

**SUPPORTING STATEMENT FOR  
PAPERWORK REDUCTION ACT 1995 SUBMISSION  
Evaluation of the Reintegration of Ex-Offenders—Adult  
Program (RExO)**

The U.S. Department of Labor (DOL), Employment and Training Administration (ETA) is seeking approval from the Office of Management and Budget (OMB) to collect information from program participants and staff in the evaluation of RExO. This evaluation aims to examine the impact of comprehensive employment-centered services on formerly incarcerated individuals' employment, earnings, and recidivism. The evaluation will rely on a comparison of the outcomes for RExO service recipients with those for eligible individuals who are randomly assigned to the control group and do not receive RExO services. Information will come from two rounds of surveys of participants in the treatment and control groups, which will include questions about relevant respondent characteristics as well as employment, earnings, and offending after random assignment.

RExO began in 2005 as a joint initiative of DOL, the Department of Justice (DOJ), and several other federal agencies. The purpose of the program is to provide employment-centered services as well as case management, mentoring and a range of other supportive services to nonviolent offenders who are newly released from prison. RExO grantee programs follow a three-stage reentry framework that begins with pre-release services, progresses through structured community-based reentry programming, and culminates in community reintegration with a reduced need for program services. A typical grantee program participant receives services for about three months with continued follow-up of up to a year. Funding was first awarded to 30 community-based organization grantees in April 2005 and renewed in 2006 and 2007. In 2008, DOL conducted a limited competition for the fourth year of funding, as a result of which 24 of the 30 grantees received awards and agreed to participate in this random assignment study. These 24 programs obtained additional funding in 2009.

In 2009, ETA contracted with Social Policy Research Associates (SPRA), a research, evaluation and technical assistance firm located in Oakland, California, to carry out an impact evaluation of RExO. MDRC and NORC at the University of Chicago are serving as SPRA's subcontractors, with the former involved in the administration of random assignment in the 24 participating sites as well as site visits to learn about the program's implementation, and the latter conducting the survey of study participants.

Between February 2010 and January 2011, sixty percent of eligible clients at the grantee sites were assigned to the program group receiving RExO services, and the rest were assigned to the control group and could receive other services available in their communities. Altogether 4,660 participants were assigned to one of the two groups. The impact evaluation design relies on the comparison of the employment, earnings and recidivism outcomes between these two groups.

## **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 24 RExO grantees enrolled an average of 194 individuals into the study (60 percent into the program group and the rest into the control group). These 4,660 individuals comprise the entire universe of RExO participants under study. The RExO survey will use a sample design wherein the sample is a census. Thus, the survey will be administered to all eligible applicants.

Although the survey will be administered to all participants, and thus the sample is a census, it is possible that those individuals who sought services from the 24 RExO grantees differ in any number of ways than those offenders who did not. This concern is more theoretical than statistical, as the study will not be attempting to generalize to the broader offender population but, rather, to describe the impacts for those who are eligible for and interested in the RExO program. Thus, any conclusions drawn will be valid for this smaller population. Nevertheless, it should be noted that those who are eligible for and expressed interest in the RExO program may differ in meaningful ways from those who did not.<sup>1</sup> Among those who agreed to participate in

---

<sup>1</sup> In addition to individuals selecting whether they are interested in receiving RExO services, it is also theoretically possible that requiring interested applicants to

the study, however, there is no selection bias, because the random assignment process ensures that all participants have an equal probability of being selected into the treatment (or comparison) group, and the selection mechanism is made entirely at random. This is the primary reason why a random assignment design was solicited by DOL and employed in this evaluation.

We expect to obtain a very high response rate for this population of 80 percent for the first round and 70 percent for the second round of the survey, consistent with previous data collection efforts of similar nature and magnitude. This yields a sample of 3,728 for the first round of the survey and a sample of 3,262 for the second round of the survey.

Several prior projects illustrate the research team's ability to successfully survey highly mobile and low-income populations, including specifically ex-offenders, by understanding, anticipating, and responding to changes in respondent circumstances, including: the National Longitudinal Survey of Youth, which after 23 rounds of data collection since 1979 had an 82 percent response rate in 2009; a Multisite Evaluation of Foster Youth Programs with final response rates by site that ranged from 89 to 92 percent for a 24-month interview; and Wave 6 of the Chicago-based Study of Adolescent Health where a 90.6 percent response rate was attained. Additional relevant studies are provided in Section A of this statement.

