Introduction

Per the South Carolina Pay-For-Success (PFS) contract (Annex E) the Abdul Latif Jameel Poverty Action Lab – North America (J-PAL)[*] prepared this evaluation plan for measuring the impact of Nurse-Family Partnership (NFP) on the four PFS outcome metrics (preterm birth, child injury, healthy birth spacing, and low-income zip codes (LIZCs)), which were predetermined by the signatories of the contract.[†] This document describes the evaluation for the PFS reports J-PAL will create for these four PFS outcome metrics. Additionally and separately, J-PAL will conduct an academic evaluation which assesses NFP's impact on a wide array of outcomes available in administrative data and over a longer period of time, potentially including but not limited to: maternal and child health, mortality, education, abuse and neglect, employment and earnings, use of government programs and social services, criminal justice involvement, among others. For the academic reports that analyze this broader range of outcomes, J-PAL will prepare a separate analysis plan that is congruent in terms of analytical choices and sample construction but may differ in defining outcomes of interest.

As the independent evaluator, J-PAL will independently conduct all key aspects of the evaluation, including: random assignment, the collection and maintenance of administrative data used in the impact analysis, the impact analyses, and the reporting of study findings. J-PAL is ultimately responsible for maintaining and protecting the confidential data required to conduct this assessment. Per the PFS contract (Annex E, Article II, Section M), J-PAL will provide a copy of any draft presentation or publication manuscript to NFP and South Carolina Department of Health and Human Services (SCDHHS) for a 10-day review and comment period. However, the results of such review are non-binding with respect to J-PAL's rights for presentation or publication.

This plan describes the details of the evaluation to the best of J-PAL's knowledge as of January 1st, 2020. Currently, over 5,433 participants have enrolled in the study, out of which we have obtained outcome data for 3,795 treatment and control group clients, 1,249 of which belong to the control group. This report has been updated to reflect analysis decisions based on this

[*] The Abdul Latif Jameel Poverty Action Lab North America, is a research center within the Economics Department at the Massachusetts Institute of Technology that conducts rigorous impact evaluations, policy outreach, and capacity building with the mission of reducing poverty by ensuring that policy is informed by scientific evidence. For this project, the legal entity, the Institutional Review Board of record, and data receiving entity is the Harvard University T.H. Chan School of Public Health.

[†] In order to produce the PFS impact assessment reports in accordance with the reporting dates specified in the contract, J-PAL North America will produce impact assessments of preterm birth, child injury, and healthy birth spacing, and a process measure of enrollment into LIZCs based on a partial sample. Once data on the full sample becomes available, J-PAL will produce updated impact assessments using data from the full sample, which may differ from those in the initial report and those from later academic reports.

matched information. These decisions were made based on the data quality and distribution of outcome data for the control group only.

1 Impact Evaluation

1.1 Research questions

This analysis plan focuses on the PFS-specific outcomes. The primary research questions for this PFS evaluation are:

- What is the average impact of NFP for the intervention group relative to the control group on preterm birth, healthy birth intervals, and child injury?
- What percent of the intervention group resides primarily in LIZCs at the time of study enrollment?

1.2 Description of the intervention

Sample members randomized into the intervention group will be offered a spot in NFP. NFP is a prenatal and infancy home-visiting program for low-income, first-time mothers and their families. Registered nurses begin visiting clients as early in the pregnancy as possible within the first 28 weeks of gestation to help the mother-to-be make informed choices about her own health and the health of her baby, and to facilitate her overall empowerment and self-sufficiency. Nurses continue regular visits with the family until the child is two years old.

1.3 Control group conditions

Sample members randomized into the control group will receive the standard of care. They are not offered a spot in NFP, but can still receive all community and medical services to which they would otherwise be entitled. Clients on Medicaid are eligible for up to two post-partum home visits. In some sites, NFP nurses provide these home visits. J-PAL will report the share of control group and intervention group members receiving other home visiting services (to the extent that those programs are captured in the data provided to J-PAL), which may help to provide context for the NFP impact findings.

