

PART B: COLLECTION OF INFORMATION INVOLVING STATISTICAL METHODS

The U.S. Department of Labor (DOL), Employment and Training Administration (ETA) contracted with Mathematica Policy Research to implement and evaluate the Self-Employment Training (SET) Demonstration. This demonstration is a reemployment program targeted towards dislocated workers, as defined by the Workforce Investment Act (WIA), who are interested in starting or growing a business in their fields of expertise.¹ The demonstration will rely on self-employment advisors who will deliver intensive business development counseling with the goal of connecting such workers to self-employment training, technical assistance, and other services (including seed capital microgrants) to help them become more successful in self-employment. Enrollment in the demonstration will be open for an 18- to 24-month period² and each program participant will have access to SET services for up to 12 months, for a total implementation period of 30 to 36 months.

The main objective of the evaluation of the SET Demonstration (the SET Evaluation) will be to analyze the effectiveness of the self-employment assistance provided as part of the demonstration and will include two major components:

1. An **implementation study** will (1) describe the experiences of service providers at each of the study sites, (2) document responses to the program of up to 32 individuals participating in the program, and (3) provide a descriptive analysis of the baseline characteristics of study participants using quantitative data from study application forms. The analysis for the implementation study is structured to gain knowledge about the contextual features of the SET demonstration sites; the characteristics of the organizations providing SET services; how the SET program model is implemented; the characteristics of the study population; and,

¹ To receive training services under Title I of WIA, a dislocated worker is an individual who (1) (A) has been terminated or laid off or has received a notice of termination or layoff from employment, and (B) (a) is eligible for or has exhausted unemployment insurance or (b) has demonstrated an appropriate attachment to the workforce, but is not eligible for unemployment insurance, and (C) is unlikely to return to a previous industry or occupation; (2) has been terminated or laid off or received notification of termination or layoff from employment as a result of a permanent closure or substantial layoff, or is employed at a facility where the employer has made the general announcement that the facility will close within 180 days; (3) was self-employed but is unemployed as a result of general economic conditions in the community or because of a natural disaster; or (4) is a displaced homemaker who is no longer supported by another family member. Individuals will be considered eligible for the SET Demonstration if they meet any of these four qualifications, irrespective of whether they register for staff-assisted services with a WIA American Job Center.

² Intake into the demonstration will proceed until the demonstration reaches its participation target (3,000 eligible applicants) across participating study sites.

through participant case studies, qualitative information about the experiences of selected program participants.

2. An **impact analysis** will use a rigorous experimental design in which approximately 3,000 applicants to the program in the study sites are randomly assigned to a program group or a control group with equal probability. Members of the program group will have ongoing access to a self-employment advisor, as well as training and other assistance related to their specific self-employment needs, for up to 12 months. Program group members who achieve key program participation milestones (such as completing a business plan) will have the opportunity to apply for seed capital microgrants of up to \$1,000 to help pay for inventory, equipment, licenses, or other business establishment costs. The control group will not have access to SET services during the program period and will be ineligible for the demonstration's microgrants. Both groups will be able seek out and make use of other self-employment services offered by existing community providers, although program group members will have such services partially subsidized through the demonstration. Impacts will be measured 18 months after randomization. Random assignment will enable the evaluation to obtain causal evidence on the effects of SET Demonstration services relative to what might be obtained by members of the target population from existing community providers. The primary outcome measures for the impact analysis will be (1) self-employment status, (2) employment in any kind of job, and (3) total earnings. An exploratory analysis will examine the effect of the program on important intermediate business development outcomes such as gaining access to seed capital, registering a business, and completing a business plan and/or a marketing plan. Additional secondary outcome measures that will be considered include participation in intensive business development counseling services, receipt of training and technical assistance, labor market experiences, and economic well-being. The exploratory analysis will also examine whether impacts vary across participants, based on initial demographic and socioeconomic characteristics, and across sites, based on contextual and programmatic features.

Additional information about the program model and the research questions that will be examined in these two components of the study is included in Part A.

This package requests clearance for five data collection efforts to be conducted as part of the SET Evaluation:

1. **An Application Package.** This package includes: (1) a **consent form** for study enrollment, which will be administered to ensure that applicants are fully informed about the study's goals and expectations; (2) a **dislocated worker screener form**, which will

allow the evaluation team to determine whether applicants qualify for the demonstration under one of the WIA-defined dislocated worker categories; (3) a **background information form**, which will be used to obtain baseline information about applicants' demographic characteristics and previous work experiences; (4) a **business idea form** requesting detailed information about the applicant's proposed business and how it relates to his or her prior work experience; and (5) a **contact information form** requesting information on three relatives or friends who could help locate the applicant for follow-up data collection (only if needed). The application package forms are included as Appendix A.

2. **Program Participation Records.** These data will be collected to better understand the flow of individuals through the SET program and will include selected information from the following forms: (1) **participant tracking forms** that will be used by the program's self-employment advisors to record participant status and assessment at intake, the types and quantities of specific training and assistance received by program group members each month, and the business development milestones reached, (2) a **seed capital request form** that will be used for qualifying SET program participants to request seed capital funds to cover the costs of approved business expenses, and (3) a **service termination form** to be filled out by the provider in the event that a participant exits the program. Appendix B includes copies of the forms that will be used to collect program participation records.
3. **A Follow-Up Survey.** The follow-up survey of study members will be administered 18 months after random assignment to gather information about their economic outcomes (Appendix C).
4. **Site Visit Interviews.** Two rounds of in-person visits to each study site will provide information about implementation of the SET Demonstration program. Interviews will be conducted with SET program staff, workforce staff, and others involved in the SET demonstration program. A master protocol is included as Appendix D.
5. **Case Study Interviews.** Qualitative information about the experiences of selected study participants who were assigned to the program group will be gathered through a series of case study interviews. These interviews will provide an opportunity to explore in greater depth the patterns of service utilization and self-employment trajectories among selected members of the program group. A master protocol is included as Appendix E.

1. Respondent Universe and Sampling

The SET Demonstration will be implemented in up to eight study sites at which recruitment will target dislocated workers likely to meet the study's

eligibility criteria (described below). It is expected that up to 4,000 individuals will apply to the SET program after completing an online orientation session describing the program. The evaluation team will select 3,000 individuals meeting the eligibility criteria—the application process will be closed once this target is reached.³ Successful applicants will be randomly assigned to a program group, which will have access to SET demonstration services, or to a control group, which will not be allowed to receive the services offered through the demonstration but will have access to any other services available in their communities.

Study sites and sample members will be selected based on the factors described in the first two subsections that follow. A third subsection describes the expected sample sizes for the evaluation’s analyses. Figure B.1 summarizes visually the respondent groups affected by the evaluation’s data collection efforts and gives an overview of the associated burden.⁴

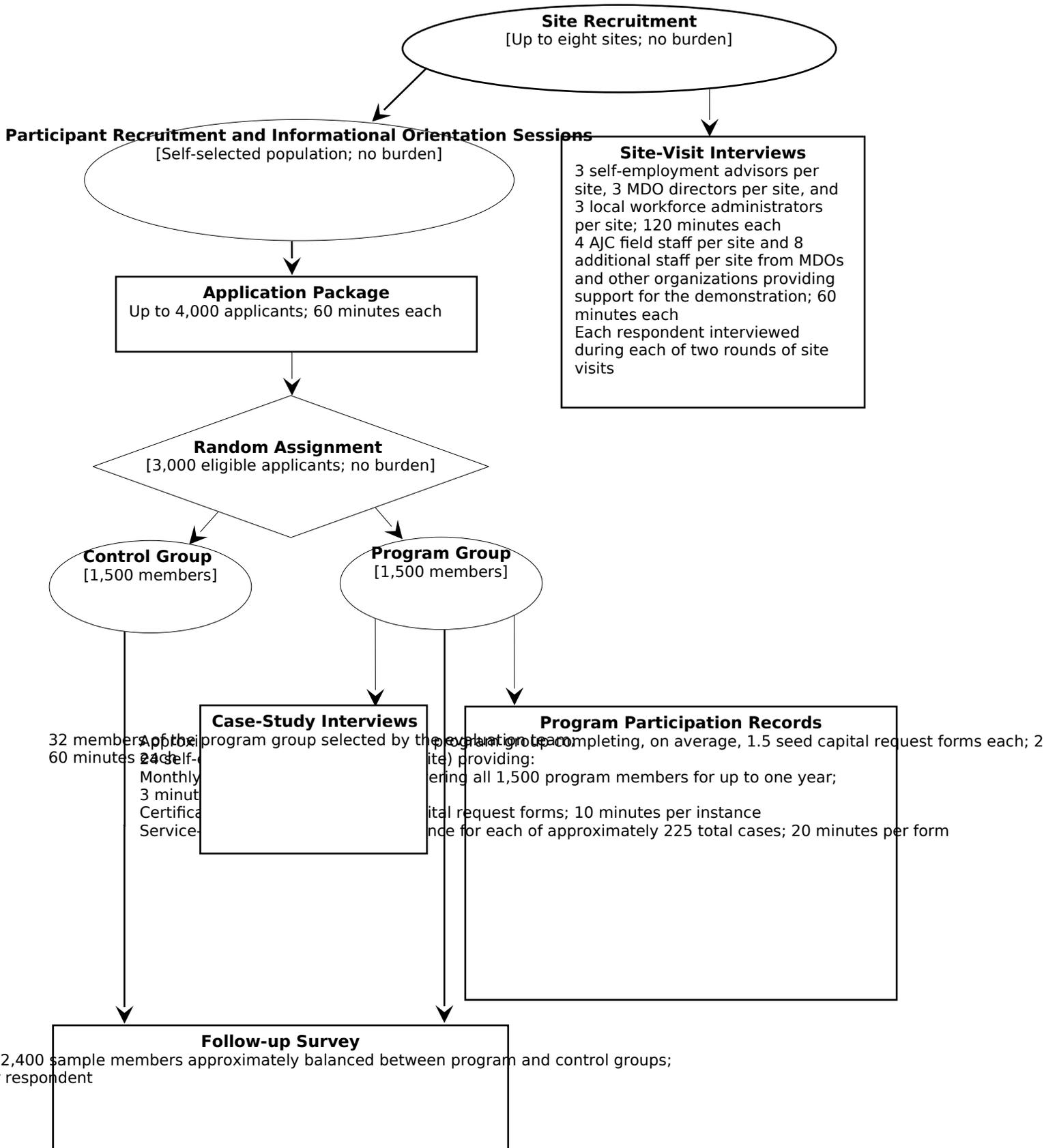
Selecting study sites. ETA is currently working with the contractor to identify states and local sites with sufficient demand to permit the evaluation to meet the study enrollment target and where the capacity of microenterprise providers and the workforce system will allow the demonstration to deliver a strong intervention. The evaluation team is focusing on large metropolitan statistical areas (MSAs) with the following characteristics:

