

Emergency Clearance

Supporting Statement Part B for OMB Approval

Mother and Infant Home Visiting Program Evaluation (MIHOPE)

January 2012

PART B. COLLECTIONS OF INFORMATION EMPLOYING STATISTIC METHODS

B.1. Sampling Methods

Twelve states with 85 sites will be included in the evaluation. The average site will include 60 families (30 assigned to the home visiting program and 30 assigned to the control group). The number of families was chosen to provide enough statistical power to investigate how the effects of home visiting programs vary by program features and for key subgroups, as described later in this section. The number of sites was chosen to provide enough statistical power to investigate the link between features of local programs and impacts. The sites would be concentrated in 12 states to reduce the costs of conducting the evaluation, including costs related to site recruitment, implementation research, surveys of families in the study, and state administrative data.

Statistical power. The evaluation will include 85 sites to allow it to explore the relationship between program features and program impacts. Program features could include any aspects of the community context, implementation system, service models, organizational influences, or home visitor characteristic. For example, this analysis could explore how program impacts vary with the duration of home visits, the background and training of home visitors, the support provided by supervisors for home visitors, the clarity of the goals of the local program, the intended targets of the national model being used, and so on.

A framework for exploring the links between program features and program impacts is described in Greenberg, Meyer, Michalopoulos, and Wiseman (2003). Within this framework, the precision of the estimated relationships between program features and program impacts depends on a number of factors, including (1) the number of sites in the evaluation, (2) the precision of impact estimates within each site (which will increase with the number of families in the site), (3) the variation in characteristics across sites, (4) the number of program features to be investigated, and (5) how related the various program features are to each other. It is easier to detect differences by program feature if there are more sites, if there are more families in each site, if different sites vary more across the program feature being examined, if fewer program features are being examined at any one time, and if the program features are not closely related to one another. As an example of the last point, it may be very difficult to distinguish the effect of planned duration of home visits from the effect of actual duration, since the two are likely to be closely related in a particular site.

Table B.1 shows the minimum detectable effects of program features for several scenarios. The upper half of the table shows results for a program feature that is binary and takes on one value in half of the sites and a different value in half of the sites. For example, half of the sites might plan to visit families weekly while half would visit only every other week. The lower half of the table shows results for a continuous program feature, such as how many weeks home visits would take place. In each panel, results are presented depending on whether 10, 20, or 30 program features would be examined at one time. As noted above, the ability to detect the effects of program features will worsen as more features are examined. Finally, results for each scenario are presented for three assumptions about how highly correlated various program features are with one another. As noted above, the ability to detect the effects of program features worsens as features become more highly correlated with one another.

Mother and Infant Home Visiting Program Evaluation

Table B.1
Minimum Detectable Effects of Program Features

Type of variable	No. of variables representing program features	Correlation across program features	Administrative data	Survey or observational data
Binary, half of sites have the feature				
	10	Low	0.203	0.231
		Medium	0.213	0.243
		High	0.226	0.258
	20	Low	0.231	0.263
		Medium	0.264	0.300
		High	0.317	0.361
	30	Low	0.268	0.305
		Medium	0.348	0.397
		High	0.626	0.713
Continuous				
	10	Low	0.101	0.115
		Medium	0.107	0.122
		High	0.113	0.129
	20	Low	0.115	0.131
		Medium	0.132	0.150
		High	0.158	0.180
	30	Low	0.134	0.153
		Medium	0.174	0.198
		High	0.313	0.356

Results are the smallest true impact that would generate statistically significant impact estimates in 80 percent of studies with a similar design using two-tailed t-tests with a 10 percent significance level. No adjustment for multiple comparisons is assumed. Results are based on fixed effects estimates. Administrative data are assumed available for all families, while survey or observational data would be available for 80 percent of families. The correlation across program features is based on the R^2 statistic when one program feature is regressed on all other program features. For purposes of the calculations, a low correlation means the R^2 increases by .01 with every added feature, by .02 with every added program feature for a medium correlation, and by .03 for a high correlation.

Consider the first row of Table B.1, which shows the case where 10 program features are being examined simultaneously and there is a low correlation across them. For outcomes measured using administrative data, the model would be able to detect differences of 0.203 standard deviation between sites of one type and sites of another type. If the overall effect on an outcome were 0.15 standard deviation, for example, the study would have an 80 percent chance of finding a statistically significant relationship between the program feature and impacts if the true impact were 0.252 standard deviation in one set of sites and 0.048 standard deviation in the other set of sites.

