

---

## Part B: Statistical Methods

Part B of the Supporting Statement for the evaluation of Round 4 TAACCCT grants – sponsored by the Chief Evaluation Office within DOL – considers the issues pertaining to Collections of Information Employing Statistical Methods. Abt Associates Inc. (Abt) is the prime contractor for the study. The evaluation of Round 4 TAACCCT grants will evaluate the national TAACCCT grants program, using a multi-pronged approach including 1) an impact study of selected Round 4 grantees, 2) an implementation analysis and performance assessment, and 3) a synthesis of the independent third-party evaluations that all Round 4 grantees are required to procure.

In Round 4, a total of 71 grants were awarded. For the impact study, the evaluation team plans to select a purposeful sample of up to five grantees. At the selected grantees, applicants will be given randomized access to the local training program created with the TAACCCT grant. Baseline data will be collected on applicants who give informed consent to participate in the study. Administrative and follow-up survey data will be collected to measure the impact of the local training program.

TAA-eligible individuals and veterans who meet DOL’s priority of service criteria will not be randomly assigned, per DOL policy (Training and Employment Guidance Letter (TEGL) 10-09), and will be able to access TAACCCT services. However, DOL has a strong interest in learning as much as possible about services designed to help veterans and TAA-eligible workers and toward that end will request that these individuals consent to baseline data collection and complete the BIF and SAQ.

This submission seeks clearance for three data collection instruments:

- Baseline information form (BIF) (impact study)
- Self-Administered Questionnaire (SAQ) (impact study)
- Semi-structured interview guides for site visits at up to nine Round 4 TAACCCT grantees (impact and implementation studies)

The implementation analysis will be conducted on a slightly larger set of nine grantees. This includes those selected for the impact analysis (up to five), as well as a few others with interesting features, as determined through review of the original grant applications. No statistical methods will be used in the implementation analysis and discussions of the results will be carefully phrased to make clear that no generalization is intended.

Subsequent OMB submissions will seek clearance for follow-up data collection activities.

### B.1 Respondent Universe and Sampling Methods

#### *Baseline Data Collection for Impact Study*

The potential respondent universe for the impact study baseline data collection is the individuals who apply to TAACCCT-funded programs selected for the impact study and who agree to participate in the study. The evaluators estimate that 5,500 individuals will participate in the study across up to five grantees. All individuals who agree to participate in the study will be asked to complete the baseline data collection, including both the BIF and SAQ, as described in Part A. The samples of participants will be self-selected with no formal probability sampling or subsampling. As discussed above, however, there will be two groups of participants subject to different admission rules. For purposes

of this application, the two groups may be considered strata. One stratum consists of all TAA-eligible workers and veterans. If they qualify for the program, the grantee will admit them on a 100-percent basis. All other qualified applicants will have to participate in a lottery in order to gain admission. These sample sizes and response rates are shown in Exhibit B.1. As is discussed below, the grantees will be selected partially on the basis of being able to recruit at least 1,000 individuals. Once a grantee recruits more than 1,000, study enrollment will stop. It is more difficult to project the number of TAA-eligible workers and veterans. The evaluation team estimates no more than 500 will be enrolled in the study, or up to 100 per grantee.

**Exhibit B.1: Sample Sizes and Response Rates by Stratum**

Stratum	In Study	Completed BIF	Completed SAQ	BIF Response Rate	SAQ Response Rate
TAA-eligible workers and veterans	500	400	400	80%	80%
All Others	5000	5000	4995	100%	99.9%
Total	5500	5400	5395		

The response rate for everyone who is neither TAA-eligible nor a veteran will be 100% on the BIF since completing it is a condition of their being eligible for the study. Study participants will also complete the SAQ at the same time as the BIF.<sup>1</sup> Though the SAQ forms are “self-administered,” based on contractor experience from similar procedures, the forms are on rare occasions totally blank. Accordingly, the response rate for the SAQ among randomized subjects is projected to be about 99.9%. For TAA-eligible workers and veterans, both the BIF and the SAQ will be voluntary. We anticipate a response rate of 80 percent.

