

Evaluation of Youth CareerConnect (YCC)

OMB SUPPORTING STATEMENT PART B

The Employment and Training Administration (ETA), U.S. Department of Labor (DOL), is undertaking the Evaluation of Youth CareerConnect (YCC). The overall aims of the evaluation are to determine the extent to which the YCC program improves high school students' educational and employment outcomes and to assess whether program effectiveness varies by students' and grantees' characteristics. The evaluation will be based on a rigorous random assignment design in which program applicants will be randomly assigned to a treatment group (who will be able to receive YCC program services) or a control group (who will not). ETA has contracted with Mathematica Policy Research and its subcontractor, Social Policy Research Associates (SPR), to conduct this evaluation. With this package, clearance is requested for four data collection instruments related to the impact and implementation studies to be conducted as part of the evaluation:

1. Baseline Information Form (BIF) for parents (presented in Attachment A)
2. Baseline Information Form (BIF) for students (presented in Attachment B)
3. Grantee survey (presented in Attachment C)
4. Site visit protocols (presented in Attachment D)

Additionally, a parental or guardian consent form and a student assent form are included in Attachment E.

An addendum to this package, to be submitted at a later date, will request clearance for the follow-up data collection of study participants. The full package for the study is being submitted in two parts because the study schedule requires random assignment and the implementation study to begin before the follow-up instruments are developed and tested.

1. Respondent universe and sampling

The universe of potential sites for this evaluation includes the 24 grantees awarded YCC grants in 2014. For the impact study, we plan to select a purposeful sample of approximately 10 sites from this universe of potential sites to test the effects of the YCC program on participants' academic and employment outcomes. The grantee survey will be administered to all 24 grantees to collect national information on service delivery models, staffing, staff development, partnerships, and implementation of the core program elements. The in-depth site visits will be conducted in the 10 impact study sites only, to look more closely at program operations, challenges, and successes that can help support the impact study. The site visits will involve interviews with up to 90 senior grantee staff members and 80 members of partner organizations (such as high school leaders, community college administrators, and workforce investment board members).

The universe (and sample) of students for the random assignment impact evaluation will consist of all students in the 10 study sites who apply to the YCC program during the sample intake period and who are determined eligible for the program based on existing program rules. The student universe does not consist of eligible program applicants in all 24 grantees, because the study sites will be a purposeful sample (not a random sample) and thus, the student sample

will not necessarily be representative of all YCC students nationally. Similarly, the universe (and sample) of parents for the evaluation will consist of parents of students in the student universe.

a. Site selection

The objective of the site selection is to identify up to 10 sites that we deem to be suitable candidates for participating in a random assignment evaluation. We will purposefully select the sites based on what can be learned from them and the feasibility of implementing random assignment. Impact study grantees must meet three criteria: (1) they must have enough student demand to generate a control group, (2) it must be feasible to implement study random assignment procedures at the site, and (3) there must be a significant contrast between the services available to the treatment and control groups. If more than 10 grantees meet these criteria, we will maximize the diversity of programs in the impact study to gain a greater understanding of the relationships between program models, services, and outcomes.

Sufficient number of interested students. To ensure a sample size that will yield sufficient statistical power to detect impacts, each participating site must be able to recruit and enroll enough participants to fill the treatment and control groups (about 200 to 400 eligible youth per site, as discussed further in Part B.1b below). Thus, our top priority will be to identify programs in which student demand is much larger than available program slots. As part of our site selection data collection, we will obtain information on each grantee’s planned enrollment for the remainder of the grant period. We will also seek to determine whether sites have the ability to recruit additional participants to fill a control group.

