counted in the control population, even if they inadvertently are enrolled in GI. The TOT analysis estimates the impact of the program for those who actually receive the GI intervention. For this evaluation, we plan to define the treated group as those who receive at least one GI payment but will also examine the implementation data to see if dosage varies. ITT estimates are of interest to policymakers who want to know whether offering an intervention is effective at addressing the problem it was chosen to solve. TOT estimates are of interest to program and practice stakeholders who want to know how the program impacted those who actually received the services, but are subject to selection bias if there are systematic differences between participants randomized to the treatment population who receive at least one GI payment and those who do not.

For the former foster youth population, we plan to pool the two sites conducting an RCT. For the pregnant population, we plan to pool the four sites conducting an RCT. We believe that there will be enough uniformity between sites that pooling will be possible. We will also conduct site-specific analyses to examine how outcomes varied by site.

In addition to our ITT and TOT analyses, we will conduct growth curve analyses to examine participant growth trajectories over the multiple survey waves. A growth curve modeling approach will allow us to more thoroughly understand patterns in the data over time from our ITT analysis.

Intent to Treat

The ITT estimate is measured as the average participant outcomes for the treatment population less the average participant outcomes for the control population. We control for pre-randomization covariates using a regression framework. Specifically, the ITT estimate, π_Y, 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 participant, *i*, that was randomly assigned, T_i is an indicator equal to 1 for families who were assigned to the treatment group and 0 for 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. We plan to include the following covariates for the former foster youth models: site, race, gender, baseline income, baseline household size, and child welfare history (e.g., age at first removal, lifetime days spent in care, number of placements). We plan to include the following covariates for the pregnant population models: site, race, baseline income, pre-pregnancy risk factors (e.g., previous pre-term birth, chronic health condition), and child welfare history (both in their own childhood and as a parent). The sample will also 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 $p <= .05$, that is, with at least 95 percent confidence they are different, will be included as additional covariates in the regressions.

Treatment on Treated

All eligible participants assigned to the treatment group will be given the opportunity to either accept or decline GI, and it is possible that some may decline due to the impact of GI on their benefits or other reasons. These participants 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 actually received GI. To estimate the effect of GI for participants who actually receive GI payments we will also estimate the TOT estimate using an "instrumental variable" estimation procedure (IV) (Angrist, Imbens, & Rubins, 1996). The IV estimate is per participant served, among those who comply with their referral assignment, which accounts for the fact that some participants referred to GI may opt out and that some people in the control group may inadvertently end up receiving GI. For example, all study participants can be divided into three

types of individuals: (1) those who will always receive GI regardless of whether they are referred to it or not; (2) those who will never receive GI even if they are referred to it; and (3) those who comply with whatever referral assignment they are given, whether it is to receive GI or to remain in the control group. The IV estimate represents the effect of receiving GI 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 GI. Conceptually, we will estimate the effect of referring a family to GI on receiving GI 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 participant, 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 received at least one GI payment. T_i is an indicator equal to 1 for individuals assigned to the treatment group and 0 for individuals 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^n is a vector of pre-randomization covariates, and δ^n is the vector of coefficients on the covariates, X^n. ε is the regression error term. The IV estimate is the ratio of the two estimates:

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.

Subgroup Analyses

In addition to our main analysis, we plan to conduct subgroup analyses. Possible subgroups include gender, race/ethnicity, urbanicity, aspects of placement history, history of homelessness or

educational attainment at baseline, and mother's age at childbirth. Subgroups to be analyzed will be decided in collaboration with the CDSS and the GI pilot program Guidance Committee.

There are many reasons that GI may affect participants differently across sites. One is that program implementation could vary widely. For instance, one site could provide many support services alongside GI, whereas another site could provide few support services. Additionally, sites have different eligibility criteria leading to differences in the population being served. 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.

Adjusting for Multiple Hypothesis Testing

We plan to test numerous outcomes, which increases the chance that we will make a Type I error (i.e., incorrectly reject a null hypothesis). We will adjust for multiple hypothesis testing through a Bonferroni correction or another similar method.