## **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.**

---

participate in the study in order to have the chance to receive RExO services may have some effect on the population that opts to participate. While it is impossible to know in any statistical sense what this effect might be, it should be noted that this requirement may have had some effect on the study population. Not requiring participation in the study would have likely meant a smaller overall sample size and a resulting sample that was not representative of the overall RExO population.

a) *Sample Selection.* The survey will employ a sample design wherein the sample is a census. Thus, the survey is administered to all members of the universe of eligible RExO applicants assigned to either the treatment or the control group.

b) *Estimation Procedures.* As described in A16 above, the estimation procedures will be conducted as follows.

**Overall Analysis.** The impact analysis will begin with establishing the extent to which the outcomes of RExO program participants differ from those of the control group, which had access to other community services but not to RExO.

Our impact analysis will employ methods that are appropriate and accessible. Because the two randomly assigned groups exhibit similar socioeconomic, demographic and criminal history characteristics and differ only along the dimension of interest (RExO service receipt), we will primarily compare the averages and distributions of the outcome variables between them. Standard statistical tests such as the two-group t-test (for continuous variables) or chi-square tests (for categorical measures and distributions) will be used to determine whether estimated effects are statistically significant at the 1, 5, or 10 percent level (Greene, 1999).<sup>2</sup>

Since we will analyze multiple outcomes, we will explore the possibility of adjusting estimates to account for the multiplicity of hypotheses. One option is to use the Bonferroni correction (Darlington, 1990).<sup>3</sup> This correction, however, is quite conservative in that it makes it rather difficult to reject the null hypothesis and find a significant difference between the groups. Accordingly, we also plan to consider less conservative techniques, including Sidak's correction (which assumes that the various tests are independent of one another),<sup>4</sup> sequential Bonferroni correction methods (such as Holm's or the Simes-Hochberg methods, which eliminate rejected hypotheses from the number of comparisons, thereby increasing the power of the tests), or the false discovery rate, originally discussed by Benjamini and Hochberg (1995).<sup>5</sup>

---

<sup>2</sup> The chi-squared test is derived from:  $\chi^2 = \sum \frac{(O_{er} - E_{er})^2}{E_{er}}$ , while the t-test is derived from:  $t = M_T - M_C / \sqrt{(\text{Var}_T/n_T + \text{Var}_C/n_C)}$

<sup>3</sup> The Bonferroni correction is given by:  $\alpha_B = \frac{\alpha_{FWE}}{C}$ . In simplest terms, this correction multiplies the number of tests by the observed probability of a specific test. Thus, if the probability of a test is .012, but there are ten tests being conducted, the Bonferroni correction would yield a probability level of .12.

<sup>4</sup> In Sidak's correction, the adjusted p-value is equal to  $1 - (1 - \text{unadjusted p-value})^k$ , where k is the number of comparisons being made.

We will use regression adjustment to increase the power of statistical tests, while closely monitoring any implications it may have for impact estimates. Where appropriate, we will explore more sophisticated statistical methods such as discrete choice regression for categorical outcomes (Maddala, 1986); Poisson regression for outcomes that can be counted (Amemiya, 1985); spell analyses (Lancaster, 1990); and panel data methods for outcomes that are measured at several points in time such as quarterly earnings (Hsiao, 1990).

Because we are primarily interested in the average effect of RExO for the 24 grantees that were part of the initial funding for the program (all of which are included in our study) and are not trying to predict what effects would be of some alternative grantee implementing the program, we will include fixed effects for each grant program in our regression specification.

**Variation by subgroup.** We will estimate impacts for key subgroups defined by age, race/ethnicity, gender, and criminal history. We will estimate subgroup impacts in three ways. First, we will use “split-sample” subgroup analyses; under this approach, the sample is divided into mutually exclusive groups, and impacts are separately estimated for each group. In addition to determining whether the intervention had statistically significant effects for each subgroup, Tukey-Kramer’s q-statistics are used to determine whether impacts differ significantly across subgroups (Hedges and Olkin, 1985).<sup>6</sup> A related type of subgroup analysis uses regression methods to see if the effects of the intervention vary significantly with a continuous baseline measure (or one that takes on many values) such as age. Finally, we will employ “conditional” subgroup analyses, which take the regression approach one step further by controlling for the effects of other baseline characteristics when estimating the relationship between a particular subgroup and program effects. For example, in estimating whether the programs have larger effects for older sample members, conditional subgroup analysis controls for gender, type of offense, criminal history, and so on.<sup>7</sup> By estimating the impacts by subgroups

<sup>5</sup> The false discovery rate is given by:  $E[V/(V+S)] = E[V/R]$ , where V is the number of false positives, S is the number of true positives, and R is an observable random variable.