The ability to detect an effect of a program feature is only slightly worse if the features are more highly correlated or if 20 program features are being examined. The statistical power gets considerably worse, however, if more features are being examined and the correlation across features is high. For example, the minimum detectable difference is 0.317 standard deviation (for an effect of 0.309 standard deviation in one set of sites compared with -0.009 standard deviation in the second set of sites) if 20 program features are being examined and the correlation across them is high, and the minimum detectable difference is 0.348 standard deviation if 30 features are being examined and the correlation across them is medium.

The lower half of Table B.1 shows minimum detectable effects if the program feature is continuous and normalized to have a variance of 1.0 standard deviation across sites. Because there can be greater variability in continuous variables than in binary ones, the design would have a greater ability to detect differences for such measures. For example, for a study examining 10 program features that are not highly correlated, the minimum detectable effect size of the program feature would be 0.101 standard deviation using administrative data and 0.115 standard deviation using survey data. Even for the most extreme case shown in the table — 30 highly correlated program features — the design could detect differences in impacts of 0.313 standard deviation using administrative data and 0.356 standard deviation using survey data.

These power calculations are not directly affected by the number of states that are included in the evaluation, but the number of states could affect the variation in program features that are observed across sites if sites within a state share features of their program implementation. Including more states might increase the variation in program features but would also increase the cost of the evaluation. The evaluation will include 12 states to balance the gains in statistical power from including more states with the costs to the evaluation of doing so. As discussed below, the evaluation will aim to include states where local programs vary in features such as the evidence-based model that is being used, the urbanicity of the local site, and the type of local implementing agency.

Choosing states and sites for the evaluation. As the first step for selecting the 12 states, the study team reviewed and analyzed all submitted MIECHV state plans and deemed 30 states as high priority based on the following criteria:

- Rate states as low priority if they are implementing only one of the program models, since the research design calls for diversity in models within a state.
- Rate states as low priority if they are supporting fewer than five eligible sites.
- Give a state higher priority if there is specific mention of intent to serve military families, because the authorizing legislation lists military families as a priority subgroup.
- Only count as eligible those sites offering one of the four of the evidence-based models, as only these four are being implemented in at least ten states (Early Head Start, Healthy Families America, Nurse Family Partnership, Parents as Teachers).

The study team plans to contact these 30 high priority states to gather information on the key characteristics of each MIECHV supported program site to determine their fit for the evaluation.

Using criteria based around geography, urbanicity, program model, and operating experience, the study team will select approximately 18 states for further consideration. The study team will then make in-person visits to these states and their local program sites to learn more about the characteristics of their home visiting programs. Follow-up visits will be made to 12-15 of the most promising states to gather additional information about their feasibility for participation in the study.

From that information, a list of potential local programs will be compiled. Eligible local programs will meet several criteria: (1) having two or more years experience with one of the four evidence-based home visiting service models that were selected by at least 10 states receiving MIECHV funds, (2) excess demand for their services so that they can provide enough families for a control group, (3) the ability to enroll 30 families in their program over a period of about a year, and (4) locations where there are few other home visiting services in order to ensure a strong service differential between the program and control groups.

States will be classified in terms of which of four clusters of ACF/HRSA regions the state is in, the number of local sites that appear to be eligible for the evaluation, and the diversity within and across the states in terms of program models, urbanicity, and demographics. Once this information is compiled, the study team will choose states so they meet the following criteria: each of four clusters of regions will be represented, the four evidence-based models are represented roughly evenly across the sites, and sites represent both urban and rural locations. Once 12 states are chosen, 85 sites will be chosen from within those states to meet the same criteria (for example, having the four evidence-based models represented roughly evenly across sites).

Within a site, the evaluation will enroll families where the mother is pregnant or where the family has a child under six months old. Home visiting programs will identify families who appear to be eligible for the study and a field staff person from the research team will go to the family's home to explain the study and obtain informed consent. Families will continue to be recruited until 60 families have been enrolled into the study.

B.2. Procedures for collection of information

All selected states and local program sites will be asked to participate in the telephone and in-person meetings and respond to the study team's questions. Sampling cannot be used because it does not fulfill the purpose of site recruitment, which is to collect information about the remaining universe of states and determine which ones best meet the needs of the study.

Site liaison teams, composed of one senior and one junior staff from the study team, will be assigned to states to make the telephone and in-person meeting contacts. These staff members are experienced in the process of site recruitment for large-scale studies such as MIHOPE.

The remainder of this section describes the study team's procedures for contacting states and local programs in order to make the selection:

Introduce the evaluation to state administering agencies (*January 2012*).