- **High unemployment rates.** To meet the recruitment goal of 3,000 sample members (with an average of 375 potential participants in each site), the study team will concentrate on large MSAs in which the unemployment rates are high.

³ The purposive selection factors described in this section, in conjunction with self-selection of applicants to the demonstration based on an unknown mechanism, mean that the study population cannot be construed as being sampled from a larger target population with well-defined probabilities. As discussed in Section B.1.b, this implies that it will not be possible to draw statistical inference about any larger population than the respondents included in the demonstration.

⁴ Respondent burden is discussed in greater detail in Part A of this package.

Figure B.1. Respondents Affected by SET Demonstration Data Collection Efforts



- **A dislocated worker population with diverse industry experience.** Choosing sites with diverse types of dislocated workers will increase the relevance of the demonstration to other states and reduce the odds that study participants will be faced with excessive competition in launching a small business.
- **A strong network of American Job Centers (AJCs) and workforce system partners.** Identifying states where local sites have strong AJC and workforce system partnerships will enable the demonstration to draw effectively on existing capacity for program marketing, intake, and referrals.
- **A strong presence of MDOs.** MDOs and other similar CBOs are likely to have self-employment advisors who would be familiar with delivering the program's intensive business development counseling services described in Part A of this submission. Consequently, the study team is using information from the Aspen Institute's microTracker database and other sources to identify sites with a strong MDO presence.

To support adequate implementation of the demonstration's program activities, the evaluation contractor will provide modest compensation to MDO and workforce system partners in each study site. Local MDOs that partner with the demonstration to deliver SET services to members of the program group will be compensated for delivering such services, providing information on participant engagement and service receipt for the evaluation, and cooperating with other evaluation data collection activities, such as site visit interviews. Compensation will be provided according to the terms of memoranda of understanding negotiated between these organizations and the evaluation contractor. Partner workforce agencies and state UI agencies will also receive modest compensation to cover the costs of outreach and recruitment activities undertaken in direct support of the SET demonstration, as well as cooperation with evaluation activities. Compensation terms for these organizations will be negotiated and established via formal memoranda of understanding between such agencies and the evaluation contractor.

Selecting the study population. The services offered as part of the SET Demonstration will be concentrated on dislocated workers who, at baseline, already have established behaviors suggesting that they will be responsive to and benefit from self-employment training.⁵ To identify

⁵ The SET Demonstration differs notably in this regard from previous self-employment demonstrations such as the Growing America Through Entrepreneurship project (Project GATE), which enrolled *all* individuals who completed an application. The evaluation of Project GATE, which used a randomized design, found that the program had a small impact on business ownership in the early quarters after program enrollment, but that this effect eroded over time (Benus et al. 2009). The evaluation also found that Project GATE had no significant impact on total earnings at any point during the five years after randomization and no impacts on receipt of Unemployment Insurance (UI) benefits, receipt of public assistance benefits, or household income.

dislocated workers who are likely to benefit from the program, applications to the SET Demonstration will be screened based on prior work experience related to the applicants' proposed business idea. Part A of this clearance request provides additional information on the practical and research-based motivations for selecting potential participants based on related work experience, as well as discussion of options for implementing the screening criteria.

Study recruitment will occur after potential applicants attend a mandatory online orientation session. The orientations will explicitly state the demonstration's eligibility criteria and inform potential applicants that (1) applications not meeting the eligibility criteria will be screened out, and (2) meeting the eligibility criteria qualifies them only a 50 percent chance to enter the SET program, based on the outcome of the random assignment lottery.

Approximately 375 applicants who meet the eligibility screens will be selected for random assignment at each of the eight study sites, on average, yielding a total sample 3,000.⁶ Although eligibility criteria will be explicitly outlined in the pre-application orientation sessions, it is assumed that approximately one in four applications will be screened out. Thus, achieving a total prerandomization study enrollment of 3,000 individuals implies that application packages would need to be obtained for up to 4,000 applicants.

As previously discussed, the 3,000 individuals meeting the eligibility screens will be randomly assigned at each to the program and control groups with equal probability. The program group ($N \approx 1,500$) will be eligible to receive services through the SET demonstration. The control group ($N \approx 1,500$) will not be eligible for such services. Both groups will have access to other existing services available through AJCs and community providers of standard self-employment assistance and training. As noted in Part A, the evaluation team will select partner MDOs that will help support the integrity of the evaluation's control-group design by providing SET services only to the members of the program group for the duration of the program period.

Sample sizes. Individual-level sample sizes vary across the analyses conducted as part of the SET Evaluation. For the implementation study, the evaluation will do the following:

- Site visit data will be collected from all eight study locations to describe the context of the demonstration, the program partner organizations, and the implementation of the SET model.

⁶ Although eligibility criteria will be explicitly outlined in the preapplication orientation sessions, it is assumed that approximately one in four applications will be screened out. Thus, achieving a total prerandomization study enrollment of 3,000 individuals implies that application packages will be obtained for approximately 4,000 applicants.

- Baseline data collected from the application materials will be available for all 3,000 sample members to describe the study population.
- Program participation data from all 1,500 members of the program group will be used to describe participants' experiences with the program.
- Case study data will be collected for a purposively selected sample of program participants (selected based on program participation records for each study site), for a total of 32 case study interviews. These study participants will consist of a mix of those with successful and unsuccessful self-employment outcomes in each site, with success defined as completion of important program participation milestones (such as completing a business plan) and establishing a business or becoming self-employed.⁷

The impact analysis will be based on outcomes data collected from follow-up surveys, which will be initiated with all study members who went through random assignment. Based on the experience of the contractor in fielding surveys for similar study populations, it is expected that the response rate for the follow-up survey will be 80 percent or higher, resulting in a sample size of 2,400 respondents.⁸ This group of individuals will be referred to as the analysis sample. Section B.3.c describes the statistical methods that the study team will use to analyze and potentially account for nonresponse bias by applying sampling weights to the analysis sample.⁹

2. Analysis Methods and Degree of Accuracy

The methods used for the implementation study and impact analysis are presented separately in two following subsections. The main research questions of the each component of the SET Evaluation and the data used to answer them are summarized by Table B.1 and described more fully in Part A of this Office of Management and Budget (OMB) package.

⁷ As noted in Section B.2.a, findings from these purposively selected cases will be used for illustrative purposes only and cannot be generalized to any larger sample of program group members.

⁸ A justification for this expected response rate is presented in Part A of this package (Section A.10).

⁹ Because applicants to the SET Demonstration will not be recruited from a sampling frame with known probabilities (that is, applicants will be self-selected from an unknown population), American Association for Public Opinion Research (AAPOR) guidelines would suggest using the rate of participation, rather than the rate of response, when describing the fraction of the original random assignment sample completing the follow-up survey; the latter term is typically associated with probability sampling (AAPOR 2011). However, the text of this OMB package submission will continue to use response and nonresponse to avoid confusion with participation in the SET Demonstration program by individuals who were randomly assigned to the program group.

Table B.1. Research Questions by Data Source

Research Question	Data Source				
	Application Package	Program Participation Records	Follow-Up Survey	Site Visit Interviews	Case Study Interviews
Implementation Study					
1. What is the context in which the SET Demonstration is implemented?				X	
2. What organizations participate in SET service delivery and what are their responsibilities?				X	
3. What services are offered to SET program group members and what are the other services available to them?		X		X	
4. How well was the program implemented and how did implementation vary across sites?		X		X	
5. What are the characteristics of the SET Demonstration study population?	X				
6. What were the experiences of SET participants with the program?	X	X	X	X	X
Impact Analysis					
1. What is the net impact of the SET Demonstration program on participants' overall employment status and total earnings?			X		
2. Does the SET Demonstration increase self-employment?			X		
3. Does the SET program improve intermediate business development outcomes?		X	X		
4. How does participation in the SET Demonstration affect economic well-being and participation in other programs?	X		X		
5. Do program impacts differ for subgroups of participants defined by baseline characteristics?	X		X		
6. Through what programmatic mechanisms might the SET Demonstration's program influence participant outcomes?			X	X	X

a. Implementation Study

The implementation analysis will include descriptive analyses of (1) the contextual and operational characteristics of the SET program in each study site, (2) detailed case study information about the experiences of selected program participants based on case studies, and (3) the prerandomization characteristics of the pool of applicants enrolled in the study. Each is described next.