<sup>6</sup> This statistic is given by:  $\bar{d}_T = q_T \sqrt{\frac{MS_{S/A}}{n}}$  in which  $q_T$  is the studentized range statistic,  $MS_{S/A}$  is the mean square error from the overall F-test, and n is the sample size for each group.

<sup>7</sup> In notation, the basic impacts are calculated from a regression of the form  $y_i = \alpha + \beta_1 E_{i1} + \beta_2 E_{i2} + \delta X_i + \varepsilon_i$  where  $y_i$  is the outcome for individual i,  $E_{ij}$  equals one for those assigned to alternative j (j can be 1 or 2) and 0 otherwise, and  $X_i$  is a set of baseline characteristics. The parameter  $\beta_1$  measures the effect on program group 1,  $\beta_2$  measures the effect on program group 2, and  $\beta_1 - \beta_2$  measures the

using multiple approaches, we can ensure that the findings from these analyses are robust under different sets of assumptions that underlie the differing methods.

c) *Degree of Accuracy.* Because sampling will not be employed, results should accurately reflect the relevant universe, subject to the constraints of reporting error and non-response bias. As a result, comparisons of means are sufficient to examine impacts of the program.

Basic characteristics of all participants are known from existing sources (e.g., the RExO management information system (MIS), as well as the random assignment system used to assign participants to either the program or control group). These data will show whether non-respondents differ in any substantial way from respondents and can be used to develop weights for respondents in order to adjust for such differences. Table 4 shows the minimum detectable effects (MDEs) given our expected sample, for several of the key impact measures to be assessed. These measures include whether an individual was: arrested in the period following random assignment; convicted during this period of a crime committed since random assignment; incarcerated since random assignment (for any reason [i.e., parole violation] or for a new crime committed since random assignment); and employed during the period since random assignment. These MDEs are derived using the following equation:  $MDE = \frac{M \cdot \sqrt{Var(\text{impact})}}{sd_{tot}}$  measured in terms of  $sd_{tot}$  where M is a multiplier of the standard error of the estimate,  $Var(\text{impact})$  is the variance of this estimate, and  $sd_{tot}$  is the standard deviation of the outcome measure across all individuals in the population. These MDEs are calculated using pooled within-group variance.

Given that subgroup analysis relies on smaller groups in each subgroup, the MDEs are necessarily increased when conducting subgroup analyses. Table 5 shows the MDEs for the primary subgroup analyses we propose to conduct using actual data for each subgroup drawn from the participants in this study. As can be seen in this table, power is sufficient for each subgroup analysis shown, and would be sufficient for other subgroup analyses provided no subgroup is less than 15 percent of the total sample.

---

difference in effects of the two alternatives. For subgroup analysis with a continuous subgroup measure, the regression would take the form  $y_i = \alpha + \beta_1 E_{i1} + \beta_2 E_{i2} + \gamma_1 Z_i E_{i1} + \gamma_2 Z_i E_{i2} + \delta X_i + \epsilon_i$ . Here,  $\gamma_1$  and  $\gamma_2$  would indicate how impacts vary with the baseline characteristic, and  $Z_i$  is a particular baseline characteristic for which subgroup impacts are being estimated. Conditional subgroup analysis can be represented by the equation  $y_i = \alpha + \beta_1 E_{i1} + \beta_2 E_{i2} + \gamma_1 Z_i E_{i1} + \gamma_2 Z_i E_{i2} + \delta_1 X_i E_{i1} + \delta_2 X_i E_{i2} + \epsilon_i$ .

**TABLE 4**  
**Minimum Detectable Effects Observable for RExO Evaluation**

Total Population Size	4,660	
MDEs Based on What Type of Data	Admin. Records	Survey
Total Sample Size	4,660	3,728
Minimum Detectable Effect Size with 80% Response	.074	.094
Arrested, Year 1	3.2%	3.6% (0.43)
Convicted, Year 1	2.8%	3.2% (0.38)
Incarcerated, Year 1	2.6%	2.9% (0.35)
Incarcerated (New Crime), Year 1	1.3%	1.4% (0.17)
Employed, Year 1	3.7%	4.2% (0.50)

Note: MDE's are calculated based on data from a recent experimental study of returning prisoners (the Center for Employment Opportunities evaluation). Specific standard deviations used for each measure are shown in parentheses. MDEs are calculated based on a power level of 80 percent with a statistical significance threshold of  $p=.05$  (one-tailed test).