Work on selecting programs for the impact study is currently underway. It will be a purposive sample. Selected programs will need to meet a variety of conditions that make probability sampling of programs infeasible. Among other conditions, programs selected for the impact study will need to:

- Serve large numbers of participants in a compact set of physical locations over a 12-18 month window,
- Offer services with a strong contrast to the status quo, and
- Offer TAACCT-funded training and services not so widespread that control group members will have no other reasonable training opportunities within local community college systems.

*Site Visits for Implementation Study*

<sup>1</sup> Sample members will have the option of skipping any item except a few key items on the BIF and skipping any item on the SAQ. Completion of the BIF therefore refers to completing the few key items that must be answered. Completion of the SAQ refers to returning the SAQ in the sealed envelope

---

As mentioned above, there will be nine grantees selected for site visits. These nine will include those programs selected for the Impact Study and a few others selected purposefully based on interesting features or strategies proposed in their grant applications. No statistical methods will be used in the selection or inference.

## **B.2 Procedures for Collection of Information**

### **B.2.1 Sample Design**

#### *Baseline Data Collection for Impact Study*

For this evaluation, the evaluation team expects the study sample to include approximately 5,000 individuals who are neither veterans nor TAA-eligible workers but who apply to participate in TAACCCT-funded programs operated by up to five TAACCCT grantees. This section describes how the sample will be recruited. No probability sampling will be conducted.

The Round 4 TAACCCT Solicitation for Grants Announcement indicated TAACCCT grantees were expected to participate in the national evaluation, as well as the impact study (should they be selected). For planning purposes, the team is assuming that five grantees will participate in the impact study. Grantees will be selected purposively for the impact study based on program features, timeline for implementation of key components and the expected number of eligible applicants who could be randomly assigned, as described under B.1.

All eligible individuals who apply for TAACCCT training at the five participating grantees and who give informed consent to participate in the impact study will be randomly assigned to one of two study arms (the treatment or control group). Those applicants who do not consent to study participation will not be allowed to participate in the services supported by the TAACCCT grant but will be able to participate in other services provided in the community.

The BIF and SAQ will be administered after applicants give informed consent but prior to randomization. The BIF will either be self-administered or completed via interview with a staff member while the SAQ will always be self-administered. The study intake period will vary across grantees, but the goal is to have recruitment and randomization up and running at as many sites as possible by October 1, 2015, and at all sites selected for the impact study by the end of December 2015. The study intake period is scheduled to end in all sites no later than May 31, 2017 to ensure that the evaluation can be completed on schedule.

In each the five sites that participate in the impact study, the evaluation team anticipates a baseline sample of 1,000:

- 500 individuals in the no-TAACCCT control group; and
- 500 individuals in the TAACCCT treatment group.

Across all five grantees, evaluators anticipate that a total of 5,000 individuals will complete the baseline surveys and be randomly assigned to one of the two study arms.

Separately, the evaluation team will collect parallel baseline data on all enrollees who are veterans or TAA eligible and therefore exempt from randomization. The evaluation team expects the number of such enrollees to be smaller than 500. The BIF and SAQ are provided in Appendices C.1. and C.2., respectively.

---

### *Site Visits for Implementation Study*

The data will be collected through semi-structured interviews held at selected grantees. The interview protocols are provided in Appendices C.3 – C.7.

#### Statistical methodology for stratification and sample selection

No statistical methods will be used to select the grantee organizations for the site visits as the sample is intended to be neither random nor representative.

#### **B.2.2 Estimation Procedures**

The five grantees selected for the impact study will be analyzed separately. As discussed in the next section, a sample of 1,000 per grantee should provide adequate power to detect such effects as are plausible to anticipate. Although power could theoretically be boosted by pooling across sites, it is likely that pooling would not answer any clear research question. The issue is that innovations in program design under development by TAACCCT grantees are very diverse. Even within grantees, the variety of activities supported by the grants make it difficult to give a short coherent explanation of the program that would be replicable. By keeping analyses separate by grantee, it will be possible to give more structure and meaning to the word “treatment” than simple receipt of a TAACCCT grant.