Analysis of program operations using site visit data and program participation records. Data collected through site visits and other periodic contacts will be used to describe program operations in each study site, the services offered to SET program group members and how these differ (if at all) from entrepreneurship services generally available in the communities, the partnerships developed to provide SET services (including client flow between the American Job Center system and the MDO providers), and contextual factors that influenced the program's operations. Site visit interview data will be analyzed in a two-stage process.

The first stage—a within-site implementation analysis—will involve preparing summary narratives for each of the demonstration sites. The study team will use these narratives to document the topics noted earlier (and discussed further in Part A). Site visitors will prepare their summary narratives following a common organizational framework, to ensure that all topics are covered in a consistent and comprehensive fashion.

The second stage will draw on this narrative information as raw data to conduct a cross-site analysis that will help inform the impact analysis. Research assistants will use Atlas.TI and a systematic coding scheme prepared and refined by senior staff to code raw data by theme, site, and type of respondent.

Using the coded site visit data together with selected data drawn from the program participation records, the analysis team will conduct a cross-site analysis to describe common elements and differences across sites in the implementation of the SET program. The team will examine variations in services across study sites from both SET providers and other organizations providing self-employment services and characterize the degree to which there is fidelity to the model (high, medium, or low) in each site, using a predeveloped rating scheme. In addition, the site visit data will be used to identify factors or considerations that might help understand why the impacts of the SET program vary from one site to the next.

An important part of ensuring the accuracy of the implementation study's conclusions derived through analysis of site visit data will be ensuring that the data are collected reliably. As described in more detail in Section B.3.d,

strategies to ensure that the data are reliable and as complete as possible include using a flexible approach to schedule visits and ensuring respondents that the information they provide will remain private. Furthermore, using structured, predetermined protocols to collect the data, thoroughly training the site visitors in the use of such protocols in the preparation of systematic summary narratives that cover *all* key topics, and conducting ongoing review of summary narratives by senior staff during the data collection period will help achieve a high degree of accuracy in the data. Because most questions will be asked of more than one respondent during a visit, the analysis will allow for comparisons and triangulation of the data so that discrepancies among different respondents can be interpreted.

Case study analysis of participants' experiences. Based on in-depth telephone interviews with selected SET program participants, this analysis will seek to provide illustrative information on participants' experiences with the program. Using the program participation records submitted by SET service providers in the study site, evaluation staff will select a mix of treatment group members with different patterns of service receipt, overall program participation, and business development progress and invite them to participate in these in-depth interviews. Case study participants will include both individuals who engage strongly with the SET program and reach important participation milestones (such as completing a business plan and/or establishing a business) and individuals who discontinue their participation because they decide not to pursue self-employment any longer and/or focus instead on wage or salary employment. These same factors will be used to select replacements if any of the initially selected participants decline to be interviewed. Although the case study interviews will not be representative of the general population of SET program group members, the findings from this portion of the study could yield new information about correlates of success and failure in self-employment as well as about the flow of participants through the SET program.

The researchers conducting the telephone interviews will prepare case study profiles (one per interview) summarizing each participant's background, original business idea, experiences with the SET demonstration program, and, rich qualitative information about his or her early business start-up experiences and/or reasons for desisting from self-employment.¹⁰ Research staff will code these interviews thematically and by type of respondent using a coding scheme that includes receipt of key types of assistance (for example, mentorship or assistance with finances) and common issues faced in starting a business (such as difficulties obtaining start-up capital).

Coded raw data will be analyzed to identify common themes and catalog their prevalence among successful and unsuccessful SET program participants. Baseline data from the background information forms,

¹⁰ Field researchers will not have to collect background information as part of the interview, because they will have access to the participants' SET program applications.

quantitative information on contact with providers from the program participation records, and data on outcomes from the follow-up survey will be used to round out the descriptive analyses. Vignettes drawn from the case study profiles will help provide illustrations of the identified themes for the evaluation's final report.¹¹

Description of the sample of applicants enrolled in the demonstration. The study team will analyze data from the baseline application form to describe the characteristics of the overall study population. Simple descriptive statistics will provide an overview of participants' demographic and socioeconomic characteristics, their prior work experiences, and the areas in which they seek to pursue self-employment. Additional analyses comparing different subgroups will also be conducted. These will include simple comparisons of selected characteristics of the program and control groups. The significance of between-group differences for each characteristic will be assessed using *t*-tests for continuous measures (including proportions) and chi-square tests for categorical measures. This analysis will serve as an early check that random assignment has been properly conducted and is an extension of the procedure used to monitor the assignment process described in Section B.3.a. A more comprehensive nonresponse analysis will also be conducted later in the study, as described in Section B.3.c.

b. Impact Analysis

The objective of the impact analysis is to provide statistically valid and reliable estimates of the effects of the SET Demonstration on the economic outcomes of the dislocated workers served by the program. A classical experimental design, in which applicants are assigned randomly to program and control groups, will enable the evaluation team to calculate estimates of the causal impact of the SET program. The measured impacts will be internally valid for the eight study sites. However, because the study sites will be chosen purposively and the pool of applicants to the demonstration will be self-selected and then purposively selected as a quota sample, the evaluation's results cannot be generalized to a wider population with a known degree of statistical precision.

A description of the study's outcome measures and discussion of the methods that will be used to estimate the program's impacts and compute variances for the point estimates follows, after which is a description of the expected precision of the estimates by characterizing the minimum detectable impacts (MDIs) of the program that are likely to be obtained using data from the follow-up survey.

Study outcome measures. The primary study outcomes to be examined in the impact analysis include the following:

¹¹ Any vignettes developed will be carefully reviewed to safeguard the participants' identities.

1. Self-employment at the time of the follow-up survey
2. Employment in any job at the time of the survey
3. Total earnings during the one-year period between random assignment and the date of the survey

These outcomes will be used to summarize the effectiveness of the program. Measuring the program's impact on self-employment is an important goal of the demonstration because of the nature of services being delivered. Additionally, self-employment is of particular interest because of the autonomy that self-employed workers are expected to achieve. The other two primary outcomes—employment in any type of job and total earnings—capture the demonstration's *overall* success at helping participants become reemployed, which is the major objective of ETA for the SET Demonstration.

In order to better understand whether and how the SET program works, the evaluation will also consider how effectively it encourages participants to take steps associated with self-employment success. The study will specifically consider intermediate milestones such as whether participants were able to gain access to startup capital, register their businesses, and develop and complete a business and/or marketing plan. Additional, secondary outcomes that will be considered include: receipt of self-employment services; achievement of important intermediate business development milestones; earnings from self-employment and from wage/salary employment; availability of fringe benefits; receipt of unemployment insurance (UI) payments and exhaustion of UI benefits; participation in other government programs; household income, including receipt of transfer payments; and measures of financial distress. (See Part A for further details.) Exploratory analyses of these outcomes will seek to shed light on the mechanisms by which the SET program operates and the diverse set of effects the program might have. Further, as described in following sections, the exploratory analyses will seek to examine how program impacts vary across subgroups. Results from the exploratory analysis will be treated cautiously because of the large number of comparisons being made.

Calculating estimates of program impacts. Random assignment will enable estimation of the net impact of the SET Demonstration by comparing average outcomes across the program and control groups.¹² These estimates will assess the impact of the offer of SET program services, rather than the impact of services received, as some individuals in the program group could chose not to use the business development counseling provided by the demonstration's self-employment advisors. In addition to capturing the direct effects of SET services, the impact estimates also implicitly measure the effects of differences in the quantity and quality of other self-employment

¹² Sections B.2.a, B.3.a, and B.3.c describe the extensive set of analyses that will be conducted to verify that random assignment was properly conducted.

services received, such as classroom training and one-on-one technical assistance, as a result of the SET program.¹³

The core statistical approach for estimating net impacts predicts the outcome of interest as a function of program group membership, site, and a set of background characteristics. The general form of this model for a continuous outcome variable is

$$(1) y_{is} = \beta_s p_i + \gamma \mathbf{x}_i + \alpha \mathbf{q}_{is} + \varepsilon_{is},$$

where y_{is} is the outcome of interest for individual i in site s , p_i is a binary variable indicating membership in the program group, and \mathbf{x}_i is a vector of baseline characteristics of individual i measured before random assignment. The (8×1) vector \mathbf{q}_{is} denotes a set of dummy variables for each study site—for individual i at site s , the s th element of \mathbf{q}_{is} is equal to one and all other elements are equal to zero—and so α represents a set of eight site-specific intercept terms.¹⁴ Finally, ε_{is} is an individual-level random error term that denotes the effects of unobserved factors that influence the outcomes. Because of the randomized design, the error term is expected to have a mean of zero within each site, conditional on the program assignment status of individual i (p_i). The main coefficient of interest in equation (1) is β_s , which measures the average effect of the SET Demonstration program on participants' outcomes at site s . Estimates of program effects using equation (1) are based on the *offer* of demonstration services and are estimated using all sample members in the program and control groups, irrespective of their actual utilization of SET services, in a classical intention-to-treat (ITT) framework.