**TABLE 5**  
**MINIMUM DETECTABLE EFFECTS FOR SUBGROUP ANALYSIS**

Outcome	Total	Gender		Race			Age		HS Completion	
		M	F	Black	White	Other	<35	35+	No	Yes
Proportion of Sample	1.0	.81	.19	.51	.33	.16	.47	.53	.46	.54
Arrested, Year 1	3.2	3.5	7.3	4.4	5.5	7.9	4.6	4.4	4.7	4.3

Convicted, Year 1	2.8	3.1	6.5	4.0	4.9	7.1	4.1	3.9	4.2	3.8
Incarcerated, Year 1	2.6	2.9	5.9	3.6	4.5	6.5	3.8	3.5	3.8	3.5
Incarcerated (New Crime), Year 1	1.3	1.4	2.9	1.8	2.2	3.2	1.9	1.7	1.9	1.7
Employed, Year 1	3.7	4.1	8.5	5.2	6.5	9.3	5.4	5.1	5.5	5.1

Note: Sample proportions for each subgroup are drawn from the actual results from the RExO evaluation. These MDEs are shown assuming the use of administrative records, so that 100 percent of the sample is available for each analysis.

TABLE 6  
SAMPLE SIZE BY SITE

<b>Location</b>	<b>Total</b>	<b>Expected Sample Size with 80% Response Rate</b>
Baltimore, MD	201	161
Baton Rouge, LA	185	148
Boston, MA	183	146
Chicago, IL	113	90
Cincinnati, OH	209	167
Dallas, TX	204	163
Denver, CO	217	174
Des Moines, IA	200	160
Egg Harbor, NJ	200	160
Fort Lauderdale, FL	191	153
Fresno, CA	200	160
Hartford, CT	179	143
Kansas City, MO	149	119
New Orleans, LA	202	162



TABLE 6  
SAMPLE SIZE BY SITE

<b>Location</b>	<b>Total</b>	<b>Expected Sample Size with 80% Response Rate</b>
Philadelphia, PA	260	208
Phoenix, AZ	200	160
Pontiac, MI	144	115
Portland, OR	204	163
Sacramento, CA	209	167
San Antonio, TX	204	163
San Diego, CA	205	164
Seattle, WA	197	158
St. Louis, MO	199	159
Tucson, AZ	210	168
	<b>4,660</b>	<b>3,728</b>

d) *Unusual problems.* There are none.

e) *Periodic Data Collection.* The survey will be administered twice to gauge the evolution in participant outcomes between the 12- and 36-month marks after random assignment.

**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.**

**A. *Methods for Maximizing Response Rates***

Several strategies will be used to achieve a high response rate to the surveys. First, when participants were randomly assigned and entered

the study, they were notified that they would be asked to complete a survey approximately one year later.

Second, about a week before the sample is released to the computer-assisted telephone interviewing (CATI) call scheduler, a letter will be mailed to prospective respondents to remind them of the purpose and sponsorship of the survey (see Appendix E). The letter will also request up-to-date contact information and provide a toll-free call-in number.

Third, staff members from the contractor's experienced pool of interviewers will be recruited and trained. They will receive thorough instruction on data collection procedures, including techniques to achieve respondent cooperation. Interviewers especially skilled at encouraging cooperation will be available to persuade reluctant respondents to participate and will be assigned to respondents who initially refuse (except for hostile refusals). Bilingual interviewers will also be available for conducting interviews in Spanish.

Fourth, call scheduling will allow respondents to select the time that is most convenient for them to be interviewed. We plan to conduct this survey using CATI, which ensures control of sample releases, call scheduling, and questionnaire logic and completeness.

Fifth, the subcontractor will make extensive use of various online databases, including Accurint, to try to locate sample members who have moved. In addition, we will attempt interviews with both respondents and nonrespondents to previous interviews, because this approach can only increase interview response rates.

We expect that these techniques, combined with the \$50 monetary incentive (and an additional \$25 "early-bird" incentive), will yield an 80 percent response rate to the first and 70 percent to the second round of interviews, as detailed in Part A of this submission.