Feasibility of implementing study random assignment procedures. Some school enrollment processes may affect the program’s suitability for the study. For instance, random assignment will be easier to implement if the site’s program intake process is centralized rather than if there are multiple points of program entry. Similarly, study implementation will be more feasible if intake staff have easy access to computers so that they can input student information required for random assignment into the web-based random assignment system that will be developed for the evaluation. There are other program features that may make it more difficult to implement study procedures. For example, some schools have an absolute preference for students who live near the school and give all such students first priority. Other school programs are linked to a middle school and offer a guaranteed slot to all students coming from that school. Both practices reduce the number of students who could be randomly assigned. It would also be difficult to conduct random assignment in YCC schools that participate in “universal” or multischool lotteries in which students rank their preferred schools and receive a random lottery number, and a computer matching algorithm attempts to match students with their preferred choices.

Information on the feasibility of implementing study procedures will be obtained through a systematic examination of extant materials on all 24 grantees, telephone discussions with staff in each grantee, and initial recruiting site visits to promising sites that have been approved by OMB (OMB Approval No.: 1205-0436).

Significant contrast between treatment and control conditions. If YCC program services are similar to other services available locally, services available to the treatment and control

groups may not be distinct enough to warrant a grantee's inclusion in the impact study. This concern may be particularly relevant for YCC models that have treatment students and control students enrolled in many of the same classes. This concern is also likely to be relevant if the non-YCC classes attended by the control students are similar to the YCC classes in terms of (1) an integrated academic and career-focused curricula around one or more industry theme(s), (2) internships and work experience opportunities with employers, (3) individualized career and academic counseling that builds career and postsecondary awareness and exploration of opportunities beyond the high school experience, and (4) a small learning community.

We will define a significant treatment-control contrast in a site based on the plausibility that program impacts can be observed in the site that are of a size that the study is powered to detect (see Part B.2b below). Information on the treatment-control contrast will be obtained through a systematic examination of extant materials on all 24 grantees, telephone discussions with staff in each grantee, and initial recruiting site visits to promising sites.

The study team will rate each site using these three criteria during a multistep site selection process, with the goal of identifying approximately 10 study sites that will be required for the impact evaluation to obtain precise estimates of YCC program effects on key student outcomes (see Part B.2b below). First, we will begin with a systematic examination of extant materials on all 24 grantees, including the grantee applications and progress reports, and clarify this information with grantees. We will use a template for recording this information, complete a standardized write-up on each grantee, and record results directly into a database. The evaluators will then work with DOL to analyze this database to identify up to 16 sites with the highest ratings in the three considered categories. Based on our initial review of grantee applications and prior knowledge of similar programs, we expect that approximately 16 sites will be plausible candidates for the random assignment study. We will then conduct site visits to these 16 promising sites focusing on the services provided by the programs and on the process by which participants are recruited and enrolled into the program, to obtain an initial idea of the feasibility of implementing random assignment (these visits have been approved under OMB Approval No.: 1205-0436). After the visits, we will update the grantee database to indicate the suitability of the grantees based on the three selection criteria defined above, and conduct follow-up phone calls as necessary. It is expected that up to 6 of these sites will not be suitable for the study based on the updated information. We expect to successfully recruit the 10 remaining suitable sites to participate in the impact study.

b. Selection of students within sites

All students who meet the program eligibility requirements and consent to be part of the study will be subject to random assignment. According to the grant applications, on average, grantees plan to serve about 200 students in 2016 (the range is 40 to 800); however, actual enrollment may differ. An important part of the site selection data collection process will be to learn how actual enrollment differs from planned enrollment. We plan to purposefully select for the study those sites that have sufficient enrollment to accommodate a control group. Thus, assuming that about 200 to 400 eligible youth apply to each selected study site during the study intake period and will be subject to random assignment, we estimate that the respondent universe for the evaluation will include about 2,000 to 4,000 youth in the study sites, and that 80 percent will provide written study consent.