For ease of exposition, the outcome variable is assumed to be continuous throughout this section. When considering binary outcomes, Equation (1) could be respecified as a nonlinear probit or logit model. However, a regression coefficient from a linear probability model often provides a reasonable approximation to the marginal effect of a variable that would be obtained from a nonlinear binary response model (Wooldridge 2002). Because of its advantages for interpreting point estimates, the linear model will be used if the regression estimates are similar to the marginal effects obtained from the nonlinear model.

¹³ As discussed in Part A, one of the major functions of the self-employment advisor is to help SET participants identify and marshal the most appropriate and effective training resources that are already available in the community.

¹⁴ Sites will be purposively selected based on the criteria described previously, thus statistical inference will be valid for the set of study sites only and cannot be generalized to any broader population. Consequently, site-level intercepts will be specified as fixed effects, rather than random error components.

Point estimates. Equation (1) can be estimated using ordinary least squares (OLS) to obtain the estimated impact of the program at each sites s , $\hat{\beta}_s$ within the analysis sample (that is, the set of individuals that completed surveys). However, the goal of the evaluation is to draw inferences about the effects of the SET Demonstration on the full study population of individuals who were randomized at baseline. As discussed further in Section B.3.c, an analysis will be conducted to assess the extent to which there is the potential for nonresponse bias in the estimates obtained from the analysis sample. In the event that nonresponse adjustments are required, Equation (1) will be estimated using weighted least squares (WLS), with individual-level nonresponse factors used as the elements of a diagonal matrix of regression weights. Equations (4.10) and (4.31) in Cameron and Trivedi (2005) provide the formulas that will be used to calculate the OLS and WLS point estimates, respectively. Irrespective of the estimation technique, estimates of $\hat{\beta}_s$ will be reported separately for each site.

Combining estimates across sites. It is also reasonable to estimate a pooled effect of the program across all sites because each site will be asked to implement the same program model. In addition, one of the key criteria in selecting sites is that the AJC and MDO infrastructure is sufficient to effectively deliver the program. The estimated pooled effect ($\hat{\tau}$) is computed as a weighted average of the estimated effects in each site, where the weights are set equal to the proportion of the sample located in each site. That is,

$$(2) \hat{\tau} = \sum_s f_s \hat{\beta}_s.$$

Without nonresponse adjustments, f_s is equal to the fraction of the analysis sample from site s ; when applying nonresponse adjustments, f_s is equal to the fraction of the baseline sample from site s . Because program assignment within each site will be independent of baseline characteristics, $\hat{\tau}$ will be approximately equal to what would be obtained by estimating a regression in which the impact of the program is constrained to be the same in every site. Thus, the pooled estimate $\hat{\tau}$ can be interpreted as the average effect of the SET Demonstration program across all sample members. Sensitivity analyses will consider whether results differ when sites are weighted equally or are weighted by the inverse of the site-specific variances when calculating the pooled estimate.

Covariates included in the regression. If random assignment has been properly implemented and there are no concerns about nonresponse, it is not strictly necessary to control for baseline characteristics (\mathbf{x}_i) in the regression. However, including these variables in the regression is advantageous because doing so will improve the precision of the estimated

program effects. This occurs because baseline measures that are predictive of the sample members' outcomes will absorb some of the variability in the outcome measures, resulting in a greater signal-to-noise ratio when estimating the impact of the program.

In addition the model described by Equation (1), an alternative approach that will be considered is to allow the relationships between the baseline characteristics and the outcome (that is, the parameters in \mathbf{Y}) to vary across sites. This set-up could potentially improve the precision of the impact estimates, $\hat{\beta}_s$, because the baseline characteristics will be allowed to explain more of the site-specific variation in the outcome. However, this approach implies estimating a substantially larger numbers of parameters, leading to a smaller number of degrees of freedom, which could, all else being equal, reduce the precision of the impact estimates. Thus the net effect on precision of allowing the coefficients on the baseline characteristics to vary across sites is ambiguous. The study team will consider both approaches; the decision about which approach is preferred will be guided, in part, by information on the relationship between survey response rates and baselines characteristics. If that relationship differs substantially across sites (see Section B.3.c), it could be advantageous to allow the coefficients on the baseline characteristics also to vary across sites.

Potential baseline characteristics that could be controlled for include measures of demographics (age, sex, race/ethnicity); family structure (marital status and number of dependents); education level; receipt of UI benefits at the time of random assignment; and baseline measures of employment status and earnings from both self-employment and wage/salary jobs. The specific characteristics included will be selected based on the substantive knowledge of the evaluation team or, alternatively, through a stepwise variable-selection procedure (Neter et al. 1996).

Subgroup analyses. Additional analyses will consider the extent to which the effects of the program differ across different groups of sample members defined by baseline characteristics and whether the effects differ according to specific contextual or programmatic factors measured at the site level. Subgroup impacts will be measured using a straightforward modification to Equation (1).

For ease of exposition, consider the case in which two subgroups of interest are defined by different levels of a binary variable d_i . For example, d_i could be set to one for individuals receiving UI benefits at baseline and to zero for individuals not receiving UI benefits. In this case, subgroup impacts would be estimated using the model

$$(3) \quad y_{is} = \beta p_i + \eta(p_i \times d_i) + \theta d_i + \mathbf{Y}' \mathbf{x}_i + \alpha' \mathbf{q}_{is} + \varepsilon_{is}.$$

Equation (3) differs from Equation (1) in two ways. First, an interaction term between assignment status and the subgroup indicator is included, $p_i \times d_i$. (For clarity, the uninteracted measure of subgroup membership, d_i , has been denoted separately from the other baseline characteristics, \mathbf{x}_i .) Second, the coefficients β and η are not allowed to vary across sites.¹⁵ With this set-up, the average effect of the program on the subgroup for which $d_i = 0$ (for example, individuals not receiving UI benefits at baseline) across all sites is measured by β . The average effect of the program on the subgroup for which $d_i = 1$ (for example, recipients of UI benefits at baseline) across all sites is measured by $\beta + \eta$.

Potential subgroups of interest include those defined by the baseline characteristics controlled for in the regression, as discussed previously. In addition, subgroups can be formed based on different levels of contextual or programmatic factors particular to each site, in which case d_i would be replaced in Equation (3) with a site-level measure of those factors. Another potentially beneficial approach is to focus on UI recipients and form subgroups according to factors associated with their likelihood of exhausting their available benefits. As discussed in Part A, such an analysis might provide useful information to states interested in examining how Worker Profiling and Reemployment Services (WPRS) systems are used to identify candidates for new or existing Self-Employment Assistance programs. The specific subgroups analyzed will be determined by the contractor in conjunction with ETA based on findings from the implementation study, evidence from prior self-employment assistance demonstration projects, and results from other research on the correlates of success in self-employment (for example, Evans and Leighton 1998; Fairlie and Robb 2008).

Variance estimation. Because the SET Demonstration sites will be chosen purposively and the study population will not be sampled probabilistically from a known population, inference will be limited to the baseline sample of individuals who went through random assignment in the eight study sites. Therefore, variances can be straightforwardly estimated using fairly standard linear regression formulas. A Huber-White “sandwich” estimator will be used to account for potential heteroskedasticity of the error term (Huber 1967; White 1980). Asymptotic formulas for heteroskedastic-consistent estimates of the variance-covariance matrix for coefficients calculated using OLS and WLS are given by Equations (4.21) and (4.32) and the surrounding discussion in Cameron and Trivedi (2005). Estimated

¹⁵ This simplifying decision was made because, based on sample sizes, it is not expected that site-specific subgroup differences can be measured with a reasonable degree of precision. Allowing β and η to vary across sites would also imply allowing the basic coefficient on the subgroup indicator, θ , to also vary across sites. This site-interacted specification would further reduce the precision of the subgroup impact estimates through a reduction in the number of degrees of freedom.

variances will be computed based on these formulas using a standard statistical package, such as Stata, that incorporates the scalar “HC1” degrees-of-freedom correction, described by McKinnon and White (1985), as a finite sample adjustment.

When conducting inference on the multisite pooled estimates, which is calculated as a sample-weighted average, the estimated variance of the pooled estimate will take into account the potential correlations among the site-specific estimates. Those correlations are non-zero when the coefficients on the baseline characteristics are constrained to be the same across sites. The variance formula for the pooled estimate given by Equation (2) is

$$(4) \quad \hat{V}(\hat{\tau}) = \sum_s \left[(f_s)^2 \times \hat{V}(\hat{\beta}_s) + \sum_{s \neq s'} \left[f_s f_{s'} \times \hat{C}(\hat{\beta}_s, \hat{\beta}_{s'}) \right] \right],$$

where $\hat{V}(\hat{\tau})$ and $\hat{C}(\hat{\beta}_s, \hat{\beta}_{s'})$ represent estimated variances and covariances and f_s is as defined above.