## ***B. Addressing Nonresponse***

Upon the completion of each wave of survey administration, we will conduct an analysis, as noted in our response to B.2.c above, to assess whether the survey sample is representative of the initial population of RExO study participants. This analysis will be done using MIS and random assignment system data,<sup>8</sup> which will be available for all study participants, whether or not they respond to the survey. These data will

---

<sup>8</sup> The MIS is a system used by all RexO grantees and maintained by DOL. The random assignment system was developed for the RExO study and is maintained by MDRC, a subcontractor for the evaluation.

include demographic variables, release status (probation, parole, or no supervision), and other personal characteristics (e.g., education and housing status at entry).<sup>9</sup>

In particular, for both waves of the survey, we will conduct statistical tests (chi-squared and t-tests)<sup>10</sup> to gauge whether treatment and control group members who respond to the interviews are representative of their groups. Noticeable differences in the characteristics of survey respondents and nonrespondents may indicate the presence of nonresponse bias. If there is any evidence of nonresponse bias, we will test whether the baseline characteristics of *respondents* in the two research groups differ from each other. Although it is likely that there will be a lower overall response rate to the 36 month survey than the 12 month survey, simply because greater time will have elapsed which therefore will allow sample members to have moved more frequently and otherwise changed their contact information in a manner that will make locating them more difficult, the analysis of potential nonresponse bias will be similar for each wave.

If any nonresponse bias is observed, we will take several steps to correct it in the estimation of program impacts. First, we will adjust for observed differences between program and control group respondents using regression models. Second, because this regression procedure will not correct for differences between respondents and nonrespondents in each research group, we will construct sample weights so that the weighted observable baseline characteristics of respondents are similar to the baseline characteristics of the universe. For each survey round, we will construct separate weights for program and control group members using the following three steps:

1. Estimate a logit model predicting interview response. The binary variable indicating whether or not a study participant is a respondent will be regressed on baseline measures.
2. Calculate a propensity score for each individual in the universe. This score is the predicted probability that a study participant responds to the survey and will be constructed using the parameter estimates from the logit regression model and the person's baseline characteristics.
3. Construct weights using the propensity scores. Individuals will be ranked by the size of their propensity scores, and divided into several groups of equal size. The weight for a study participant will be inversely proportional to the mean propensity score of the group to which the person is assigned (Hirano and Imbens, 2001).

---

<sup>9</sup> Note that nonresponse differs from attrition, which could occur if individuals drop out of the RExO program before completing it. Attrition of this sort, however, would not preclude the individual from taking part in the survey, and we would still make the same effort to survey those who did not "complete" the program. A second form of attrition—those who ask to have their names withdrawn from the study—is deemed to be trivial in this instance, as only three individuals (out of 4,660) have done so.

<sup>10</sup> The chi-squared test is derived from:  $\chi^2 = \sum \frac{(O_{er} - E_{er})^2}{E_{er}}$ , where  $O_{er}$  equals the observed value, and  $E_{er}$  equals the expected value of an observation, while the t-test is derived from:  $t = \frac{M_R - M_N}{\sqrt{\text{Var}_R/n_R + \text{Var}_N/n_N}}$ , where  $M_R$  and  $M_N$  are the means for respondents and nonrespondents, and  $\text{Var}_R$  and  $n_R$  are the variance and number of respondents, and  $\text{Var}_N$  and  $n_N$  are the comparable values for nonrespondents.

This propensity score procedure will yield larger weights for those with characteristics associated with lower response rates (that is, those with smaller propensity scores). Accordingly, the *weighted* characteristics of respondents should be similar, on average, to the characteristics of the universe, addressing the nonresponse bias.

It is important to note that the use of weights and regression models adjusts only for *observable* differences between survey respondents and nonrespondents in the two research groups, and, using Little and Rubin’s (2002) terminology, for data that are missing “at random.” The models do not account for data that is missing in an unmeasured fashion (or “not at random,” using Little and Rubin’s terminology). Should the analysis of nonresponse bias indicate that the nonrespondents differ in a manner “not at random” we will use multiple imputation methods (Rubin, 1987) as an alternative means of correcting for non-response.

Although in practice we would use a routine such as PROC MI and PROC MIANALYZE in the statistical software package SAS, the procedure’s basic steps are as follows:

1. Estimate a regression of the outcome on any relevant observed information, including any covariates that will be used in the impact analysis regression adjustment, and post-random assignment data obtained through administrative records.<sup>11</sup>
2. Predict the mean value for each individual, conditional on the covariates included in the regression. Using only the mean predicted value, however, results in inconsistent estimates of the variances and therefore incorrect statistical inferences, because it does not account for random error associated with individual values.<sup>12</sup> Thus, to generate the right distribution of the outcome and hence the right statistical tests, the standard error of the regression is used to add an additional error component to each imputed value.
3. Repeat this process multiple times. Adding random noise to each imputed value generates a distribution of estimates across the multiple imputations that, in theory, provides more accurate statistical inferences. Little is gained by creating more than five to ten imputed datasets, however, even with fairly large amounts of missing data (say, up to 30 percent of values missing. Rubin, 1987).