2. Analysis methods and degree of accuracy

Statistical methods will not be used to select the sample of sites for the random assignment study. Instead, we will select sites for study participation from among the universe of grantees and schools served by grantees based on their number of potential participants, service model, treatment differential, and suitability for implementing a random assignment design, as described in Part B.1 above. This will result in a sample of 2,000 to 4,000 students on which to base the impact estimates to achieve our target precision levels (see the power analysis in Part B.2b below).

a. Analysis methods for impact estimation

The central feature of the impact study is the random assignment of program-eligible youth in the selected sites to a treatment group that will be eligible to receive YCC program services or a control group that will not be able to receive YCC services but could receive other services available in their schools and communities. We will use experimental statistical methods to yield unbiased estimates of the impacts of the YCC programs by comparing the mean outcomes of treatment and control group members over time. Outcomes will be measured as continuous variables (for example, test scores), binary (0/1) variables (for example, whether the student stayed in school), or categorical variables (for example, math proficiency levels). Impacts on study outcomes could increase, decrease, or remain constant over time. Impacts will be estimated not only for the full sample, but also for policy-relevant subgroups defined by student, program, and site characteristics from the BIFs and other data sources. The analysis will be conducted using the SAS and STATA software programs.

Assessing baseline equivalence. If a random assignment design is conducted properly, there should be no systematic observable or unobservable differences between research groups except for the services offered after random assignment. We will assess whether randomization was conducted properly by conducting statistical *t*-tests to assess mean differences in the baseline measures of treatment and control groups using data from the BIFs. Because parent and student BIF data will be collected *prior* to random assignment, there should be no differences in data quality or response between the treatment and control groups. We will conduct *t*-tests on each baseline measure in isolation and will also conduct a joint *F*-test to assess the joint significance of the baseline differences.

Estimating impacts for the full sample. With a random assignment design, simple differences in the mean values of outcomes between students assigned to the treatment and control groups will yield unbiased impact estimates of program effects, and the associated *t*-tests can be used to assess statistical significance.

Key follow-up outcomes will be constructed using school records and student survey data collected two to three years after random assignment. These outcomes will include measures of (1) education success (school and program retention, attendance and behavior, school engagement and satisfaction, test scores and proficiency, postsecondary credits and the number of Advanced Placement classes, and educational expectations); (2) employment success (such as work-readiness skills, work experience in paid and unpaid jobs, employment expectations, and

knowledge of career options); (3) life stability (such as involvement with the criminal justice system and drug use); and (4) the school climate.

Impacts will also be estimated using regression procedures that control for baseline covariates from BIFs. This approach will improve the precision of the net-impact estimates, because the covariates will explain some of the variation in outcomes both within and between sites. Also, covariates can adjust for the presence of any differences in observable baseline characteristics between research groups due to random sampling and, for survey-based analyses, interview nonresponse.

Our benchmark model for the analysis will be a regression model in which an impact is calculated for each site, adjusted for students' baseline demographic characteristics from the parent and student BIFs:

$$(1) y_i = \sum_{k=1}^{10} \beta_k * Block_{i,k} + \sum_{k=1}^{10} \delta_k * YCC_i * Block_{i,k} + \gamma * X_i + \varepsilon_i,$$

where y_i is the outcome of student i ; $Block_{i,k} = 1$ for students in YCC program k and 0 otherwise; $YCC_i = 1$ for treatment students offered YCC program entrance and 0 for control group students; X_i represents measures of student prior achievement and demographic characteristics; ε is the error term (school-level random effects do not exist because the unit of random assignment is the student, allowing us to treat school effects as fixed); and β_k , δ_k , and γ are parameters to be estimated.

The average impact of YCC is $\sum_{k=1}^{10} \delta_k / 10$. Differences in impacts across schools can be assessed using a joint F -test of the site-level impacts (δ_k) and by comparing them to each other. We will also explore the extent to which the results are sensitive to different weighting schemes where, for example, each sample member is weighted equally or each site is weighted according to the size of the program eligible population—all are valid approaches but will provide slightly different estimates if grantees differ in size and have heterogeneous impacts. In addition, we will conduct sensitivity analyses to alternative methods to account for missing BIF data to create the baseline covariates (for example, multiple imputation or mean imputation with missing dummy variables).

