

**SUPPORTING STATEMENT B**  
**FAMILY SELF SUFFICIENCY EVALUATION**  
**OMB 2528-0296**

**B. COLLECTION OF INFORMATION USING STATISTICAL METHODS**

**B1. Respondent Universe, Sampling Selection, and Expected Response Rates**

The HUD-funded Family Self-Sufficiency (FSS) evaluation has enrolled 2,551 households across 18 public housing authorities (PHAs). The 36-month follow-up survey sample includes all eligible households participating in the FSS study.

The expected response rate to the 36-month follow-up survey is 80 percent. Historically, MDRC has targeted and achieved an 80 percent response rate to fielded surveys. MDRC will use a variety of strategies to produce this response rate, which are detailed in section B3 below.

**B2. Procedures for Data Collection and Statistical Analysis**

The 36-month follow-up survey will be administered by MDRC's subcontractor M Davis and Company (MDAC) via phone using computer-assisted survey interviews (CATI).

**Statistical Impact Analysis**

Impact analysis will assess the overall and independent effects of the FSS program by comparing the key outcomes of this treatment group to the outcomes of the control group. The study will track both the program and the control groups for a number of years using administrative and survey data to measure outcomes.

The impact analysis will examine the program's effects on a wide range of outcomes. Key clusters of outcomes measured through the 36-month follow-up survey are detailed below (the full survey, and an item-by-item list are included in appendixes in Supporting Statement A).

**Education and Work:** MDRC will use the survey to collect data on employment, earnings, job characteristics, and work search behaviors. Discussions with PHAs have revealed that some programs take a human capital development approach to self-sufficiency and thus emphasize degree, diploma and certification achievement. MDRC will track educational attainment among study participants through survey data.

**Income, assets, finances, and rent burden:** If FSS affects participants' disposable income, it may help them accumulate assets. With survey data, MDRC will assess the effects of the program on household finances and financial behaviors (such as savings, access to credit, and debt reduction). Data on income combined with housing authority and survey data on tenant rent and utilities payments would be used to construct measures of rent burden.

**Material hardship, and family well-being:** Increases in disposable income, may produce reductions in material hardships, including housing-related hardships such as disconnection of phone and utilities, and reductions in food insufficiency. (MDRC observed such effects on

poverty and hardship in its study of NYC’s conditional cash transfer program.) The 36-month follow-up survey includes measures of the frequency of a variety of material hardships.

## **Statistical models**

As noted in Supporting Statement A and detailed in the original OMB submission for the FSS evaluation (OMB control number 2528-0296, Expiration Date 07/31/2016), this evaluation is a randomized control trial. The power of this experimental design comes from the fact that random assignment ensures that the treatment and control groups are alike in all aspects of the distribution of observed and unobserved baseline and pre-baseline characteristics. As a result, any post-baseline differences between the two groups can be interpreted as effects of the intervention.

The estimation strategy for survey-based outcomes is the same as that used for those collected from administrative records. We will use regression adjustment to increase the power of statistical tests that are performed, in which the outcome, such as “employment during Year 1” is regressed on an indicator for program group status and a range of other background characteristics.

The general form of the regression models which will be used to estimate program impacts is as follows:

$$Y_i = \alpha + \beta P_i + \delta X_i + \epsilon_i$$

where

$Y_i$  is the outcome measure for sample member  $i$ ;

$P_i$  equals one for program group members and zero for control group members;

$X_i$  is a set of background characteristics for sample member  $i$ ; and

$\epsilon_i$  is a random error term for sample member  $i$ .

The coefficient  $\beta$  is interpreted as the impact of the program on the outcome. The regression coefficients,  $\delta$ , reflect the influence of background characteristics.

We may vary the functional form and estimation method depending on the scale of measurement of the outcome for which impacts are estimated; for example, continuous outcomes will be estimated using ordinary least squares (OLS) regression. We can use a more complex set of methods depending on the nature of the dependent variable and the type of issues being addressed, such as: logistic regressions for binary outcomes (e.g., employed or not); Poisson regressions for outcomes that take on only a few values (e.g., months of employment); and quantile regressions to examine the distribution of outcomes for continuous outcomes.

The evaluation will examine many outcomes across a number of domains. When multiple outcomes are examined, the probability of finding statistically significant effects increases, even when the intervention has no effect. For example, if 10 outcomes are examined in a study of an

ineffective treatment, it is likely that one of them will be statistically significant at the ten percent level by chance.