We do not plan to use multiple imputation unless it is clear from the data that nonresponse bias is a concern, and will do so only after attempting to adjust for this bias using the regression techniques described above.

### **C. Reliability of Data Collection**

The draft survey questionnaire for the present study (Appendix B) draws extensively on questionnaires developed for other studies of former offenders, including the Center for Employment Opportunities

---

<sup>11</sup> For more discussions on imputation model specification, see Little and Rubin (2002) and Allison (2002).

<sup>12</sup> Little and Rubin (2002).

study (Redcross et al., 2009) and the Transitional Jobs Reentry Demonstration (Redcross et al., 2010). The questions were designed to be easily understood by respondents. Revisions were made to the draft questionnaire based on an internal review, a review by DOL, and a pretest.

The use of CATI to conduct the survey also helps ensure the reliability of collected information. It controls question branching/skip patterns (reducing item nonresponse due to interviewer error), modifies wording (providing memory aids and probes and personalizing questions), and constructs complex sequences in skip patterns and logic that are not possible to produce or are less accurate in hard-copy surveys. The probes, verifications, and consistency checks are built into the system to standardize procedures. These procedures ensure the reliability of both the data collection methods and the data collected.

Contractor staff will monitor ten percent of each interviewer's work using silent call-monitoring equipment and video monitors that display the interviewers' screens, comparing the interviewer's performance and recording of the data with the information given by the respondent. Because the interviewer will be unaware which of the interviews are being monitored, this will serve to ensure the reliability of the data.

Should ongoing monitoring identify certain interviewers whose data are unreliable (i.e., more than five percent errors in a single survey), they will receive follow-up training to remind them of the proper procedures for data collection. As part of this training, key survey staff will review the questions, probes, and verification checks that are built into the survey and probe to determine what challenges the interviewer is experiencing in properly using the instrument. Those individuals who continue to show reliability deficiencies after this intervention will be subject to more frequent monitoring and may ultimately be dismissed if they do not improve their performance to a satisfactory level.

**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 tests may be submitted for approval separately or in combination with the main collection of information.**

Nine pre-tests of the current survey were conducted with offenders who received services from grantees funded under later rounds of RExO funding. The pre-tests assessed the content and wording of

individual questions, the organization and format of the questionnaire, respondent burden time, and potential sources of response error. The pre-test results were used to modify the questionnaire only slightly, mostly to ensure the correct skip patterns and logic were implemented.

**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.**

KEY TECHNICAL STAFF

Name	Affiliation	Telephone Number
Statistical Consultants		
None		
Individuals Responsible for Data Analysis		
Dr. Andrew Wiegand	SPR	(510) 763-1499
Dr. Ronald D'Amico	SPR	(510) 763-1499
Dan Bloom	MDRC	(212) 340-8611
Cindy Redcross	MDRC	(212) 340-8817
Dr. Chuck Michalopoulos	MDRC	(212) 340-8817
Individuals Responsible for Data Collection		
Dr. Candace Johnson	NORC	(301) 634-9319
Pam Loose	NORC	(301) 634-9319

## References

- Allison, P.D. (2002) *Missing Data*. Thousand Oaks, CA: Sage.
- Hirano, K., and Imbens, G. (2001). "Estimation of Causal Effects using Propensity Score Weighting: An Application to Data on Right Heart Catheterization," *Health Services and Outcomes Research Methodology* 2: 259-278.
- Little, R.J.A. and Rubin, D.B. (2002) *Statistical Analysis with Missing Data*, 2nd edition. New York: Wiley.
- Redcross, C., Bloom, D., Azurdia, G., Zweig, J., and Pindus, N. (2009). *Transitional Jobs for Ex-Prisoners: Implementation, Two-Year Impacts, and Costs of the Center for Employment Opportunities (CEO) Prisoner Reentry Program*. New York: MDRC.
- Redcross, C., Bloom, D., Jacobs, E., Manno, M., Muller-Ravett, S., Seefeldt, K., Yahner, J., Young, Jr., A., and Zweig, J. (2010) *Work After Prison: One-Year Findings from the Transitional Jobs Reentry Demonstration*. New York: MDRC.
- Rubin, D.B. (1987) *Multiple Imputation for Nonresponse in Surveys*. J. Wiley & Sons, New York.