1.4 Sample identification, selection, and assignment

1.4.1 Referral sources

Potential NFP clients are identified through two main referral channels. First, referral partners, such as local health care providers, schools, and WIC agencies, directly refer potential clients to an implementing agency (IA) with the client's permission. In these cases, the potential clients know to expect a call from an IA and may know something about NFP. Second, the SCDHHS regularly sends a list of newly enrolled, first-time pregnant women from the Medicaid eligibility

database to NFP. After a referral is received, the IA attempts to contact the potential client and determine their eligibility for the program.

In addition to these two main channels, some clients are self-referred or referred by a family member or friend. NFP launched a community outreach campaign in South Carolina to raise awareness of the program. The campaign includes digital and print advertisements as well as outreach posters and flyers.

1.4.2 Eligibility

IA staff conduct an initial eligibility screening over the phone and schedule an intake visit for those who are interested. The intake visit and final eligibility determination are completed in person. To be eligible for the study, potential clients must meet the following eligibility criteria:

Eligibility Criteria	Process for assessing criteria
Female	Potential Client's self-report
No previous live births	Potential Client's self-report
Age at least 15 years old	Potential Client's self-report[1]
Non-institutional residency	NFP staff's assessment that the Potential Client is not incarcerated or living in a lock-down facility[1]
Currently pregnant	Potential Client's self-report
Gestation period less than 27 weeks, 7 days	Potential Client's self-report of estimated due date
Income level meets Medicaid eligibility criteria	NFP staff's verification of enrollment and/or eligibility in QuickCheck (State eligibility determination program) based on the Potential Client's self-reported income. If not currently enrolled in Medicaid, Potential Client must actively apply.[2]
Live within an area serviced by an NFP Implementing Agency	Potential Client's self-report of current mailing address
Not currently enrolled in the study	Potential Client's self-report, with verification check that no client with the same name and date of birth have been previously enrolled.[3]
No language barrier	NFP staff's assessment that the program would be therapeutic for the client if she were randomly assigned to receive NFP.

[1]Potential clients who are fourteen (14) years of age and younger and those who are incarcerated/living in a lock-down facility are the only potential clients who can receive NFP services without randomization during the study period. They are not counted as sample members.
[2] Women must be enrolled or have completed an application (except for income-eligible women who are not citizens or permanent residents).
[3] NFP staff follow up with Potential Clients following warning messages. If NFP staff believe that the match detected by SurveyCTO is not a true match, study enrollment continues.

1.4.3 Consent process

If NFP program staff determine that the applicant is not eligible for the study, they will provide the applicant with a list of other available resources in the community, and their interaction will end there. If the applicant is eligible, the program staff will invite her to participate in the study. During the informed consent process, NFP program staff review the consent form with the potential participant, check her understanding of what it means to be in the study, and answer any questions she has about the study. Once the applicant is fully informed, the program staff will ask her to electronically sign and date the informed consent form if she agrees to participate in the study. The program staff conducting study enrollment also will sign and date the consent form electronically. If she does not agree to participate in the study, the program staff document her decision not to participate.

1.4.4 Baseline interview

Prior to randomization, eligible applicants who provide written consent will be asked to complete a brief thirty minute baseline interview. Using encrypted tablets and Version 2.40 of SurveyCTO (2018)[‡], the NFP program staff will ask the participant questions about her demographics, health, feelings, use of social services, and what she hopes to get from the home visiting program. For completing the baseline interview, each participant will receive a $25 gift card as compensation for her time. Data collected at baseline will be used to describe the characteristics of the study sample, to find their administrative records, to assess the baseline equivalence of the intervention and control groups at the point of randomization, and to provide baseline covariates for the impact models.

1.4.5 Randomization

The study uses an individual-level random assignment design. All eligible program applicants who provide their written consent and complete the baseline interview are randomly assigned "on-the-spot" to an intervention group that is offered access to the NFP program or to a control group that is not offered the opportunity to enroll in NFP. Program staff use a pre-programed randomization function on their tablets to conduct the random assignment. Two-thirds of those

[‡] SurveyCTO [Technology for digital data collection]. (2018). Cambridge, MA. Dobility, Inc. Retrieved from http://www.surveycto.com

who consent to participate in the study will be randomized to the intervention group and one-third to the control group.

Program staff explain to the participant that the computer chooses which study group she will be assigned to, that her assignment is not dependent on any personal traits or characteristics, or answers to the baseline survey, and that she has the same chance to get NFP as everyone else.