Minimum detectable impacts. Table B.2 presents MDIs calculated for the three primary study outcomes measured: (1) self-employment at the time of the follow-up survey, (2) employment in any job at the time of the survey, and (3) average quarterly total earnings (from all sources) during the four quarters between random assignment and the survey. The MDIs have been calculated using the following assumptions:

- The level of statistical power is 80 percent and inference will be conducted using a two-tailed test with the significance level set to 5 percent.
- The overall prevalence of self-employment will be 40 percent, the prevalence of employment in any job will be 70 percent, and the standard deviation of quarterly total earnings will be \$9,000.¹⁶
- At baseline, the sample members in each site are assigned with equal probability to the program or control groups.
- The response rate for the follow-up survey is 80 percent in both groups.

¹⁶ Because the pool of applicants to be included in this evaluation is expected to be more focused and more experienced than those in the Project GATE evaluation, the rate of self-employment is assumed to be slightly higher than what was seen in the 18-month follow-up for Project GATE for individuals who were unemployed at baseline. The rate of employment in any job is assumed to be approximately equal to the average of the 6- and 18-month rates for initially unemployed members of the Project GATE sample. Likewise, the standard deviation of total quarterly earnings is based on the average of standard deviation of total earnings since random assignment for the baseline-unemployed sample at the Project GATE 6- and 18-month follow-up surveys; this number is expressed in 2011 dollars. All of these estimated sample statistics from Project GATE are reported in Benus and Michaelides (2010, Tables 4 and 5).

- Baseline measures included in the regression explain 20 percent of the variance in the outcome.
- Point estimates are based on an unweighted regression.
- Variance estimates do not account for heteroskedasticity.

The final two assumptions were made so that an analytic expression for the MDI could be derived. Specifically, using formulas (1) and (5) from Schochet (2008), MDIs are calculated using the approximation:

$$(5) \quad MDI = [T^{-1}(1 - \{\alpha/2\}, df) + T^{-1}(\beta, df)] \times \sqrt{\frac{1}{Np(1-p)}} \times SD$$

In this expression: $T^{-1}(\cdot)$ represents the inverse of the student's t distribution function; α is the significance level for the test, β is the level of statistical power; df is the number of degrees of freedom, which is equal to the number of respondents minus the number of groups minus the number of sites; N is the number of respondents; p is the fraction of respondents assigned to the treatment group; and SD is the standard deviation of the outcome.

In addition to presenting MDIs for the full sample, Table B.2 also displays MDIs for a 50 percent subsample and a 33 percent subsample—which could shed light on the impacts that could be detected in subgroup analyses—as well as MDIs for a single site. Using the full sample obtained from all study sites, the expected MDIs based on these assumptions are 5.0 percentage points for self-employment, 4.7 percentage points for employment in any job, and \$921 for quarterly total earnings. As might be expected, the subgroup and single-site MDIs are higher than the MDIs calculated for the full sample and pooled across all sites.

Table B.2. Minimum Detectable Impacts for Key Outcomes

Sample	Number in Program Group at Baseline	Number in Control Group at Baseline	Outcome Variable (Units)		
			Self-Employment (Percentage Points)	Employment in Any Job (Percentage Points)	Quarterly Total Earnings (\$)
All Sites					
Full Sample	1,500	1,500	5.0	4.7	921
One-Half Subsample	750	750	7.1	6.6	1,303
One-Third Subsample	500	500	8.7	8.1	1,505

Sample	Number in Program Group at Baseline	Number in Control Group at Baseline	Outcome Variable (Units)		
			Self-Employment (Percentage Points)	Employment in Any Job (Percentage Points)	Quarterly Total Earnings (\$)
le					9 7
Single Site					2
Full Sample	188	188	1 4 2	13.3	6 1 3

Note: MDI calculations were calculated using equation (5) based on the following assumptions: (1) the level of power is 80 percent and a two-tailed test will be applied at a 5 percent significance level; (2) at the follow-up survey, the overall prevalence of self-employment will be 40 percent, the prevalence of employment in any job will be 70 percent, and the standard deviation of quarterly total earnings will be \$9,000; (3) individuals at each site are assigned to the program and control groups with equal probability; (4) 80 percent of the individuals in each assignment group complete a follow-up survey; (5) 20 percent of the variance in the outcome is explained by baseline covariates included in the regression; (6) point estimates are based on an unweighted regression; and (7) variance estimates do not account for heteroskedasticity.

To put the MDIs in Table B.2 in perspective, they can be compared with actual impacts found in a randomized evaluation of the Enterprise Project, a demonstration program that provided self-employment assistance to UI recipients in Massachusetts during the early 1990s (Benus et al. 1995).¹⁷ The Enterprise Project increased self-employment by 11 to 17 percentage points during the 21 months after random assignment. Over the same period, Enterprise Project program group members were 11 to 13 percentage points more likely to be employed in any job. The MDIs in Table B.2 indicate that the SET Evaluation could detect such effects even when analyzing a 33 percent subgroup and, possibly, when estimating the impact of the SET Demonstration at a single site. The Enterprise Project also increased total earnings, although this was largely due to increases in wage/salary earnings rather than self-employment earnings. The findings reported in Benus et al. (1995) suggest that the Enterprise Project increased total quarterly earnings

¹⁷ Benus et al. (1995) also evaluated a second demonstration program in Washington State, the Self-Employment and Enterprise Development (SEED) Project, which also provided self-employment assistance to UI recipients. However, the results from SEED evaluation were not considered to benchmark the MDIs calculated for the SET Demonstration for two reasons. First, although the SEED project specified that sample members could “cash out” their remaining UI entitlement, receiving a lump-sum payment after achieving certain benchmarks and business milestones, as described in Part A, the SEED program’s lump-sum payments were: (a) substantially larger than the microgrants offered in the SET demonstration; and (b) only offered to that participants that had already secured adequate financing, which will not be required for participants to access the SET microgrants. Second, the SEED Project was open to all UI recipients. Based on the WIA eligibility criteria noted previously, it is expected that the dislocated workers enrolled in the SET Demonstration will more closely resemble the likely UI exhaustees enrolled in the Enterprise Project.

of the program group by approximately \$1,657 (in 2012 dollars). Based on Table B.2, the SET Evaluation could detect such an impact for the full sample, as well as and the 50 and 33 percent subsamples when calculating pooled estimates across sites. Thus, the SET Evaluation has the potential to statistically detect program effects of a realistic size.¹⁸

3. Methods to Maximize Response Rates and Data Reliability

The methods to maximize response and data reliability are discussed for each data collection effort that is part of this request for clearance in the following subsections.

a. Application Package

The application package will have five components: (1) a consent form for study participation, (2) a dislocated worker screener form, (3) a background information form, (4) a business idea form requesting detailed information about the applicant's proposed business and how it relates to his or her prior work experience, and (5) a contact information form to help locate the applicant for follow-up data collection (if needed).

Response rates. Individuals interested in the SET Demonstration will receive detailed information about the program and associated evaluation from AJC staff and during online orientation sessions. During the orientation sessions, the program's eligibility criteria and study participation requirements (for example, consenting to random assignment) will be clearly and explicitly explained. Staff at AJCs will also be trained on the demonstration's operational procedures and receive both an operational procedures manual for the SET Demonstration and contact information for members of the research team (in case they are asked questions or encounter issues that they are uncertain how to handle). Members of the research team will contact site staff periodically to monitor implementation of the demonstration and provide technical assistance as needed.

At the end of the orientations, potential participants in the SET Demonstration will be offered a hardcopy of the application for their reference and will be given directions to access the secure website hosting the online application (Appendix A). Thus, prospective applicants will have an opportunity to assess their likelihood of qualifying for the program's services and choose whether to complete the application package. Although eligibility criteria will be made clear before applications are distributed, it is assumed that up to one in four applications could be screened out after submitting an application. Thus, achieving a total prerandomization study enrollment of

¹⁸ Applying nonresponse weights would reduce the precision of the SET Evaluation's impact estimates, due to a design effect from unequal weighting. However, as described in Section B.3.c., the contractor conducting the evaluation's follow-up survey will use a variety of proven techniques to maximize response rates for important subgroups.

approximately 3,000 individuals implies that approximately 4,000 applications will be screened.

The application forms are designed to be easy to complete. The forms are written in clear and straightforward language. The time required for customers to complete all three forms is estimated at 45 minutes, on average.

Data reliability. All forms required at intake are unique to the current evaluation and will be used across all SET program sites, ensuring consistency in the use of the forms and in the collected data. The forms have been extensively reviewed by project staff and staff at ETA and will be thoroughly tested in a pretest involving approximately nine individuals from nonparticipating sites who have backgrounds similar to anticipated SET Demonstration participants. The web implementation of the survey will seek to maximize the reliability of the data entered by applicants through skip-pattern logic and checks for consistency and validity.

Ensuring the integrity of the random assignment procedure. The contractor will develop a random assignment system, as described previously, to be implemented for the demonstration. As assignment occurs, the contractor will monitor the process using selected application data that has been transferred into the sample management system (SMS) to ensure that the following four conditions are met:

- 1.** All people who reach the point of random assignment should be randomly assigned.
- 2.** A person can be randomly assigned only once. The validity of demonstration procedures is compromised if people can be randomly assigned again if they do not like their initial assignment, or if they reappear and are not recognized as part of the sample and are randomly assigned again.
- 3.** All people who are randomized to the program group are offered intensive business development counseling through the SET program, and no member of the control group should be offered SET program services or have access to one of the program's self-employment advisors.
- 4.** All individuals assigned to the SET program group should remain identified as members of the program group, regardless of their actual use of services.