Regional project officers from HRSA will send an email to the state administrators overseeing the MIECHV programming to introduce the study and its goals, introduce the team that will be doing the study on HHS's behalf, and alert state administrators that a study team member may be in contact to explore whether their state would be a good fit for the evaluation.

Telephone contact with state administrators to gather information (January-April 2012).

After the regional project officers from HRSA have sent the initial email, the site liaison team will call state administrators to schedule a longer telephone appointment to collect the minimum information necessary that allows us to select those that best meet our selection criteria and proceed to the next stage of site recruitment. The site liaison team will confirm the appointment by email and attach a list of the information to be collected during the phone call. The appointment confirmation will include several attachments: (1) a project description, which explains the study, the process for selection and enrollment, the project timeline; (2) a set of frequently asked questions, which responds to potential questions state administrators may have about the study; (3) a site participation overview, to provide states with an understanding of what participation in the study would entail for their local home visiting programs and the process for their involvement; and (4) the information that will be discussed during the telephone appointment.

Using a protocol, site liaison team will initiate the longer telephone appointment to answer any questions the state administrator might have regarding the study and ask for a few key characteristics of each MIECHV supported program site. This will enable us to understand the number of local MIECHV programs, using the study's definition of a local program.¹ The information collected will also help the study team classify these sites according to three main characteristics: geographic region, program model, and urbanicity. The study team will use this information, to select approximately 18 high priority states that best suit the evaluation needs.

In-person visits and teleconferences to key states and sites for detailed discussion about the evaluation (March-December 2012).

To recruit and reach agreement with approximately 12 states and 85 local program sites from among the high priority states, the study team plans to visit a state three times. Site liaison teams will meet in-person and by phone to discuss the evaluation with state and local program site staff. These visits and telephone calls will be used to collect information needed to determine which pool of states and sites best meet the criteria for site selection. After each visit, the study team will narrow the pool of eligible states and sites based on the information collected. This could mean first round visits to 18 states, follow-up visits to 12-15, and teleconferences and visits with roughly 120 sites to ensure that we will have 85 from which to choose. Visits to the states will also be opportunities to meet with some prospective sites and introduce them to the study (using a PowerPoint presentation). Using semi-structured protocols, conversations with state staff will be used to gain an understanding of the processes for accessing state administrative records, and to underscore the state administrators' importance in helping to recruit sites. Important topics in discussions with program sites concern their administrative

¹At this time, we define a site as a home visiting program with local administration (separate office and supervision), but the study team will use these conversations to try and understand how the definition may vary across states.

structures, their programmatic experience, when they plan to begin MIECHV services, the community service context, and their program size. Initial visits may include groups of sites, but the study team would eventually meet with each site individually (although not always in person) to understand their program flow, respond to questions and concerns, and discuss the terms of an agreement. The average state will contribute seven sites to the evaluation, but the actual numbers may vary from as few as five to as many as 12 sites in a state.

B.3. Maximizing response rates

Maximum response rate is critical to ensuring that the study team selects the most appropriate states and sites for the evaluation. As a condition of receipt of the MIECHV funds, states had to provide assurances that, if asked, they would participate in the legislatively-mandated evaluation. Therefore, the response rate is expected to be 100 percent for telephone and in-person meetings with state representatives and administrators. The response rate from the program sites is expected to be close to 100 percent but it is possible that a site may avoid contact attempts from the research team because they are unwilling to participate in the study or because program services at that site are terminated after they have been chosen for data collection. High response rates should be achievable with minimal leveraging of the funding requirement by having the study team develop strong relationships with the states and sites.

Regional project officers from HRSA, who are already familiar to the states, will introduce the evaluation and the study team to state administrators via email. From that point forward, the state will be contacted by their assigned site liaison team, composed of one senior and one junior staff member from the study team, for the telephone and in-person meetings. Each state and program site will be asked to designate an individual who will be the site liaison team's primary contact. The development of a close relationship between the state, program sites, and site liaison team will ensure that responses from states and program sites are timely and thorough and that any questions or concerns that may delay or prevent a response are quickly addressed by the site liaison team.

The senior site liaisons have had significant experience in working closely with state and local agencies on previous evaluations. In addition, all site liaisons will receive a training to ensure that states and program sites are engaged in a consistent manner.

B.4. Pre-testing

The data collection instruments will not be pre-tested. Previous large-scale evaluations such as the Department of Labor's YouthBuild Evaluation and OPRE's Head Start CARES Project have used nearly identical instruments in the site recruitment process with success.

B.5. Consultants on statistical aspects of the design

There are no consultants on the statistical aspects of this design. We have drawn on the expertise of Charles Michalopoulos and Howard Bloom of MDRC in designing the study to include 85 sites and 12 states.