While the statistical community has not reached consensus on the appropriate method of correcting for this problem, we propose to address it by being parsimonious in our selection of outcome variables. In particular, we plan to identify a set of “primary” outcomes and subgroups before beginning the impact analysis. All other outcomes and subgroups will be considered “secondary” and will be used to provide context for the primary impact findings or to generate hypotheses about impacts. Schochet (2008) suggests that this strategy is flexible enough to credibly test the key hypotheses about the program, while at the same time allowing the analyst to examine a range of outcomes in a more exploratory manner in order to uncover policy-relevant information.

The main analysis will be conducted over the pooled sample across all PHAs. While sample sizes will not permit PHA-specific impacts, MDRC will examine impacts for certain clusters of programs, such as smaller versus larger programs, programs that have a strong focus on employment, or programs that are operating in strong versus weak local labor markets. The ability to conduct this type of analysis depends on the variation in program features and contexts across participating PHAs, which will be critical to capture in the implementation analysis.

### **Minimum Detectable Effect Size**

MDRC has enrolled 2,551 households, across the 18 PHAs participating in the evaluation, with half assigned to the program group and half assigned to a control group. As noted above, we derive a research sample size of 2,551 households from the full sample. MDRC will work to achieve a survey response rate of at least 80 percent, creating an effective survey sample of approximately 2,000 households. A sample size of 1,000 per research group is large enough to detect policy relevant impacts on outcomes measured through the survey both for the survey sample as well as key subgroups.

It is useful to consider the concept of Minimum Detectable Effects (MDEs) to explore the size of program impacts that are likely to be observed or detected for a set of outcomes and a given sample size. Since these are estimates, the actual MDEs may be smaller or larger than what is shown here. The estimates shown are likely to be conservative, since they assume that baseline variables are not used in the impact model to improve precision. Pre-random assignment values of key outcomes, such as employment and earnings, are likely to be highly predictive of post-random assignment values of the same outcome. In this case, the increased precision brought about by including these variables in the impact model can reduce the MDEs considerably.

Table 1 presents MDEs for the proposed sample size. The first column presents data for the survey sample and the second column presents data for a subgroup within the survey sample, assuming that subgroup makes up half of the larger sample. The first several rows present MDEs for work, education and earnings. For the survey sample, the evaluation could detect effects (increases or decreases) as small as 5.5 percentage points on employment rate at the time of survey. Because sample sizes are more likely to be smaller for subgroups, the MDE for a subgroup is somewhat larger, at 8.0 percentage points. MDEs for earnings are shown in the table

but are harder to predict, given the difficulty of predicting the variance of earnings. The final row presents MDEs in terms of effect sizes (or the impact on a given outcome divided by the standard deviation of that outcome). Effect sizes are a useful way to present and compare impacts on outcomes that are measured in different units, such as family well-being scales. For each of the proposed sample sizes, the effects sizes are typically considered small to moderate in the evaluation literature.

In sum, the proposed sample size is adequate for detecting effects on a range of outcomes that are relatively modest but still but meaningful from a policy standpoint. This pattern holds for the survey sample and for key subgroups.

**The Family Self-Sufficiency Evaluation**  
**Table 1**  
**Minimum Detectable Effects**  
**Research Group i versus Research Group j**

	Survey Sample (1000 per group)	Subgroup (500 per group)
<b><u>Percentage point effects</u></b>		
Employed at time of survey	5.5	8.0
Has any degree or diploma	5.3	7.7
<b><u>Dollar effects</u></b>		
Earnings	1,002	1,417
<b>Effect size</b>	0.11	0.16

Notes: MDEs are calculated based on a two-tailed significance test and assuming an R-squared in the impact model of 0. Average values for employment, education and earnings outcomes are taken from the Opportunity NYC Work Rewards sample. Average values are 44% for employment at time of survey; 63 percent for degree or diploma, and \$6,874 (standard deviation of \$9,500) for earnings. Effect sizes are measured as the impact on a given outcome divided by its standard deviation

### **B3. Maximizing Response Rates and Issues of Nonresponse**

Because of the recognized mobility of low income populations and the need to ensure high and comparable response rates for both the control and program groups, tracking is considered a critical component to ensure the success of the project’s data collection efforts. Tracking efforts will occur in the interim period between an individual’s random assignment and their re-contact for the 36-month follow-up survey approximately three years later. Multiple methods will be employed during the interim period to update sample-member contact information to help ensure response-rate goals are achieved.