If the applicant is assigned to the control group, the program staff explain that she was not chosen by the computer to receive NFP services, but she and her children may continue to receive services in the community that she would otherwise be entitled to receive. NFP staff provide a list of services available in the community, but will not make any specific referrals to other home visiting programs or distribute any additional materials. Control group sample members' direct involvement in the study ends at this stage.

If the applicant is assigned to the intervention group, the program staff will obtain the applicant's consent to participate in the NFP program, and either deliver or schedule her first home-visiting appointment. The program staff will also provide intervention group members with a list of other resources and programs that are available in the community.

2 Data Collection

2.1 NFP home visiting

J-PAL will use NFP program data to assess whether or not clients receive home visiting services, which is defined as receiving at least one completed NFP visit after study enrollment (see section 3.1). J-PAL will link baseline survey data to NFP program data using an NFP program ID. The program ID is generated by NFP when a client is referred. Nurses then enter program IDs into the baseline survey for both control group and intervention group mothers at study intake. In order to maintain accurate tracking of program implementation, NFP program ID numbers may be updated post-randomization for both control and intervention group mothers. In order for a mother to be included in the analytical sample for any of the three impact measures (preterm birth, healthy birth intervals, and child injury), her baseline survey must be linked to the NFP program data. Any mother for whom we cannot perform this linkage will be excluded from the impact analysis.[§]

[§] The importance of NFP program data to the impact analysis necessitates a high match rate of NFP program IDs for both the control and intervention group mothers. This means that timely reconciliation of non-matching NFP IDs (within three months of enrollment) is critical for success of the analytical strategy.

2.2 Outcome measures

J-PAL will assess NFP's impact on four outcome measures as designated by signatories of the contract. The PFS payment depends only on the point estimates, which are the calculated relative percent change in the Preterm Birth, Healthy Birth Intervals, or Child Injury PFS Outcome Metrics between the intervention and control groups, or the calculated percentage of sample members in the intervention group residing in a LIZC at the time of study enrollment. J-PAL will also estimate 95% confidence intervals to provide context for interpretation of the precision of the point estimates, noting that some of the outcome measures, especially preterm birth, may be underpowered to detect a statistically significant impact of the intervention.

1. ***Preterm Birth:*** A preterm birth is defined as a live[**], singleton[††] birth where the obstetric estimate of gestation is less than 37 completed weeks. Data will be obtained from a probabilistic match between the study sample and vital statistics birth records. The obstetric estimate of gestation may be implausibly short or long, therefore the study team will consider a record to be matched if the birth is within 120 days before or after the estimated due date reported on the baseline survey. If this record is matched as such, we consider this the index birth. This definition is designed to identify the birth that results from the pregnancy that was in gestation at the time of the baseline survey.[‡‡] Records that are outside of this eight-month window will be considered not matched and will not be included in the analysis. We will consider obstetric estimates of gestational age between 21 and 42 weeks as plausible. For any matched vital record with gestational age outside of these bounds, including matched records where the obstetric estimate of gestational age is missing, we will calculate gestational age based on the self-reported due date in the baseline survey as compared to the birth date reported in the matched vital statistics record. Any remaining outliers (still appearing with a gestational age less than 21 or greater than 42 weeks) are coded as missing and will not be included in the analysis.

   Specifically, the preterm birth outcome is coded as a 1 if a) there is a matched birth record in the vital statistics data, b) this match is listed as a "singleton" birth based on the vital statistics record, and c) the obstetric estimation of gestation listed on the vital statistics record is less than 37 complete weeks (259 days) and is equal to or greater than 21 weeks (147 days). If the obstetric estimation of gestation is not present on the vital statistics record, the preterm birth outcome is coded as a 1 if the calculated

[**] In South Carolina, birth certificates are generated for live births only.
[††] A singleton birth has a matched birth certificate record with a "singleton" indicator.
[‡‡] This strategy may capture birth from a pregnancy that shortly followed a miscarriage, but these cases should be extremely rare.