The random assignment system will have built-in features that flag possible violations, such as duplicate entries that might result if the same person applies at two sites or reapplies after being assigned to the control group in an attempt to be assigned to the program group. Contractor staff will adjudicate such cases and assign them to the proper research group

(program or control). Staff will also look for other irregularities and seek out missing data.

In addition, the contractor will periodically assess whether the characteristics of individuals randomized to the program and control groups differ by assignment status using data from the SMS. Such characteristics include the same variables used in the impact analysis to form subgroups based on demographics, family structure, receipt of UI benefits, and earnings and employment (in both self-employment and wage/salary employment). Substantial or statistically significant differences (based on t-tests and chi-square tests, as appropriate) in characteristics across subgroups and assignment status could reveal a problem with the implementation of random assignment at local sites that the contractor would seek to address.

Addressing item nonresponse. Although all potential participants in the demonstration are expected to submit complete applications, some item nonresponse on the baseline information form is possible. In such cases, evaluation contractor staff will contact applicants to obtain the missing data when the incomplete application is submitted. Applicants who refuse to provide missing information on characteristics used to determine eligibility for the SET Demonstration and/or monitor random assignment will be considered ineligible for the study. For missing data on other, less essential characteristics, the study team will consider the feasibility of imputing the missing values using, for example, a hot-deck procedure similar to what is used in the Current Population Survey (U.S. Census Bureau 2006).

b. Program Participation Records

The contractor will maximize response rates and data reliability of the program participation records through a combination of three factors. First, the use of advanced technologies (Section A.3) and the use of carefully designed recordkeeping forms (Appendix B) is expected to minimize the burden on staff at MDO partner organizations of transmitting program participation records. Second, organizations will be selected, in part, based on their commitment to evaluating the SET program model and willingness to provide information to assist with this effort. Third, the contractor will carefully monitor the flow of information from MDOs to ensure completeness and accuracy.

c. Follow-Up Survey

The contractor will use well-established methods to maximize response rates and data reliability for the follow-up survey. These methods have been used by the contractor in other data collection efforts, such as the Trade Adjustment Assistance Study Follow-Up Survey (OMB number 1205-0460) and the Individual Training Account 2 (ITA2) Follow-up Questionnaire (OMB 1205-0441). Following a discussion of approaches for maximizing response rates and ensuring data reliability is a description of (1) the methods that will

be used for addressing item non-response on the survey and (2) a detailed description of plans for analyzing and addressing individual-level survey nonresponse.

1) Response Rates for the Follow-Up Survey

The strategy for maximizing response to the SET follow-up survey will be based on the approaches described in following sections. The methods employed will address all types of individual nonresponse, including failure to locate the sample member or his or her refusal to participate in the survey.

Multimodal Administration of the Survey. Based on the pervasive use of the web by a cross-section of the general population, it is anticipated that a substantial number of sample members will choose the web, because many of them are likely to be more comfortable with this self-paced, self-administered approach. It is estimated that 70 percent of the completed surveys will come from the web.

Contact with sample members. The contractor will send an advance letter on DOL letterhead to sample members shortly before the fielding of the survey begins to provide information about the content of the follow-up survey and average administration time, and explain how to access the web-based instrument. This letter will (1) explain the voluntary and private nature of participation, (2) extend the incentive offer, (3) provide web survey log-in information, and (4) give a toll-free number for telephone calls. The contractor will work with partner organizations in the study sites to encourage participation in the survey by sample members. The envelope for hardcopy advance letters will be printed with the DOL logo to capture the sample members' attention and to communicate the legitimacy of the study. Electronic copies of the advance letter will also be mailed to study members who provide an email address at baseline. The contractor's return address will be used to facilitate the processing of returned mail and locating procedures. The advance letter will be followed up with timed reminders offering the option to complete the survey via the telephone or the web. A draft copy of the advance letter that will be sent to sample members is included as Appendix F.

Before the mailing of these materials, interviewing staff, such as interviewers, project supervisors, monitors, and locators at Mathematica's Survey Operations Center (SOC) will be thoroughly trained on how to address respondents' questions about the study and questionnaire. A list of frequently asked questions and answers (FAQs) will be developed for the self-administered web survey, and web survey respondents will have access to them throughout the survey. Other FAQs will be included in the operational procedures manual for the questionnaire administered via computer-assisted telephone interviewing (CATI), and integrated into the CATI instrument. Interviewers will be able to access the FAQs at any time during an interviewer-administered survey.

Locating sample members. A key component to obtaining a high response rate is locating sample members. The process of locating members of the SET study population will begin before sending out the first mailing. This locating process will involve the use of an independent vendor that will check the full sample against current address databases. This first step is critical given that some sample members could have moved since the date at which they submitted their applications. Extensive tracking and locating procedures that have proven successful in other Mathematica studies will be used for sample members whose mail is returned as undeliverable. These include using other independent databases, checking with neighbors and family members, and searching social networking sites. When talking with contacts, the specific purpose of the call will not be disclosed, but it will be stated that the effort to reach the sample member is for an important study being sponsored by the government.

Gaining and maintaining cooperation. A key component to achieving high response rates is gaining cooperation after locating respondents. Mathematica's interviewers are highly trained in establishing rapport with gatekeepers, gaining cooperation, and avoiding refusals. Sample members who are difficult to contact and who have not yet completed the survey on the web will be sent a reminder postcard one week after the advance letter and a follow-up postcard two weeks later. A reminder letter will be sent at the midpoint of the data collection period and again three to four weeks before the end of data collection to remaining nonrespondents. To those sample members who refuse to participate, a targeted refusal-conversion letter that will address their specific concerns will be mailed first. Next, expert refusal-conversion interviewers will make follow-up calls to try to gain the sample members' cooperation.

Multilanguage survey administration. During telephone contact, interviewers will identify Spanish-speaking respondents and connect or schedule them to speak with a bilingual interviewer. When necessary, translators for languages other than Spanish will be used. Mathematica employs staff who speak a wide range of languages and have experience conducting interviews in a number of languages.

Incentives for survey participants. Offering an incentive for the SET follow-up survey could be important for obtaining the desired response rates and reducing overall survey costs. According to Singer et al. (2000), incentives can help to achieve high response rates by increasing the sample members' propensity to respond. By doing so, incentive payments were been found to contain evaluation costs by significantly reducing the number of calls required to resolve a case. Incentives also may increase the likelihood of participation from subgroups with a lower propensity to cooperate with the survey request. This can be an important component of ensuring the representativeness of the survey respondents and the quality of the data being collected. For example, Jäckle and Lynn (2007) found that incentives increased the participation of sample members more likely to be

unemployed. There is also evidence that incentives bolster participation among those with lower interest in the survey topic (Schwartz et al. 2006; Jäckle and Lynn 2007; Kay 2001), resulting in data that are more complete. Furthermore, paying incentives did not impair the quality of the data obtained (such as item nonresponse or the distribution of responses) from groups that would otherwise be underrepresented in the survey (Singer et al. 2000).

Part A of this clearance package provides additional discussion about the potential benefits of incentive payments for response rates and data quality. As discussed there, the evaluation team plans to consider offering incentives to survey respondents, but will test the effectiveness of this incentive for improving response rates through an auxiliary analysis using an experimental design—see Section A.9 for details. To fully assess and leverage the benefits of offering incentives in the SET evaluation’s follow-up survey, the advance letter to study participants receiving incentives will be customized based on the incentive scheme they were selected for and explicitly mention the payment. Such sample members who elect to complete the survey via the telephone will also be reminded of this incentive by the interviewers when contact is first established.

Survey length. The SET follow-up questionnaire is designed to be easy to complete. The questions are written in clear and straightforward language. The average time required for the respondent to complete the survey, either on the web or by telephone, is estimated at 60 minutes.

Interviewer training. Study members opting to complete the survey over the telephone will be interviewed by trained members of Mathematica’s survey operations staff who are experienced working on previous studies conducted for DOL as interviewers, supervisors, and monitors. Most of these staff are familiar with similar questionnaire content and sensitive to the difficulties faced by jobseekers and unemployed individuals, as well as aspiring business owners. All survey operations staff assigned to the study will participate in both general training (if not already trained) and an extensive project-specific training. Interviewers will not work on the study until they have been certified as prepared. The project-specific training will include role-playing with scenarios and other techniques to ensure that interviewers are ready to respond effectively to sample members’ questions. They will also focus on developing skills for securing respondents’ cooperation and averting and converting refusals.

Targeted response rate. Employing these procedures, a response rate of at least 80 percent for the SET follow-up survey is anticipated. When the survey is completed, an analysis that compares response rates in the program and control groups will be conducted to assess whether there are systematic differences between the groups in the likelihood of nonresponse and in the characteristics of individuals responding to the survey. This analysis will use data from the baseline information form, which will be

available for all sample members. These data will include the same variables from the SMS used to monitor the random assignment process, as discussed in Section B.3.a. If it appears that the survey respondent sample is not representative of the study sample, weights to adjust for nonresponse will be developed using propensity scoring methods.