Estimating impacts for subgroups. These same analytic methods can be used to obtain impact estimates for subgroups defined by student characteristics from the parent and student BIFs (for example, student risk factors, education and employment expectations, gender, and work experience history). We will also estimate impacts for subgroups defined by program features using data from the implementation analysis (for example, program organization structure, types of services, and career focuses). Subgroup analyses will address the question of whether access to grantee services is more effective for some subgroups than for others.

Impacts for subgroups will be estimated using a straightforward modification to Equation (1), where the model includes terms formed by interacting subgroup indicators with the treatment

status indicator variable and using F -tests to assess whether differences in impacts across subgroup levels are statistically significant. In addition, Equation (1) can be reformulated as the first level of a multilevel model or hierarchical linear model. In the second level, we can regress the estimated impact of each site ($\hat{\delta}_k$) on various characteristics of the site obtained from the planned implementation analysis, including key site practices and policies, to examine best practices at the grantee programs.

The analytical framework can be extended to account for treatment students who do not take up the offer of program services and control students who receive program services (known as crossovers). We will use an instrumental variable approach for this analysis by replacing the YCC_i indicator in Equation (1) with an indicator variable $PART_i$ that equals 1 for those who received YCC services and 0 for those who did not, and we will use the lottery number as an instrumental variable for $PART_i$. In addition, because students may receive different amounts of YCC services, the analysts will estimate program effects by dosage level (measured at the end of the follow-up period), using methods such as principal stratification and propensity score matching.

b. Precision calculations of the impact estimates

Based on anticipated study sample sizes of 200 treatment and 200 control group members per site, the evaluation will have sufficient statistical power to detect meaningful impacts on key study outcomes that are similar to those that have been found in impact studies of similar types of interventions. For all study grantees, we can expect to detect a significant impact on dropping out of school (a short-term outcome) if the true program impact is 3.1 percentage points or more (Table B.1); this minimum detectable impact (MDI) is similar to the rate found in an experimental evaluation of career academies (Kemple and Snipes 2000) and considerably smaller than the 12 percentage point impact found in Mathematica's evaluation of Talent Search (Constantine et al. 2006). For the subgroup of students at high risk of dropping out, the MDI is 8.2 percentage points, lower than the 10.9 percentage points found in Kemple and Snipes (2000). The MDI on achievement test scores is 0.08 standard deviations, which is much lower than the 0.25 standard deviation MDI typically targeted in school-based evaluations funded by the U.S. Department of Education.

Table B.1. Minimum detectable impacts on key outcomes

Outcome and subgroup	Minimum detectable impact (MDI)
Dropout (percentages)	
Overall	3.1
Low dropout risk	3.6
Medium dropout risk	4.0
High dropout risk	8.2
Math achievement (standard deviations)	
Overall	0.08
Low dropout risk	0.15
Medium dropout risk	0.11
High dropout risk	0.18

Note: The intention-to-treat (ITT) analysis is based on the offer of YCC services. The MDI formula for the ITT estimates is as follows:

$$factor \times design \times \sigma \sqrt{\left(\frac{1 - R^2}{r} \right) \left(\frac{1}{N_T} + \frac{1}{N_C} \right)},$$

where σ is the standard deviation of the outcome measure, r is

the survey or school records response rate, R^2 is the explanatory power of regression variables, and N_T and N_C are sample sizes for the treatment and control groups, respectively (see Schochet, 2008). Based on analysis of data from similar populations, we set $r = 0.80$ for survey outcomes. For outcomes based on school records data, we set $r = 0.87$ overall and to