obstetric age (using self-reported gestation from the baseline survey and the date of birth from the matched vital statistics record) is less than 37 complete weeks (259 days) and is equal to or greater than 21 weeks (147 days). The preterm birth outcome is coded as 0 if a) there is a matched birth record in the vital statistics data, b) this match is listed as a "singleton" birth based on the vital statistics record, and c) the obstetric estimation of gestation listed on the vital statistics record is greater than or equal to 37 complete weeks (259 days) and is less than or equal to 42 complete weeks (294 days). If the obstetric estimation of gestation is not present on the vital statistics record, the pre-term birth outcome is coded as 0 if the calculated obstetric age (using self-reported gestation from the baseline survey and the date of birth from the matched vital statistics record) is greater than or equal to 37 complete weeks (259 days) and less than or equal to 42 complete weeks (294 days).

If there are vital statistics birth records for a study participant, but the matched records are coded as "multiples" in the vital statistics data, the preterm birth outcome is coded as missing. We exclude multiple births in our measure of preterm birth because pregnancies with multiples face different risk factors and often have very different distributions in gestational age at birth from singletons.

If no vital statistics birth data are matched to the study participant, the value of preterm birth is considered missing. This may happen for a number of reasons. The study participant's pregnancy may not have ended in a live birth; she may have moved out of state; the identifying information provided at study intake was not accurate; or, some other unknown reason. We cannot differentiate between these possibilities in the data.

Mothers who are at least 15 weeks (105 days) past their expected due date as reported in the baseline survey[§§], have a non-missing preterm birth outcome, and have non-missing program participation data are considered our reporting sample for the preterm birth impact estimation. We restrict the sample to mothers who are 15 weeks past their expected due dates to allow for potential mis-estimation of expected due dates (15 day buffer), to account for the time it takes to obtain matched birth certificates (approximately 90 days), and to ensure that the characteristics of the reporting sample are the same for treatment and control mothers (allowing for the possibility that participation in NFP alters gestational length).

Healthy Birth Interval: Healthy birth intervals are defined as having no subsequent live

[§§] Measured as the time between the date on which matched outcome data is requested from the South Carolina Department of Revenue and Fiscal Affairs (RFA) and the anticipated due date as reported in the baseline survey.

births observed within 24 months of the index birth. In the case of multiples, birth spacing will be measured as the time between the date of birth of the last child from the index pregnancy to the date of birth of the first child from the subsequent pregnancy. Data will be obtained from vital statistics birth records. There may be cases where the subsequent birth is implausibly close to the index birth. Births that occur less than or equal to 90 days from the index birth will be assumed to be multiple gestation based on guidelines provided by the National Center for Health Statistics. We will consider obstetric estimates of gestational age between 21 and 42 weeks as plausible. Subsequent births that occur between 90 days and 21 weeks (147 days) after the index birth will be considered outliers and will be dropped from the analysis.

If a study participant's index pregnancy matched to no vital statistics birth data, the healthy birth interval is coded as missing for this participant. If her index pregnancy matched to vital statistic birth data (as a live birth of either singleton or multiples), but her record matched to no subsequent birth records within 24 months of the index birth, this mother is considered having had zero subsequent birth within 24 months of the index birth.

Those mothers for whom at least 27 months and two weeks (835 days) have passed since their anticipated due date reported on the baseline survey[§§], have a non-missing healthy birth interval outcome, and have non-missing program participation data, are considered our reporting sample for the healthy birth interval impact estimation. Note that this reporting sample may include study participants whose index pregnancy resulted in a live birth with multiples. We restrict the sample to mothers who are 835 days past their expected due dates to account for the time it takes for children to reach 24 months of age (730 days), to obtain matched birth certificates (approximately 90 days), and to ensure that the characteristics of the reporting sample are the same for treatment and control mothers (allowing for the possibility that participation in NFP alters gestational length) by only observing outcomes for mothers who are at least 15 days beyond their expected due date as reported on the baseline survey.

2. ***Child Injury:*** The child injury metric is defined as the number of child emergency department visits (outpatient) and hospitalizations (inpatient) of the child(ren) from the index birth (i.e. a live birth of either singleton or multiples) due to acute injury within the 24 month period following the index date of birth. Data will be obtained from the All-Payer Health Utilization records. Acute injuries will be identified based on the following ICD-10 codes: Any of T01 to T35; T51 to T78; or T36 to T50 if the intent code is between 1 and 4; (for example, only if the sixth character of the 7-character code is between 1

and 4 such as T36.xx4x); or any S code in the primary diagnosis position or either of the first two secondary diagnoses positions of the record; or any of the ICD-10 codes above in any position when an ICD-10 external cause of injury code is indicated on the record.