The core research question for each grantee will be whether individuals enrolled in the TAACCCT-funded program experience more favorable outcomes than they would have experienced if they had not been able to enroll in the program. This is a causal inferential question. As such, standard econometric methods will be used to improve the power beyond what can be achieved with a simple comparison of mean outcomes in the two groups. These involve fitting a weighted least squares (WLS) regression model on the sample, with baseline covariates and treatment status as independent variables, and followup-nonresponse-adjusted weights as in equation 1.

$$\begin{aligned} y_i &= \mu + X_i\beta + \delta T_i + e_i \\ e_i &\sim N(0, \sigma^2) \end{aligned} \tag{1}$$

where  $y_i$  is an outcome such as earnings in the fifth quarter following random assignment,  $X_i$  is a row vector of baseline covariates,  $T_i$  is a binary indicator for treatment status,  $\mu, \beta, \sigma^2$  are unknown nuisance parameters to be estimates,  $e_i$  is a random residual error, and  $\delta$  is the effect of being allowed to enroll in the program. As discussed by Lumley et al (2002) and Judkins and Porter (2013), these methods give valid answers even if the errors are not normally distributed. The SAS procedure SurveyReg will be used to estimate the parameters in equation 1 and to estimate their standard errors. SurveyReg allows the specification of survey weights. In this case, the weights will represent inverse probabilities of response to the followup survey for survey outcomes, and they will be identically equal to 1 for administrative-based outcomes. SurveyReg also has the option to output estimated standard errors on model parameters that relax the assumption of constant residual variance. The research team will use these robust “sandwich-style” standard error estimators.

---

If statistically significant differences ( $p < 0.1$  using a one-sided test) are detected, they will be ascribed to the local program. As articulated by Rubin (Holland, 1986), this rests on the assumption of SUTVA (stable-unit-treatment-value assumption). This assumption would be violated if being randomly assigned to the control group produces different results than never being recruited for the study or if there were peer-to-peer transmission of treatment effects, such as might happen in group counseling sessions. While it is not possible to rule out these two potential SUTVA violations, there is a broad consensus that these methods are useful research tools that allow for causal inferences with fewer untestable assumptions than other methods (Angrist, 2004, Cook, 2007).

The outcomes have not been selected yet. Broadly speaking, the outcomes will measure education progress and earnings. The intention though is to select a single primary outcome for each domain. The primary outcomes will correspond to confirmatory hypotheses. Because we will have a single confirmatory outcome for each of the two domains, no multiple comparison adjustment procedures will be applied. Secondary outcomes will also be tested, but estimated effects on them will be labeled as exploratory and excluded from the highest level executive summaries and press releases. The confirmatory hypotheses will be registered in a public way prior to the commencement of data analysis.

Whether local effects of treatment in five purposively selected grantees can be used to characterize the likely effects in all the other TAACCCT grantees is a far more difficult question. The evaluation team will avoid claims that the five sets of local effects can be generalized to the full set of grantees. Consideration was given to dispersing the study sample more broadly throughout the set of grantees, but the costs of doing so would be prohibitive. Given the evaluation budget available, the evaluation team, in consultation with DOL, determined the proposed study design.

The baseline data to be collected for the evaluation will be used to describe the study sample, to provide baseline measures of outcomes to use as covariates in impact estimates to improve precision, to reduce the bias from missing data, and, possibly, to define subgroups for analysis. The minimum detectable effects presented in the next section assume the impacts of TAACCCT grantees will be estimated using regression models that include baseline variables as covariates.

The estimation procedures to describe the study sample will be simple means and proportions. There will be no pooling across programs, so clustering will not be an issue. Given the lack of probability sampling and the anticipated near perfect response rates, no weights will be used either.

#### *Site Visits for Implementation Study*

##### Estimation procedures

The site visits are designed to provide in-depth qualitative information about grantees; no estimation procedures will be used. The data analysis will be descriptive.