MDAC will employ both tracing activities that will help maintain up-to-date participant data, and tracking activities to reestablish lapsed connections. Changes will be carefully documented in a database, tracking the history of changed fields to prevent reversions to out-of-date information and maximizing the amount of information available for future tracking activities. While the tracking activity will focus on the designated head of each household, MDAC may also include other household members in tracking activities if they are also enrolled in FSS, or the sample member used the other family member as a point of contact if the sample member cannot be located.

The fundamental components of tracking include:

1. Development of a tracking database capable of integrating regular updates of information from various sources including housing authority records (extracted semi-annually), federal Housing and Urban Development records (approximately every quarter) as well as MDAC tracking efforts,
2. Passive tracking using services such as LexusNexis' SmartLinx database,
3. A welcome packet (with a trinket such as a magnet to serve as a reminder of the study and benefit of updating contact information) and an address correction request,
4. A telephone-based tracking survey to confirm full contact information, and
5. An additional tracking letter midway between the tracking survey and the 36-month follow-up survey to ensure accurate contact information and
6. A survey pre-notification letter.

MDRC will also utilize incentive payments as detailed in Supporting Statement A. Payment (\$30) upon survey completion is intended as a token of appreciation. As documented in the literature, this token of appreciation is likely to improve response rates by decreasing the number of refusals, enhancing respondent retention, and providing a gesture of goodwill to acknowledge respondent burdens. This technique is proposed in addition to many of the techniques suggested by OMB to improve response rates that have been incorporated into our data collection effort because our experience has shown that small monetary amounts are useful when fielding data collection instruments with hard-to-employ populations as part of a complex study design. In a seminal meta-analysis, Singer, et al. (1999) found that incentives in face-to-face and telephone surveys were effective at increasing response rates, with a one dollar increase in incentive resulting in approximately a one-third of a percentage point increase in response rate, on average. They found some evidence that incentives were useful in boosting response rates among underrepresented demographic groups, such as low-income and non-white individuals.<sup>1</sup> This is a significant consideration for this study. Another important consideration is the burden posed by this data collection, which will take on average 45 minutes of the participant's time for the follow-up survey.

---

<sup>1</sup> Berlin, M., L. Mohadjer and J. Waksberg (1992). An experiment in monetary incentives. *Proceedings of the Survey Research Section of the American Statistical Association*, 393-398; de Heer, W. and E. de Leeuw. "Trends in household survey non-response: A longitudinal and international comparison." In *Survey Non-response*, edited by R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little. New York: John Wiley, 2002, pp.41-54; Singer, E. and Kulka, R. *Studies of Welfare Populations: Data Collection and Research Issues*, Panel on Data and Methods for Measuring the Effects of Changes in Social Welfare Programs. Ploeg, Robert A. Moffitt, and Constance F. Citro, Editors. National Academies Press, Washington, DC, 2000, pp. 105-128.

#### **B4. Pre-Testing**

MDRC will conduct pre-testing of the survey instrument before fielding. The survey will be thoroughly tested prior to initial pretest calls, including:

1. Screen reviews by both MDAC and MDRC to prevent the release of text with typos and to ensure proper flow of the questionnaire,
2. Entry and review of up to a dozen pre-specified scenarios to ensure output responses align with what was input, and
3. Pending the capabilities of the CATI software used, generation of random responses for at least 1,000 data records to be reviewed to ensure all skip patterns operate as anticipated.

The pretest of the 36-month follow-up survey will be conducted over a 1-week period using a specially trained group of interviewers. MDAC staff will closely monitor each pretest interview to determine whether any substantial changes are needed to the questionnaire design. MDAC also will conduct an interviewer debriefing after the pretest interviews are completed to discuss the flow of the interview and any questions that came up.

During the pretest MDAC will track the minimum, maximum, and average time to complete the interview as well as the median times per section. Following the completion of the pretest, MDAC will prepare a Pretest Report to indicate these tracked timings of interview length and provide recommendations for changes to the questionnaire and data collection procedures.

#### **B5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data**

The information for the FSS studies is being collected by MDRC and its subcontractors, Branch Associates and MDAC on behalf of the Department of Housing and Urban Development. With HUD oversight, MDRC and its subcontractors, including Ingrid Gould-Ellen and John Goering, both national experts, were responsible for developing the study documents included in this submission. The statistical aspects of the study were developed in consultation with MDRC senior economist and impact analyst, Cynthia Miller.