One acute injury may appear on two different hospitalization records when a child is transferred between hospitals. Such cases will be counted as one hospitalization. Transfers between hospitals will be identified by the following criteria: a hospitalization record containing a discharge code of 02, 05, 43, 66, 82, 85, or 94; and a subsequent hospitalization record with an admission date within 24 hours of the discharge date of the original hospitalization.

If a study participant's index pregnancy matched to no vital statistics birth data, the child injury outcome is coded as missing for this participant. If her index pregnancy matched to vital statistics birth data, but no discharge records with the corresponding ICD codes identified as acute injury are matched to the index child, the child injury measure is coded as 0.

Those mothers for whom at least 29 months and two weeks (895 days) have passed since their anticipated due date reported on the baseline survey[§§], have a non-missing child injury outcome, and have non-missing program participation data are considered our reporting sample for the child injury impact estimation. Note that this reporting sample may include study participants whose index pregnancy resulted in a live birth with multiples. For those study participants whose index pregnancy resulted in multiple live births, their outcome is measured as the average number of injuries for all children from the index pregnancy for whom the child injury measure is not missing. We restrict the sample to mothers who are 895 days past their expected due dates to account for the time it takes for children to reach 24 months of age (730 days), to obtain matched hospital discharge data (approximately 150 days) and to observe mothers 15 days after their estimated due date to ensure that the characteristics of the reporting sample are the same for treatment and control mothers (allowing for the possibility that participation in NFP alters gestational length).

In order to interpret out impact estimation as the causal effect of NFP on the outcome measures, the likelihood of have a missing outcome measure has to be balanced between the control group and the intervention group. We will test this assumption to provide context for interpreting the estimation results in Table 2.

3. ***Coverage in LIZCs****:* Coverage in LIZCs is defined as whether sample members in the

intervention group report a primary residential address within a LIZC (as specified in Appendix B of the PFS Contract) at the time of study intake, prior to the randomization stage. Data will be obtained from the baseline survey file and cleaned and coded by J-PAL. We will calculate this metric based only on valid, 5-digit zip codes. Our team has built checks into the survey that only allow numeric answers with five digits; further, only zip codes assigned to the state of South Carolina are accepted (such that values are between 29000 and 29999). All other values will be coded as missing. The reporting sample for coverage in LIZCs will include all mothers whose baseline surveys have been submitted by the date of the most recent data pull. Table 1.A will include the date of this data pull.

2.3 Timing

To ensure sufficient time for data cleaning and analysis, J-PAL will include in the analysis for the Impact Report all data received by South Carolina Revenue and Fiscal Affairs Department (RFA) at least 6 months before the delivery date of the Impact Report, but will attempt to include more recent data if feasible. Per Annex F, Article V. of the PFS contract, J-PAL will report the results to the Executive Committee 45 days prior to the fourth and fifth year anniversary of the Commencement of the Service Delivery Period (February 15, 2020 and February 15, 2021, respectively).

2.4 Matching to outcome data

J-PAL will work with the South Carolina Revenue and Fiscal Affairs Department (RFA) to link study participants to vital records data and All-Payer Health Utilization data using the participants' identifying information.

For linkage to the South Carolina vital statistics data, J-PAL will send RFA the list of study participants with their study IDs and identifying information from the master data file. The identifying information includes patient name (first, middle, and last), complete date of birth, social security number, Medicaid number, and address. RFA will use all of the available identifying information to conduct a probabilistic match between the list of study participants and the administrative records and return the non-identifying data elements for all matched observations, along with their study ID, to J-PAL.

The identifying information in the master data file are collected prior to the randomization stage in the study enrollment process. No updates provided post-randomization that will affect matching to administrative outcome data (such as a correction of name, date of birth, or Medicaid ID) by any source will be incorporated into the master data file. While this may limit the opportunity to improve data quality for some study subjects, it minimizes bias by ensuring that data quality is comparable between the control group and the intervention group.