### **B.2.3 Degree of Accuracy Required**

It is important to calculate minimum detectable effects (MDEs) before beginning an evaluation to ensure that the design provides samples large enough to detect impacts that matter to policy makers and practitioners. The MDE is the smallest true effect that a study will be able to detect at specified levels of power and statistical significance. Power refers to the probability of detecting a statistically significant impact of given size when it exists (i.e., avoiding Type II error) and typically is set to 80 percent. The statistical significance level in a hypothesis test equals the probability of rejecting the

hypothesis of no impact when it is correct and there really is no impact (i.e., making a Type I error). The standard for statistical significance in the TAACCCT impact study will be 0.10, though tables will report whether results meet more stringent levels (0.01, 0.05) as well. One-sided tests will be used because a harmful program (if it existed) would be no more interesting to policy development than a program that is merely ineffective in terms of the evaluation metrics. This is true because of the many unmeasured costs (societal and personal) associated with ineffective programs.

**Exhibit B.2: Minimum Detectable Effects (MDEs) for Experiment at each Grantee (with assumptions)**

Statistic	Percent with Substantial Educational Progress	Average Quarterly Earnings
MDE given 500 T: 500 C	6.2	\$324
Control Group Mean	50.0	\$2,863
Threshold p-value for statistical significance	0.10	0.10
Power	0.80	0.80
R <sup>2</sup> for outcome in terms of BIF and SAQ variables	0.15	0.30

Note: MDEs based on one-tailed tests. The assumptions and calculations are the same as developed by Abt Associates for the evaluation of Pathways for Advancing Careers and Education (PACE)(OMB No. 0970-0397). More details are available in the Evaluation Design Report for PACE. This report has not yet been published but can be made available to OMB. The projected variance reductions due to use of baseline variables are from Nisar, Klerman, and Juras (2013).

The exhibit shows that the impact study will be able to detect program-specific impacts on the percentage making substantial educational progress (e.g., receiving specified credentials) as small as 6.2 percentage points. The corresponding MDE for quarterly earnings is \$324. These MDE estimates assume data will be available for 100 percent of sample members, as will be the case when administrative data are used to measure outcomes. MDEs for outcomes based solely on survey data will be larger—typically about 12 percent higher (in relative terms) than those shown in Exhibit B.2.

The MDE on quarterly earnings was calculated as<sup>2</sup>

$$\begin{aligned}
 MDE &= [\Phi^{-1}(.8) + \Phi^{-1}(.9)] \sqrt{(1 - R^2) \left( \frac{\sigma_T^2}{1000} + \frac{\sigma_C^2}{1000} \right)} \\
 &= [0.842 + 1.282] \sqrt{(1 - 0.3) \left( \frac{\$3187^2}{500} + \frac{\$2540^2}{500} \right)} \\
 &= \$324
 \end{aligned}$$

The MDE for substantial educational progress was defined as the smallest value of  $\delta$  such that

$$\begin{aligned}
 \frac{\delta}{\sqrt{(1 - R^2) \left( 2 \frac{(p_c + \delta/2)(1 - p_c - \delta/2)}{500} \right)}} &\geq [\Phi^{-1}(.8) + \Phi^{-1}(.9)] \\
 \frac{\delta}{\sqrt{(1 - 0.15) \left( 2 \frac{(0.5 + \delta/2)(1 - 0.5 - \delta/2)}{500} \right)}} &\geq 2.124 \\
 \frac{0.062}{\sqrt{0.85 \left( 2 \frac{(0.5 + 0.062/2)(0.5 - 0.062/2)}{500} \right)}} &\geq 2.124 \\
 \frac{0.062}{0.0291} &= 2.1307 > 2.124
 \end{aligned}$$

The evaluation team has not calculated likely precision for the separate sample of veterans and TAA-eligible workers. Sample sizes are likely to be small for this population, especially by program. Analyses of this separate sample will probably need to be pooled across programs to be useful.

*Site Visits for Implementation Study*

Statistical techniques to ensure accuracy for the purposes described in this justification

No statistical techniques will be used to ensure accuracy.