0.97, 0.92, and 0.68 for the low-, medium-, and high-risk populations, respectively. R^2 is set to 0.2 for the dropout rate and 0.5 for the academic achievement test based on our previous experience. Sample sizes are based on an initial sample of 10 schools with 400 students in each, a consent rate of 0.8 for study participation and an initial assignment rate of 0.5 to each treatment condition. For subgroup estimates, we assume that the share of the analysis sample is 0.24, 0.49, and 0.27 for the low-, medium-, and high-risk students, according to the distribution in Kemple and Snipes (2000), who defined baseline risk factors based on attendance, grade point average, credits earned, whether the student was over age, whether the student had a sibling who dropped out, and the number of transfers. The MDI calculations assume a two-tailed test with 80 percent power and a 5 percent significance level, yielding a factor of 2.80 in the MDI formula.

c. Analysis methods for implementation study

The analyses for the implementation study will pull together information from myriad data sources to obtain a comprehensive picture of YCC program operations, successes, and challenges. To facilitate these analyses, the information gathered in the grantee survey and site visits will be imported into the ATLAS.ti software package, which will be used to manage and analyze all qualitative information. We will create a hierarchical coding scheme of categories and classifications linked to the research questions and program model. The software will then be used to assign codes to specific text in the electronic files. The software will enable us to cross-check the coded information obtained from multiple sources, to search for regularities and patterns in the information, and to ensure accurate conclusions are drawn.

We will conduct two types of analyses for the implementation study. The *within-site analysis* will provide a detailed site-level narrative for each site in the impact study. This narrative will be a structured document based on specific topics that will mirror the topics site

visitors addressed in their interviews. Site visitors will begin to complete this narrative before the visit based on their review of grant applications and the grantee telephone calls and will add to it during each round of site visits, detailing observations from the visit. Site visitors will also incorporate into these narratives information from other data sources—extracts from the participant tracking system (PTS), the grantee surveys, and the reviews of quarterly progress reports containing information on both short- and long-term program performance measures that will be generated by the PTS (such as information on services received, wage at placement, and type of occupation for each job placement). Ultimately, these narratives will provide a detailed profile that thoroughly describes each YCC program model and the successes and challenges it has faced, as well as the students it serves, their experiences in the program, and short-term outcomes available in the PTS. The narratives will also contain an evaluative component that highlights key findings related to the implementation of the YCC program.

The site visit narratives will be used as inputs to the *cross-site analysis* of all 24 grantees. This analysis will draw on the information from the PTS, grantee survey, grantee applications, quarterly reports, phone calls, and visits to the grantee if they occurred. The cross-site analysis will describe the services offered and practices used by YCC grantees and will evaluate the variations in their administrative setups, service delivery models, partnerships, and relationships with employers. By identifying common patterns of implementation among the grantees, the cross-site analysis will paint a picture of how YCC programs were implemented—relative to the planned program model, the types of students they served, and the successes and challenges they faced. The evaluative component will help identify factors or considerations that might help the study team to (1) understand *why* some programs were able to more effectively implement the core elements of the YCC program model than others and (2) interpret findings from the impact study relating to *why* the impacts of the services and treatments might vary from one site to another and *why* different groups of students might experience differential impacts from the program.

d. Assessing and correcting for survey nonresponse bias

A future OMB package will request clearance for the follow-up surveys for the evaluation, which will be conducted using a staggered multi-mode format: web, followed by telephone calls, followed by in-person visits for those who have not yet responded. However, parent and student BIF information will be used to assess and correct for potential survey nonresponse, which could bias the impact estimates if outcomes of survey respondents and nonrespondents differ. To assess whether survey nonresponse may be a problem for the follow-up survey, we will use two general methods:

- **Compare the baseline characteristics of survey respondents and nonrespondents within the treatment and control groups.** We will use baseline data (which will be available for the full research sample) to conduct statistical tests (chi-squared and *t*-tests) to gauge whether those in a particular research group who respond to the interviews are fully representative of all members of that research group. Noticeable differences between respondents and nonrespondents could indicate potential nonresponse bias.
- **Compare the baseline characteristics of respondents across research groups.** We will conduct tests for differences in the baseline characteristics of respondents across the treatment and control groups. Statistically significant differences between respondents in different

research groups could indicate potential nonresponse bias and limit the internal validity of the study if not taken into account.