2.5 Attrition

Attrition due to sample member withdrawing from the study: The causal interpretation of the point-estimates depends on the assumption that attrition from the study is small and balanced. A balance table will be presented for all individuals remaining in the study. Baseline covariates will be included in the model to improve precision of estimates. Sample members who contact the evaluation team directly to request removal from the study will not be included in analysis and will also no longer receive NFP services if assigned to the treatment group.

Attrition due to data availability: The main data sources for outcomes are administrative data, and there may be "data attrition" as a result of outcome data not being available for some study participants. The possibility of data attrition is reduced by the effort to collect detailed identifying information from mothers at study enrollment, including names, SSNs, and Medicaid IDs. However, some data attrition is inevitable. For example, if the mother's identifying information is incomplete or has changed (e.g. the mother changes her name or moves out of state), we may not be able to locate her administrative records. We do not have the ability to differentiate between the reasons that an individual's outcome data may be unavailable. We handle all such cases using the same methods (Refer to section 3.3, *Missing values*).

2.6 Duplicates

The program is only for first-time mothers and each woman can enroll in the study only once. The design of the study does not allow one woman to have more than one study ID. In practice, duplicates may happen if a woman's prior study enrollment status is not accurately tracked. We use a probabilistic matching software program to identify duplicates prior to analysis. In the rare cases where a woman has multiple study enrollment records or multiple copies of the baseline survey, the earliest record with a randomization status will be kept.

Duplicate records may exist in the South Carolina vital statistics data and the All-Payer Health Utilization Data. As of January 1st, 2020, no study-ID matches to two different ID's in the data from RFA. If there is an increased prevalence of duplicate records, we will develop a plan with RFA to systematically identify and de-duplicate these records.

2.7 Data Security

2.7.1 Overview

All parties with access to individual-level data will adhere to strict data security policies. Parties agree to comply with all laws, regulations, and executive orders relating to the confidentiality of sensitive data and will adhere to all data security policies and rules regarding the reporting of any security breaches.

2.7.2 Evaluator Data Security Procedures

All study team members at J-PAL with access to confidential information will both acknowledge

a confidentiality agreement and be appropriately trained.

Identifying information will always be kept separate from the datasets that will be used for analysis. The datasets used for analysis will be stripped of personal identifiers such as name, address, record number, and replaced with a unique, scrambled, study identifier. The study team will adhere to strict data security protocols to protect participants' confidentiality.

Confidential information will never be sent via email except in encrypted files. Designated study staff will use encrypted files and/or a secure File Transfer Portal (FTP) to transfer personal identifiers to the data agencies for selecting administrative records. After the data agency staff-person selects the administrative records for the study participants, this individual will strip off any personal identifiers, keep the study ID, and send back a limited dataset to the study team using a secure method.

Confidential information will never be stored on any personal computer or portable computing device (e.g. laptop, PDA, or smart phone). All data for analysis will be stored on an institutional stationary server to which only designated study staff will have access. A written list of those designated study staff will be disclosed to and approved by the IRB.

2.7.3 Maintenance of Backup Files

In order to prevent data loss in the event of the accidental loss/deletion of the electronic file, J-PAL will electronically back up the cumulative Master Data Files no less than once per 30 days.

3 Statistical analysis of impacts

The impact analysis will examine the extent to which NFP affected each of the PFS outcomes. In testing for these effects, we will use two-tailed hypothesis test procedures, to be neutral about the direction of any effects on outcomes. Because each of the four PFS outcome metrics constitutes a different domain and only three of them involve a comparison between the control group and the intervention group, we will not correct for multiple comparisons (which would affect p-values and confidence intervals but not point estimates).

In compliance with our IRB protocol, we must stop collecting administrative data on clients who have withdrawn from the study. Because most of the study withdrawals occur before we've been able to collect any administrative records on the client or her child, we will not have any data on these study dropouts to include in the impact estimates. For this reason, we have concluded that study dropouts will not be included in the impact estimates for the four PFS outcome measures: pre-term birth, birth spacing, child injury, and low-income zip code (LIZC).