Specialized sampling procedures to correct unusual problems.

No specialized sampling procedures will be used.

**B.2.4 Who Will Collect the Information and How It Will Be Done**

To enroll the sample in the impact study, participating Round 4 TAACCCT grantees will:

- **Conduct an Orientation and Information Session.** Participating grantees will conduct orientation and information sessions in either in a group setting or individually. Participants will

<sup>2</sup> The population standard deviations on quarterly earnings of \$3,187 and \$2,540 come from a small evaluation of a training program for young adults called Year Up (Roder and Elliott, 2011).

---

be given an opportunity to ask questions so that they will understand what study participation entails. The TAACCCT Round 4 Impact study team will develop training materials and conduct an in-person training with staff involved in the data collection process to assist in this explanation, so that program staff understand and are able to explain all aspects of the study clearly.

- **Obtain Informed Consent.** Eligible applicants to participating Round 4 TAACCCT programs will be asked if they would be willing to participate in the study and if so, complete the study's Informed Consent Form. Eligible applicants who complete this form affirmatively will be considered study participants. As shown in Appendix B, the consent form will explain that:
  - Study participants will be asked to complete at least three surveys—two brief surveys immediately after consent is provided and another 12 months after random assignment—and to give permission for the study team to request administrative data from the National Directory of New Hires (NDNH) and other extant data sources.
  - The program will use random assignment to determine which eligible applicants will be invited to participate in the TAACCCT program.

**Collect Baseline Information.** Once an eligible applicant reads and signs the Informed Consent Form, program staff will collect baseline data. After the BIF and SAQ are completed, grantee staff will use the web-based system to conduct random assignment. Applicants who choose not to participate in the study or who do not complete the BIF (as defined in footnote 3) will not be randomly assigned; however, they will not be able to access TAACCCT-funded services. Sites may choose to inform individuals of their random assignment status during the initial intake visit or shortly afterwards.

#### *Site Visits for Implementation Study*

The data will be collected through semi-structured interviews held at selected grantees. Two or three-person teams of experienced researchers will conduct 2-4 day site visits (depending on the size and nature of the grant).

### **B.2.5 Procedures with Special Populations**

#### *Baseline Data Collection for Impact Study*

To ensure participants can understand each of the documents, the Informed Consent Form, BIF and SAQ are designed at an 8th-grade readability level. The TAACCCT team will work with sites on ways staff can assist where translation of the Informed Consent Form and other data collection instruments may be needed.

At this point, it is difficult to project the added burden from administering the BIF and the SAQ to special populations (TAA-eligible workers and veterans) that are ineligible be randomized as part of the impact study. Past history suggests that despite their targeting, these populations represent a small proportion of the total population served by programs supported through TAACCCT grants. Much will depend on the outreach of the grantees and the appeal of their programs to these populations. The evaluation team estimates no more than 100 applicants in each grantee will be veterans or TAA-eligible for purposes of burden projection. Since these applicants will not be asked to participate in the lottery, Appendix B.2 contains a separate version of the informed consent form for them.



---

### **B.2.6 Use of Periodic Data Collection Cycles to Reduce Burden**

The site visits and the survey are one-time data collection efforts and will not require periodic data collection cycles.

## **B.3 Methods to Maximize Response Rates and Deal with Non-response**

### *Baseline Data Collection for Impact Study*

All individuals who agree to participate in the evaluation must complete the BIF in order to have the opportunity to be randomly assigned to the TAACCCT program.<sup>3</sup> Therefore, a response rate of 100 percent is expected at baseline on the BIF and nearly 100 percent on the SAQ as discussed in B.1.

The response rates on the protected classes (TAA-eligible workers and veterans) are likely to be lower. Given the setting in which applicants are being interviewed and are expected to fill out forms, the evaluation team anticipates a response rate of 80 percent. Accordingly, no special efforts are planned to maximize response rates on these populations.