2) Data Reliability for the Follow-Up Survey

The follow-up survey is unique to the current evaluation and will be used across all SET study sites, ensuring consistency in the collected data. The survey has been extensively reviewed by project staff and staff at ETA, and will be thoroughly tested in a pretest involving approximately nine individuals from nonparticipating sites with backgrounds similar to SET Demonstration participants. Potential respondents will be referred to the survey web site by the advance letter and by AJC staff. If a respondent starts the web survey but encounters problems or must complete it at a later time, the survey can be resumed either online or over the telephone with an interviewer. The responses collected by both the web and telephone versions of the survey will be stored in a single database, eliminating the need for merging and related data cleaning. Every aspect of both the web and CATI programs will be thoroughly tested before being put into production. Additionally, to ensure that respondents answer questions, all interview respondents will be ensured of the privacy of their responses to questions.

Addressing item nonresponse. The follow-up survey primarily collects data on outcome measures to be used in the impact analysis. Although the past experience of the contractor conducting surveys for similar evaluations suggests that rates of item nonresponse on the follow-up survey will be very low, some item nonresponse is inevitable. Imputation of outcome data could lead to biased estimates due to imperfect matches on observables when using a hot-deck procedure (Bollinger and Hirsch 2006). Thus, sample members with missing data on a given outcome will be omitted from the sample when analyzing that outcome.

Addressing individual-level nonresponse. As with almost any survey, some nonresponse in the follow-up survey is inevitable. Some sample members will not be located and others will not be able or willing to respond to the survey. The nonresponse analysis will use various data items from the baseline information form, including demographic characteristics, employment status, and earnings. The nonresponse bias analysis will consist of the following steps:

- 1.** Compute response rates for key subgroups. A key subgroup comparison considers the difference between members of the program group and members of the control group and additional subgroups will be formed based on characteristics monitored when conducting random assignment (see Section B.3.a).

2. Compare the distributions of respondents' and nonrespondents' characteristics.
3. Identify the characteristics that best predict nonresponse and use this information to generate nonresponse weights.
4. Compare the distribution of characteristics of respondents using response-adjusted analysis weights with the distribution of characteristics of the baseline sample.

These analyses will be conducted within and across sites to assess whether the potential for nonresponse bias differs among sites. Each of these steps is discussed in greater detail in the following subsections.

Compute response rates for subgroups. The response rate for the subgroups will be computed using the American Association for Public Opinion Research (AAPOR) definition of the participation rate for a nonprobability sample: the number of respondents who have provided a usable response divided by the total number of individuals from whom participation in the survey is requested (AAPOR 2011).¹⁹ Overall response rates will be computed for the full sample and by site. Response rates will then be computed for subgroups defined by characteristics available from the baseline information form to examine if these rates differ systematically from the overall response rate.

Compare the characteristics of respondents and nonrespondents. Next, the characteristics of respondents and nonrespondents will be calculated according to characteristics available from the baseline information form. The statistical significance of the difference between the respondent and nonrespondent subgroups will be assessed using *t*-tests. This type of analysis can be useful in identifying patterns of differences in observable characteristics that might suggest nonresponse bias. However, this approach has low power to detect substantive differences when sample sizes are small, and the large number of statistical tests conducted can also result in high rates of Type I error. Consequently, the results of this item-by-item analysis will be interpreted cautiously.

Identify the best explanatory factors of nonresponse and generate nonresponse weights. Logistic regression modeling is commonly used to develop adjustment weights for nonresponse. This approach is also known as response propensity modeling and can be viewed as an extension of the classical weighting-class nonresponse adjustment

¹⁹ As previously noted, this OMB package submission uses the terms response and nonresponse, rather than participation and nonparticipation, to avoid confusion with "participation in the SET Demonstration program" by individuals who were randomly assigned to the program group. This terminology is not intended in any way to imply that the baseline sample for the SET Evaluation is sampled with known probabilities from a known population. Applicants will be self-selected from an unknown population and the evaluation will seek to draw inference about only the baseline sample of individuals that were randomly assigned.

procedure that makes it possible to include more factors (that is, binary, categorical, and continuous factors) in nonresponse adjustments.

The logistic nonresponse model will be fitted by first identifying a pool of covariates to work from using stepwise regression and then assessing candidate models using various measures of goodness of fit and predictive ability. The covariates will include factors or attributes that can be obtained from the baseline information form and which (1) are likely to be associated with differences in the likelihood that a sample member is located and interviewed and (2) have been shown by previous research (Benus and Michaelides 2010; Fairlie and Robb 2008) to be related to the outcomes of interest for this study among individuals seeking self-employment. Specific examples include demographics (age, sex, race/ethnicity); family structure (marital status or number of dependents); education level; receipt of UI benefits at the time of random assignment, and baseline measures of employment status and earnings from both self-employment and wage/salary jobs. Another important variable to be included in this analysis is the assignment (program or control) status of the individual.

A chi-squared automatic interaction detector (CHAID) will be used to refine the list of candidate independent variables and identify interactions among them.²⁰ The CHAID procedure iteratively segments a data set into mutually exclusive subgroups that share similar characteristics based on their effect on nominal or ordinal dependent variables. It automatically checks all variables in the data set and creates a hierarchy that shows all statistically significant subgroups. The algorithm finds splits in the population, which are as different as possible based on a chi-square statistic. It is a forward stepwise procedure, and it finds the most diverse subgrouping, and then each of these subgroups is further split into more diverse sub-subgroups. Sample size limitations are set to avoid generating cells with small counts. The algorithm stops when splits no longer are significant; that is, the group is homogeneous with respect to variables not yet used or when the cells contain too few cases. The CHAID procedure results in a tree that identifies the set of variables and interactions among the variables that have an association with the propensity of a baseline sample member to complete a follow-up survey.

The variables and interactions identified using CHAID then will be processed using forward and backward stepwise regression to further refine the candidate variables and interaction terms. After identifying a smaller pool of main effects and interactions for potential inclusion in the final model, a set of models will be evaluated to determine the final model.

²⁰ CHAID is normally attributed to Kass (1980) and Biggs et al. (1991), and its application in SPSS is described in Magidson (1993). Decisions about variables and interactions will be based on statistical tests with the significance level (alpha level) set to 0.30. The test size of 0.30 is used instead of the standard 0.05 because the purpose of the model is to improve the estimation of the propensity score and not to identify statistically significant factors related to response.

Computing nonresponse adjustment factors through this process will contribute substantially to the nonresponse bias analysis by identifying the main effects and interaction among main effects that are statistically associated with nonresponse. This information will be used in the bias analysis to form levels of categorical variables for computing response rates and point estimates of program impacts using nonresponse adjustment weights.

Compare the nonresponse-weighted distribution of respondent characteristics with the distribution for the full random assignment sample. In this last step, the weighted distribution of respondent baseline characteristics will be compared with the unweighted distribution of the original study population that went through random assignment. Comparisons will be made for the full study population and for key subgroups, as described earlier in this subsection. This analysis can highlight measures in which the potential for nonresponse bias is greatest and in which greater caution should be exercised in the interpretation of the observed findings.

d. Site Visit Data Collection

The plan to collect study data during site visits will ensure that response rates are high and that the data are reliable.

Response rates. Site visitors will begin working with staff at AJCs and SET partner organizations well in advance of each visit to ensure that the timing of the visit is convenient. The site visits will take place over a period of several months, which also will provide flexibility in timing. Because the visits will involve several interviews and activities each day, there will be flexibility in the scheduling of specific interviews and activities to accommodate the particular needs of respondents. Should scheduling conflicts prevent a meeting with all respondents while on site, follow-up telephone calls will be conducted accordingly.

Data reliability. Five well-proven strategies will be used to ensure the reliability of the data. First, two experienced site visitors will conduct a pilot site visit. During this visit, the site visitors will assess the flow and pacing of the discussion that is guided by the questions in the site visit protocol to ensure that it is feasible during a visit to collect comprehensive information that is in accord with the study's goals. As needed, revisions to the protocol will be made to facilitate the data collection effort. Second, all site visitors, most of whom already have extensive experience with this data collection method, will be thoroughly trained in the issues of importance to this particular study. This training will include techniques to probe for additional details to help interpret responses to interview questions and to ensure all interview respondents of the privacy of their responses to questions. Third, when appropriate, the protocols will use standardized checklists to further ensure that the information is collected systematically. Fourth, site visitors

will be trained in systematic documentation of data gathered on all key topics through the use of a standardized template. Finally, a senior member of the evaluation team will review each site visit report to ensure that the relevant data are collected and recorded. Site visitors will be directed to conduct follow-up with respondents to gather missing information as necessary and submit revised site visit reports for review to senior team members.

e. Case Study Interviews

The study team has devised several strategies to ensure that response rates are high and that the data are reliable.

Response rates. Interviewers will contact selected study participants to explain the purpose of the case study interview and schedule a convenient time for the interview to be conducted by telephone. To ensure high response rates, interviewers will stress the private nature of the interview, the importance of this information for future program improvements, and interviewer flexibility in selecting a time that meets the needs of the respondents. In the event of a refusal, the study team will select new respondents using the same criteria used to purposively select the initial pool of potential respondents.