3.1 Statistical model

Analysis of the preterm birth, child injury, and the healthy birth interval metrics will compare non-missing outcomes for sample members who were randomized to the intervention group to

those who were randomized to the control group. To account for the fact that all study participants randomized to the intervention group may not go on to receive NFP services, J-PAL will calculate the PFS impact estimates using an instrumental variable approach. Intervention group status is used as an instrumental variable for receiving NFP service, defined by having at least one completed NFP nurse visit (after study intake).

Consider an outcome, Y_i, such as an indicator for preterm birth. For subject i, the estimating equation is:

$$Y_i = \pi_0 + \pi_1 I(\text{Enrolled in NFP}=1)_i + \pi_2 X_i + \vartheta_i$$

where "Enrolled in NFP" means having received at least one completed visit from NFP for service delivery.

This model will be estimated using two-stage least squares (2SLS), where the first stage is:

$$I(\text{Enrolled in NFP}=1)_i = \alpha_0 + \alpha_1 I(\text{Treatment}=1)_i + \alpha_2 X_i + \omega_i$$

where $I(\text{Treatment}=1)_i$ is an indicator variable equal to one if the subject was randomized to the intervention group and zero if the subject was randomized to the control group; X_i is a vector of covariates, specified in more detail below. These covariates should be uncorrelated with the treatment indicator because of the randomization. We include them in the model since they may increase the precision of the estimates.

This linear model estimates the local average treatment effect (LATE) of NFP on intervention group members who actually participate in NFP relative to the services consumed by the control group. This estimated effect of NFP is of policy interest because it represents the impact of NFP on those clients who are likely to participate in NFP were the program to expand and offer additional program slots through a lottery. The source of non-compliance that it explicitly captures is that some mothers randomized into the intervention group may never receive NFP services (i.e. the "enrollment rate" is less than 1). According to the enrollment protocol, no mothers in the control group should be enrolled in NFP services. To the extent that some sample members in the control group receive services from similar home visiting programs that may also affect outcomes, this model estimates the effect of NFP relative to the mix of other home-visiting programs that the control group receives, rather than relative to no home-visiting service at all. J-PAL will report the share of control group and intervention group members receiving other home visiting services (to the extent that those programs are captured in the data provided to J-PAL), which the Operations Committee may use as context in interpreting and disseminating the NFP program impact findings.

Analysis of the coverage of LIZCs will report the share of sample members in the intervention group residing in LIZCs at the time of study enrollment.

3.2 Control variables

Previous NFP trials have identified a number of characteristics that may be predictive of the three PFS impact outcomes. Such characteristics include maternal age, smoking status, maternal socio-economic status (SES), marital status, maternal child rearing attitude, maternal relationship with partner, maternal psychological resources, and maternal sense of control.

In our impact models, we will include variables from the baseline survey to measure these characteristics. Binary variables will be coded as "1", "0", or missing; for coding categorical variables, we will consider distributions across the control and treatment groups including missing values. J-PAL may include additional or alternative measures, including potential variables from administrative data in academic publication and subsequent analysis.[***] All of the control variables listed below will come from the Baseline Survey:

- Implementing Agency
  - indicators for each implementing agency enrolling study participants
- Demographics:
  - Indicator for age equal to 15, 16, or 17
  - Indicator for age equal to or greater than 28
  - Race – indicators for non-Hispanic black, non-Hispanic white, other
  - Ethnicity – indicator for Hispanic/Latina
- Gestational age at time of study enrollment
  - A continuous variable of weeks to delivery as calculated by the difference in reported due date and the survey date
- Relationship with father of the child
  - Indicator for daily interaction with father of the child
- Education
  - Indicators for a) HS diploma with no higher education and b) less than high school diploma
- Employment, income, and financial resources
  - Indicator for whether or not mother is working for pay
- Social services
  - Indicator for receiving one or more social service (i.e. TANF, SNAP, SSI, unemployment benefits, and WIC)
- Housing stability

[***] In particular, our planned control variables do not currently include an indicator for Medicaid enrollment at baseline. Preliminary analysis of self-reported data on Medicaid enrollment at baseline indicates that self-reports of Medicaid coverage have a high degree of inaccuracy. For this reason, we do not include this indicator as a planned control variable for the PFS analysis. We are working to develop alternative measures of the status of Medicaid enrollment prior to randomization that may be included as a control variable in future analyses.