We will use two approaches for correcting for potential nonresponse using the baseline data in the estimation of program impacts based on survey data. First, we will use regression models to adjust for any observed baseline differences between respondents across the various research groups. Second, because this regression procedure will not correct for differences between respondents and nonrespondents, we will construct sample weights so that weighted observable baseline characteristics are similar for respondents and the full sample that includes both respondents and nonrespondents. We will construct weights for each research group separately, using the following three steps:

1. **Estimate a logit model predicting interview response.** We will regress the binary variable indicating whether a sample member is a respondent to the instrument on baseline measures.
2. **Calculate a propensity score for each individual in the full sample.** We will construct this score, the predicted probability that a sample member is a respondent, using the parameter estimates from the logit regression model and the person's baseline characteristics. Individuals with large propensity scores are likely to be respondents, whereas those with small propensity scores are likely to be nonrespondents.
3. **Construct nonresponse weights using the propensity scores.** Individuals will be ranked by the size of their propensity scores and divided into several groups of equal size. The weight for a sample member will be inversely proportional to the mean propensity score of the group to which the person is assigned.

This propensity score procedure will yield large weights for survey respondents who have characteristics associated with low response rates (that is, those with small propensity scores). Similarly, the procedure will yield small weights for respondents who have characteristics that are associated with high response rates. Thus, the weighted characteristics of respondents should be similar, on average, to the characteristics of the entire research sample.

Finally, for key outcomes, we will correct for nonresponse using multiple imputation methods (using the SAS PROC MI procedure) to examine the sensitivity of our results to alternative methods for adjusting for survey nonresponse. Rubin (1986, 1996) and Shafer (1997) discuss the theory underlying multiple imputation procedures.

Assessing and correcting for site nonparticipation. As part of the recruitment process, we will collect data on key site characteristics and compare the characteristics of the selected sites that agree to participate to the characteristics of those that do not. This information will be used to help interpret the analysis findings. However, because we will not randomly select sites, but rather will purposefully select them based on their suitability for the study, the study sites will not generalize to a well-defined universe of sites. Consequently, the benchmark approach will not adjust the impact estimates for site nonparticipation, because external validity is not a well-defined concept for this evaluation. However, we will conduct sensitivity analyses that use site-level weights constructed using similar propensity score methods to those described above.

3. Methods to maximize response rates and data reliability

This study is requesting approval for the use of a set of forms to be completed as sample members go through an intake process, for a grantee survey, and for protocols to be used during visits to study sites. No monetary or non-monetary incentives will be provided to respondents. The methods to maximize response rates and data reliability are discussed first for the intake forms and then for implementation study protocols.

a. Intake documents: The consent and assent forms and the parent and student BIFs

Response rates. The study enrollment data collection forms (informed consent, student assent, and parent and student BIFs) will be administered to all program-eligible students at the selected sites during the sample intake period. These forms will be added to the application materials that programs typically use to enroll students. The project team will work closely with program staff to identify ways to integrate the BIFs into their normal program application process (whether it is electronic or on paper) in order to minimize respondents' burden and increase response rates.

We will use methods to maximize response for the consent, assent, and BIFs that have been used successfully in many other random assignment studies to ensure that the study is clearly explained to both study participants and staff and that the forms are easy to understand and complete. Care has been taken in these forms to explain the study accurately and simply to potential participants. In addition, the forms will be available in Spanish to accommodate Spanish-speaking students and parents. Program staff will be thoroughly trained to address study participants' questions about the forms and to check that the forms have been filled out properly. Grantee staff will also be provided with a site-specific operational procedures manual prepared by the research team, contact information for members of the research team, and detailed information about the study. Based on our experience with similar data collection efforts, we expect that 80 percent of eligible program applicants will participate in the study and will complete the intake forms.