Data reliability. To ensure high-quality and reliable data collection, the following steps will be taken. First, the interview protocol will be tested and refined by senior staff to ensure that key topics can be covered in the designated time, that questions are clear and unambiguous, and that all key topics are covered. Second, interviewers who are well versed in conducting telephone interviews will be selected and trained in the use of the protocol. Training will focus on ensuring that the interviewers fully understand the interview protocols, are able to adapt the protocol based on existing information available about the respondent, are able to clarify questions and probe for additional details to gather comprehensive information on all topics, and fully understand how to document data from the interviews in a systematic and consistent fashion using a standardized template. Third, before the interview, interviewers will use data from each respondent's application form and from the follow-up survey to assemble a preliminary profile of the respondent that can guide the interview and allow for more time for efficient follow-up on key topics. Fourth, senior members of the evaluation team will participate in initial telephone calls by each interviewer to ensure that they are using the correct interview techniques and following the interview protocol with fidelity. Finally, a senior member of the evaluation team will read each case study report to ensure that the relevant data are collected and recorded. Interviewers will conduct follow-up telephone calls with respondents to collect missing data as necessary and revise case study reports.

4. Tests of Procedures or Methods

All data collection procedures, instruments, and protocols to be used in the conduct of the SET Evaluation will be tested to ensure that the procedures can be feasibly and efficiently carried out, to evaluate the clarity of the questions to be asked, to identify possible modifications to either question wording or question order that could improve the quality of the data, and to estimate respondent burden.

Application package and follow-up survey. The forms contained in the application package and the follow-up survey instrument will be thoroughly tested with up to nine individuals from nonparticipating sites with backgrounds similar to SET Demonstration participants. After each pilot test participant completes the forms, project staff will debrief each participant using a standard debriefing protocol to determine if any words or questions were difficult to understand and answer. Like actual study participants, participants in the pilot test of the follow-up survey will be given an incentive for their time.

Program participation records. The record-keeping, data-entry, and file transmission procedures associated with the program participation data will be reviewed and tested by senior staff at Mathematica before they are deployed. Once sites and partner organizations have been selected, these procedures will be explained to field staff, who will conduct dry runs to test the procedures. Based on early implementation feedback from AJC counselors and self-employment advisors at the SET partner organizations, the procedures for recording and transmitting program participation records will be adjusted, as necessary.

Site visit data collection. To ensure that the site visit protocols provide effective field guides that will yield comprehensive and comparable data across the eight SET study sites, site visit protocols will be based on those used for related evaluations and senior research team members will conduct the first site visit as a pilot test, before launching the full round of site visits. This pilot site visit will help ensure that the protocol will appropriately assist site visitors in delving into the topics of interest and does not omit relevant topics of inquiry. Senior research staff will also assess the site visit agenda—including the data collection activities to be conducted and how these activities are structured—to ensure that they can be feasibly conducted as part of the site visits and yield the desired information. Adjustments to the site visit protocols will be made as necessary based on the results of the first visit.

Case study interviews. The case study protocol will be carefully designed and tested in order to ensure that case study interviews yield high-quality data that provide richer detail compared with data from the application forms and the follow-up surveys. The case study interview protocols will be tested by senior members of the evaluation team in the first

two interviews with study respondents and subsequently refined. Careful attention will be paid to whether key topics of interest to the evaluation are covered by the protocol; whether all these topics can be covered in the designated time; and whether questions and probes are clearly worded, easily understood by respondents, and optimally sequenced in order to solicit responses with sufficient levels of detail. In light of initial interviews, the evaluation team will revise and streamline the interview protocol and the related templates for data recording.

5. Individuals Consulted on Statistical Methods

Consultations on the statistical methods used in this study have been used to ensure the technical soundness of the study. Specifically, ETA has contracted with Mathematica to conduct the SET Evaluation. Table B.3 displays the technical staff who were consulted in planning for the implementation and evaluation of the SET Demonstration.

Table B.3. Contractor Technical Staff

Affiliation and Name	Role on Project	Telephone Number
Mathematica Policy Research		
Dr. Irma Perez-Johnson	Project director	(609) 275-2339
Dr. Heinrich Hock	Task leader, impact analysis	(202) 250-3557
	Task leader, implementation	
Ms. Samia Amin	study	(609) 275-2375
Mr. Shawn Marsh	Survey director	(609) 936-2781
Ms. Annalee N. Kelly	Survey researcher	(609) 275-2885
Ms. Stephanie A. Boraas	Survey researcher	(202) 484-3292
University of California, Santa Cruz		
Dr. Robert Fairlie	Consultant	(831) 459-3332

REFERENCES

- American Association for Public Opinion Research (AAPOR). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. Seventh edition. Lenexa, KS: AAPOR, 2011.
- Benus, Jacob, and Marios Michaelides. "Are Self-Employment Training Programs Effective? Evidence from Project GATE." Unpublished Manuscript. Munich Personal RePEc Archive Paper No. 20883. Available at <http://mpra.ub.uni-muenchen.de/20883>. Accessed May 27, 2011.
- Benus, Jacob, Theodore Shen, Sisi Zhang, Marc Chan, and Benjamin Hansen. "Growing America Through Entrepreneurship: Final Evaluation of Project GATE." Final report submitted to the U.S. Department of Labor, Employment and Training Administration. Columbia, MD: IMPAQ International, LLC, December 2009.

- Benus, Jacob. M., Terry R. Johnson, Michelle Wood, Neelima Grover, and Theodore Shen. "Self-Employment Programs: A New Reemployment Strategy: Final Report on the UI Self-Employment Demonstration." Unemployment Insurance Occasional Paper 95-4. Washington, DC: U.S. Department of Labor, Employment and Training Administration, Unemployment Insurance Service, 1995.
- Biggs, David, Barry de Ville, and Ed Suen. "A Method of Choosing Multiway Partitions for Classification and Decision Trees." *Journal of Applied Statistics*, vol. 18, no. 1, 1991, pp. 49-62.
- Bollinger, Christopher R., and Barry T. Hirsch. "Match Bias in the Earnings Imputations in Current Population Survey: The Case of Imperfect Matching." *Journal of Labor Economics*, vol. 24, no. 3, July 2006, pp. 483-520.
- Cameron, A. Colin, and Pravin K. Trivedi. *Microeconometrics: Methods and Applications*. New York: Cambridge University Press, 2005
- Evans, David S., and Linda S. Leighton. "Some Empirical Aspects of Entrepreneurship." *American Economic Review*, vol. 79, no. 3, June 1989, pp. 519-535.
- Fairlie, Robert W., and Alicia Robb. *Race and Entrepreneurial Success: Black-, Asian-, and White-Owned Businesses in the United States*. Cambridge, MA: MIT Press, 2008.
- Huber, Peter J. "The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions." *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability 1*, edited by L.M. LeCam and J. Neyman. Berkeley, CA: University of California Press, 1967.
- Jäckle, Annette, and Peter Lynn. "Respondent Incentives in a Multi-Mode Panel Survey: Cumulative Effects on Nonresponse and Bias." Working paper presented to the Institute for Social and Economic Research, University of Essex, Colchester, United Kingdom, 2007.
- Kass, G. V. "An Exploratory Technique for Investigating Large Quantities of Categorical Data." *Applied Statistics*, vol. 29, no. 2, 1980, pp. 119-127.
- Kay, Ward R. "The Use of Targeted Incentives to Reluctant Respondents on Response Rates and Data Quality." *Proceedings of the American Association for Public Research*. Montreal, Canada: American Association for Public Opinion Research, 2001.
- Magidson, Jay. *SPSS for Windows CHAID Release 6.0*. Belmont MA: Statistical Innovations, Inc., 1993.

- McKinnon, James, and Halbert White. "Some Heteroskedasticity Consistent Covariance Matrix Estimators with Improved Finite Sample Properties." *Journal of Econometrics*, vol. 29, no. 3, September 1985, pp. 305-325.
- Neter, John, Michael Kutner, Christopher Nachtsheim, and William Wasserman. *Applied Linear Statistical Models*. New York: McGraw-Hill, 1996.
- Schwartz, Lisa K., Lisbeth Goble, and Edward M. English. "Counterbalancing Topic Interest with Cell Quotas and Incentives: Examining Leverage-Saliency Theory in the Context of the Poverty in America Survey." *Proceedings of the American Association for Public Research*. Montreal, Canada: American Association for Public Opinion Research, 2006.
- Shochet, Peter Z. "Statistical Power for Random Assignment Evaluations of Education Programs." *Journal of Educational and Behavioral Statistics*, vol. 33, no. 1, March 2008, pp. 62-87.
- Singer, Eleanor, John Van Hoewyk, and Mary P. Maher. "Experiments with Incentives in Telephone Surveys." *Public Opinion Quarterly*, vol. 64, no. 2, summer 2000, pp. 171-188.
- U.S. Census Bureau. "Current Population Survey: Design and Methodology." Technical Paper 66. Washington, DC: U.S. Census Bureau, October 2006. Available at <http://www.census.gov/prod/2006pubs/tp-66.pdf>. Accessed November 14, 2011.
- White, Halbert. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica*, vol. 48, 1980, pp. 817-830.
- Wooldridge, Jeffrey. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT Press, 2002.