    - Indicator for moving two or more times in the previous twelve months
    - Indicator for living with parents
- Access and utilization of health care, including mental health and maternal health
    - Indicator for receiving at least one antenatal care visit before time of survey
    - Indicator for having obtained care at a hospital ER in past six months
    - Indicator for receiving mental health treatment in the previous year
- Health behavior (e.g. drinking and smoking)
    - Indicator for reporting having consumed alcohol in the three months before pregnancy
    - Indicator for reporting having smoked cigarettes in the three months before pregnancy
- Psychological state/resources (measured on a PHQ-2 scale)
    - Indicator for PHQ-2 score of 3 or higher.
- Baseline measure of self-reported health
    - Indicator for self-reported health described as fair or poor.
- Maternal perceived stress level
    - Indicator for PSS-4 score of 4 or higher.
- Self-reported pregnancy risk factors
    - Indicator for pre-pregnancy weight and height yielding normal body mass index.
- Family planning
    - Indicator for responding yes to having access to a place for family planning or birth control
    - Indicator for responding yes to wanting more children one day

3.3 Missing values

Sample members who respond to survey questions with "Don't know" or "Refused to Answer" will be set to missing in the analysis. We will use case deletion for missing outcome data and dummy-variable adjustment to account for missing covariates (Puma et al., 2009). In the dummy variable adjustment method, missing covariate values are set to a constant and indicators (i.e. dummy variables) for such values are added to the impact analysis model.

3.4 Baseline balance testing

In addition to checking baseline equivalence of the full sample to assess the success of random assignment, we will also check the baseline equivalence of the reporting sample for preterm birth, healthy birth intervals, and child injury outcomes to assess whether missing outcome data affected the comparability of the intervention and control groups. The models for assessing baseline equivalence of the full and reporting samples will have the same structural form as the models that will be used to estimate impacts. Specifically, we will report a joint F-test of orthogonality on all covariates.

3.5 Descriptive information on participation in home visiting programs

In Table 3, J-PAL will report the share of mothers who have received at least one completed NFP visit as well as the average number of completed visits received for both control group and intervention group members. Table 3 will also include the share of control group and intervention group members receiving other home visiting services. Data on other home visiting services comes from The Children's Trust of South Carolina (TCT). The reporting sample for these indicators will mirror the reporting sample for preterm birth for PFS Report 1 and for birth spacing for PFS Report 2. However, the sample for this descriptive information may be smaller due to data availability.

4 Reporting

We will provide the following main table at the end of each reporting period.

Table 1 – Results Table for the PFS Impact Estimates

Dependent Variable	Control Group Mean (regression adjusted)	Intervention Group Mean (regression adjusted)	Estimated Treatment Effect (2SLS)	95% Confidence Interval (2SLS)	Sample Size for Analysis
Preterm Birth (%)					
Healthy Birth Interval (%)					
Child Injury (number of injuries)					

Table 1.A Results Table for the LIZC Metric

	Intervention Group Mean	**Sample Size for Analysis**
LIZC Coverage (%)		

We will also provide the following tables to establish a context for interpreting Table 1 and Table 1.A.

Table 2 –Table on Missing Outcomes

Dependent Variable	Control Group Share (unadjusted)	Intervention Group Share (unadjusted)	Sample Size for Analysis	P-value*
Preterm Birth Outcome Missing (%)				
Healthy Birth Interval Outcome Missing (%)				
Child Injury Outcome Missing (%)				
Zip Code Missing (%)				

***** We will report the p-value from a chi-squared test for the null hypothesis that the percentage of missing outcomes for the control group is equal to the percentage of missing outcomes for the intervention group.

Table 3 –Table on Home Visiting Services

Dependent Variable	Control Group Mean (unadjusted)	Intervention Group Mean (unadjusted)	Sample Size for Analysis	P-value*
Receiving at Least 1 Home Visit from NFP (%)	[intended to be 0]			

# of Home Visits from NFP	[intended to be 0]			
Enrolled in another Home Visiting Program** within two years after Study Enrollment (%)				

***** We will report the p-value from a chi-squared test for the null hypothesis that the percentage of missing outcomes for the control group is equal to the percentage of missing outcomes for the intervention group.

******Based on data availability.