Data reliability. All three forms required at intake are unique to the current evaluation and will be used across all YCC 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 DOL and will be thoroughly tested in a pretest.

Key student identifying information from the forms will be entered into the PTS for all students in the study sample prior to random assignment. This information will be required so that the PTS can check that the student was not previously randomly assigned to a research status. All forms will be sent to Mathematica for data entry into the PTS using pre-paid FedEx envelopes provided to all study sites. To ensure complete and accurate data capture, the PTS will flag missing data or data outside a valid range. Data problems will be identified during site visits and during the PTS data entry process and site liaisons will be contacted to help address and fix the problems.

b. Grantee survey and site visit data

Response rates. The grantee survey will be administered to all 24 YCC grantees in spring 2015 and again in spring 2017 to collect information on service delivery models, staffing, staff development, partnerships, and implementation of the core program elements. The study team will send the survey via email as a writeable PDF file and via mail in hard copy. Ongoing relationships with administrators through PTS support and grantee meetings as well as DOL support will facilitate a response rate of 90 percent or higher for each round of data collection.

Collecting implementation study data during site visits will help us obtain high response rates and reliable data. The recruitment process for study sites will include an explanation of the nature of the visits, to ensure that administrators know what is expected of them when they agree to participate. Site visitors will begin working with site staff well before each visit to ensure that the timing of the visit is convenient. Each round of site visits will take place over a period of months, which will allow flexibility in timing. During the visits, each day will involve several interviews and activities, which will allow further flexibility in the scheduling of specific interviews and activities to accommodate respondents' needs and program operations.

Data reliability. We will use several well-proven strategies to ensure the reliability of the site visit data. First, 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, including ways to probe for additional details to help interpret responses to interview questions. Second, this training and the use of the protocols will ensure that the data are collected in a standardized way across sites. When appropriate, the protocols use standardized checklists to further ensure that the information is collected systematically. Finally, all interview respondents will be assured of the privacy of their responses to questions.

4. Tests of procedures or methods

We have pretested the BIFs with five student and parent pairs from schools with YCC grants and the grantee survey with nine individuals from existing YCC grantees. After the pretest participants completed the forms, members of the evaluation team conducted a debriefing with participant using a standard debriefing protocol to determine whether any words or questions are difficult to understand or answer.

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. The following individuals consulted on statistical aspects of the design and will also be primarily responsible for actually collecting and analyzing the data for the agency:

Mathematica Policy Research

Dr. Peter Schochet	(609) 936-2783
Dr. Nan Maxwell	(510) 830-3726
Ms. Jeanne Bellotti	(609) 275-2243

Additionally, The following individuals were consulted on the statistical methods discussed in this submission to OMB:

Social Policy Research Associates

Ms. Sukey Leshnick (510) 788-2486
Ms. Kate Dunham (510) 788-2475
Mr. Christian Geckeler (510) 788-2461

University of California, Berkeley

Dr. David Stern (510) 642-0709

References

- Constantine, Jill, Neil Seftor, Emily Sama Martin, Tim Silva, and David Myers. "A Study of the Effect of Talent Search on Secondary and Postsecondary Outcomes in Florida, Indiana, and Texas: Final Report from Phase II of the National Evaluation." Washington, DC: U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service, 2006.
- Kemple, James J., and Jason Snipes. "Impacts on Students' Engagement and Performance in High School." MDRC, 2000.
- Rubin, D.B. (1987) *Multiple Imputation for Nonresponse in Surveys*. J. Wiley & Sons, New York.
- Rubin, D.B. (1996) Multiple imputation after 18+ years (with discussion). *Journal of the American Statistical Association*, 91, 473-489.
- Schafer, J.L. (1997) *Analysis of Incomplete Multivariate Data*. Chapman & Hall, London.
- Schochet, 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.