Supporting Statement for Paperwork Reduction Act Submissions
Family Self Sufficiency (FSS) Evaluation – Long-Term Follow-Up Survey
OMB Control # 2528-XXXX

B. Collections of Information Employing Statistical Methods

1. **Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection methods to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.**

The U.S. Department of Housing and Urban Development (HUD)-funded Family Self-Sufficiency (FSS) evaluation has enrolled 2,551 households across 18 public housing authorities (PHAs). A random sample of about 1,300 FSS sample members will be targeted for the long-term follow-up survey. The sample will be evenly split between control and program group members. It is expected that the program group sample will include a mix of active FSS participants, program graduates (the successful exits), as well as those that were terminated from FSS, voluntarily exited the FSS program, or are no longer receiving Housing Choice Voucher (HCV).

The expected response rate to the long-term follow-up survey is 60 to 70 percent. Historically, MDRC has targeted and achieved a 60 to 70 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.

2. **Describe the procedures for the collection of information including:**

- **Statistical methodology for stratification and sample selection,**
- **Estimation procedure,**
- **Degree of accuracy needed for the purpose described in the justification,**
- **Unusual problems requiring specialized sampling procedures, and**
- **Any use of periodic (less frequent than annual) data collection cycles to reduce burden.**

The long-term follow-up survey will be coordinated by MDRC's subcontractor M Davis and Company (MDAC) via a self-administered web-based survey and a computer-assisted telephone interview (CATI) for the non-respondents. So far, MDAC has administered all the follow-up surveys for the evaluation.

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 measures outcomes.

The impact analysis will examine the program's effects on a wide range of outcomes. Key clusters of outcomes measured through the long-term follow-up survey are detailed below (the full survey is included in an appendix in Supporting Statement A).

Education and Work: MDRC will use both Unemployment Insurance (UI) wage records, obtained from the National Directory of New Hires database, and 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 FSS long-term follow-up survey data.

Income, debt, expense, and material hardship: If FSS increases participants' disposable income, it may help participants accumulate assets and reduce their material hardships. 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, outcomes which several FSS programs focus on). MDRC will also evaluate how the program affects 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 New York City conditional cash transfer program, which included a significant housing-assisted population. The longer-term follow-up survey will allow investigation in these potential impacts.

Statistical models

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:

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

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

ε_i is a random error term for sample member i.

The coefficient β is interpreted as the impact of the program on the outcome. The regression coefficients, δ, 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 estimates; 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 identified a set of "primary" outcomes before beginning the impact analysis. All other outcomes and subgroups are 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.

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. For the long-term follow-up survey, MDRC will derive a random sample of 1,300 households from the full sample. MDRC will work to achieve a survey response rate of 60 to 70 percent, creating an effective survey sample of approximately 780 to 910 households. A sample size of about 400 to 450 per research group is large enough to detect policy relevant impacts on outcomes measured through the survey for the survey sample.

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 row present MDEs for employment at the time of the survey. For the long-term follow-up survey sample, the evaluation could detect effects (increases or decreases) of at least 6.8 percentage points on employment rate at the time of survey. The second 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. The effects sizes outlined below are typically considered 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.

	Survey sample (400-450 per group)
Percentage point effects	
Employed at time of survey [a]	6.8
Dollar effects	
Earnings [b]	167
Effect size	0.14

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 are taken from the Opportunity NYC Work Rewards sample. Average values are 44% for employment at time of survey. Effect sizes are measured as the impact on a given outcome divided by its standard deviation.
[a] The value of 6.8 indicates that the FSS program group rate of employment at time of survey would need to be at least 6.8 percentage points above the control group level to be statistically significant 80% of the time.
[b] The value 167 indicates that the FSS program group total monthly household income at the time of survey would need to be at least $167 above the control group level to be statistically significant 90% of the time with a response rate of 70% of the fielded sample.

3. **Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.**

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 long-term follow-up survey. Multiple methods will be employed during the interim period to update sample-member contact information to help ensure response-rate goals are achieved.

We will utilize survey tracking information in order to maximize response rates. MDRC collects contact information semi-annually from participating housing authorities and quarterly from federal HUD database. MDRC is forwarding the information to M Davis and Company (MDAC). Address changes come from mailings to the participants and passively tracking respondents through the U.S. Postal Service Change of Address database. This approach provides an inexpensive method for being able to collect more recent contact information for respondents.

MDAC will also conduct passive tracking using a service such as LexisNexis' Smart Linx in order to maximize response rates and deal with issues of non-response. Passive tracking may begin as early as 3 months prior to data collect and will take place through the end of fielding. Such tracking efforts will help maintain up-to-date participant data and reestablish lapsed connections. Changes to contact information 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.

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.[1] This is a significant consideration for this study. Another important consideration is the burden posed by this data collection, which will take on average 18 to 20 minutes of the participant's time for the follow-up survey.

4. **Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of test may be submitted for approval separately or in combination with the main collection of information.**

Both MDRC and MDAC will conduct pre-testing of the survey instrument – the online and CATI versions – before fielding. The survey will be thoroughly tested prior to the fielding of the survey, 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. These reviews will test different pathways through the instrument, ensuring that questions and response categories appear as intended and that skip patterns are correct, and
2. An operational pre-test of the survey instrument by MDAC to mitigate risk. The test will be a slow start that will collect up to 9 interviews to help ensure that systems are working as intended.

5. **Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.**

The information for the FSS study is being collected by MDRC and its subcontractors, MDAC, on behalf of the Department of Housing and Urban Development. With HUD oversight and input, the MDRC team is responsible for developing the latest follow-up survey for this evaluation, included in this submission. The statistical aspects of the study were developed in consultation with MDRC senior economist and impact analyst, Cynthia Miller (212-340-8693).

---

[1] 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.