### *Site Visits for Implementation Study*

For the site visits, it is expected that all of the grantee organizations approached will agree to participate.<sup>4</sup> Once selected sites have been confirmed, site visitors will work closely with the primary contact for each grantee in Employment and Training Administration records to help in scheduling the site visit. One member of the two-person site visit team will take responsibility for working with the primary contact person to handle the scheduling and logistics, e.g., identifying appropriate interview respondents. Dates for site visits will be set at least one month in advance to allow ample time to schedule interviews. Interview appointments will then be confirmed via e-mail the week prior to the visit. Should a potential respondent not be available during the visit, the research team will follow up with a time to interview the person by phone.

## **B.4 Tests of Procedures**

Most of the items in the BIF and SAQ are either identical or similar to questions used in previous DOL studies (including those conducted by Abt Associates) or national surveys. The few new items in the SAQ are drawn verbatim from well-validated studies. As such, all items have been thoroughly tested on large samples. In addition, Abt Associates conducted a pilot test on both forms with five current or former participants in TAACCCT-funded services. The instruments reflect minor formatting adjustments and item clarification based on feedback from the pilot test participants. The most significant of these was rephrasing the stem on item 28 of the BIF to emphasize the business of the employer rather than the job of the participant.

The field research discussion guides were pilot-tested as part of the evaluation of Rounds 2 and 3 TAACCCT grants.

---

<sup>3</sup> Sample members will have the option of skipping any item except a few key items on the BIF and skipping any item on the SAQ. Completion of the BIF therefore refers to completing the few key items that must be answered. Completion of the SAQ refers to returning the SAQ in the sealed envelope.

<sup>4</sup> The expected response rate by the grantees is 100 percent. Participation in evaluation activities is required as a condition of the grant award.

---

## **B.5 Individuals Consulted on Statistical Aspects of the Design**

The individuals listed in Exhibit **B5.1** below made a contribution to the design of the evaluation. Baseline data collection forms will be administered by grantees selected for participation in the impact study, under the direction of Abt Associates (and overseen by Ms. Gardiner as Project Director). Follow-up data collection for the Impact Study will be conducted by Abt SRBI, a subsidiary of Abt Associates, under the general direction of Ms. Gardiner. The data collected for the Impact Study will be analyzed under the direction of Mr. Judkins. Both the conduct and analysis of data for the Implementation Study will be under the direction of Ms. Gardiner.

### **Exhibit B5.1: Individuals Consulted**

<b>Name</b>	<b>Telephone Number</b>	<b>Role in Study</b>
Karen Gardiner	(301) 347-5547	Project Director
David Judkins	(301) 347-5952	Co-Principal Investigator
Lauren Eyster	(202) 261-5621	Co-Principal Investigator

Inquiries regarding the statistical aspects of the study's planned analysis should be directed to:

David Judkins	Co-Principal Investigator	301-347-5953
Dr. Molly Irwin	Senior Evaluation Specialist, Chief Evaluation Office, DOL	202-693-5091

---

## References

- Angrist, J.D. (2004). American Education research changes tack. *Oxford Review of Economic Policy*, **20**(2), 198-212.
- Cook, T.D. (2007). Randomized experiments in education: Assessing the objections to doing them. *Economics of Innovation & new Technology*, **16**(5), 331-355.
- Holland, P.W. (1986). Statistics and causal inference (with discussion). *Journal of the American Statistical Association*, **81**, 945-970.
- Judkins, D. R. and Porter, K.E. (2013). The remarkable robustness of ordinary least squares in randomized clinical trials. In *JSM Proceedings, Social Statistics Section*, [CD ROM], pp. 1235-1247. Alexandria, VA: American Statistical Association.
- Lumley, T., Diehr, P., Emerson, S., and Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annual Review of Public Health*, **23**, 151-169.
- Nisar, H., Klerman, J. and Juras, R. (2013). *Designing training evaluations: New estimates of design parameters and their implications for design* (working paper). Cambridge, MA: Abt Associates Inc.
- Roder, A., and Elliott, M. (2011). *A promising start: Year Up's initial impacts on low-income young adults' careers*. New York: Economic Mobility Corporation.