PART B: COLLECTION OF INFORMATION INVOLVING STATISTICAL METHODS

The U.S. Department of Labor (DOL) has contracted with MDRC [and subcontractors Mathematica Policy Research (Mathematica) and Social Policy Research Associates (SPR)] to conduct a rigorous evaluation of the 2011 YouthBuild program funded by DOL and the Corporation for National and Community Service (CNCS). The evaluation includes an implementation component, an impact component and a cost-effectiveness component. This data collection request seeks clearance for a longitudinal series of participant follow-up surveys that will be administered as part of the impact component of the evaluation. The research team plans to enroll approximately 4,200 youth into the study, with approximately 60 percent of the sample assigned to the treatment group that is eligible to receive YouthBuild services and 40 percent assigned to the control group which is not eligible to receive YouthBuild services. All study participants in 83 of the YouthBuild grantees that received funding from either DOL or CNCS in FY 2011 will participate in the impact component of the evaluation. This data collection request pertains to the 12-, 30-, and 48-month follow-up surveys. A one-time sample of 3,465 youth will be randomly selected from the full sample of 83 grantees for the fielding of these surveys. This sample of 3,465 youth will be asked to participate in each of the three follow-up surveys.

This section begins by describing the procedures that will be used to select sites for the impact component of the study and individuals for the follow-up surveys. The procedures for the study entail selecting sites for the

evaluation, enrolling eligible and interested youth at these sites into the study, randomly assigning these youth to either a treatment or control group, and randomly selecting a subset of the full study sample within these sites for the fielding of the surveys. Next, we describe the procedures that will be used to collect the survey data, paying particular attention to specific methods that will ensure high response rates. This section closes by identifying the individuals who are providing the project team with expertise on the statistical methods that will be used to conduct the empirical analyses.

B. Sampling for the Follow-up Surveys

The sample for the follow-up surveys includes 3,465 study participants across 83 sites. Youth who consent to participate in the study will be randomly assigned to one of two groups: The treatment or program group, whose members will be invited to participate in YouthBuild as it currently exists; or the control group, whose members cannot participate in YouthBuild, but will be able to access on their own other services in the community.[1]

1.   Sample Selection for the Follow-up Surveys

In May 2011, DOL awarded grants to 74 programs. After dropping three programs from the selection universe (representing less than five percent of expected youth served), the evaluation team randomly selected 60 programs for participation in the impact component of the evaluation, using

---

[1] The embargo period for this study is two years. After the embargo period, control group members can receive YouthBuild services if deemed eligible by the program.

2

probability-proportional-to-size sampling. In other words, each program had

a probability of selection that was equal to its expected enrollment over a

given program year. This method gives each YouthBuild slot (or youth

served) an equal chance of being selected for the evaluation, meaning that

the resulting sample of youth who enroll in the study should be

representative of youth served by these programs. Once a program is selected for

the evaluation, all youth who enroll at that program between August 2011 and December 2012

will be included in the research sample. In deciding on the total number of DOL-funded

programs to include in the impact component of the evaluation, we attempted to balance the

objectives of 1) maximizing the representativeness of the sample and the statistical power of the

impact analysis and 2) ensuring high quality implementation of program enrollment and random

assignment procedures. On the one hand, maximizing the representativeness of the sample and

the sample size would call for including all grantees in the study. However, substantial resources

are required to work with a study site to: a) secure staff buy-in and cooperation with the study; b)

train staff on intake and random assignment procedures; and c) monitor random assignment.

Given a fixed budget, the team determined that 60 DOL-funded programs should be selected for

the evaluation. The quality of the enrollment process (and, potentially, the integrity of random

assignment) would suffer if more than 60 sites were included. Given the expected number of

youth served per program, 60 programs was deemed adequate to generate a sample size that

would provide reasonable minimum dectectable effects on key outcomes of interest.

There are 40 programs that comprise the universe of programs that

receive CNCS but not DOL funding in 2011, and 23 of these programs

received grants of $95,000 or more from CNCS. These 23 programs were

selected for the impact component of the evaluation.

Recruitment and enrollment into the study began in August 2011 and is scheduled to continue through October 2012. Programs recruit and screen youth using their regular procedures. However, the recruitment materials do not guarantee youth a spot in the program. Program staff show applicants a video prepared by the evaluation team that explains the study. The video and site staff explains that applicants are required to participate in the study in order to have a chance to enroll in YouthBuild. Staff is provided with information to answer any questions youth may have after viewing the video presentation, before providing their consent to participate.

After youth agree to participate in the study, as discussed below, baseline data are collected from all youth before random assignment (see OMB Control Number #1205-0464).

Youth assigned to the program group are invited to start the program, which may include a pre-orientation period that youth must successfully complete in order to enroll in the formal YouthBuild program. Youth assigned to the control group are informed of their status and given a list of alternative services in the community.

In these 83 sites, we expect to enroll approximately 4,200 youth into the study. Underlying this expected number is the assumption that each site will have about 50 applicants, with 60 percent of applicants assigned to the treatment group and 40 percent of applicants assigned to the control group. From the full sample of youth, we will randomly select a subset of 3,465 youth as the survey sample.

Program impacts on a range of outcomes will be estimated using a basic impact model:

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

where: Y_j = the outcome measure for sample member i; P_i = one for program group members and zero for control group members; X_i = a set of background characteristics for sample member i; ε_i = a random error term for sample member i; β = the estimate of the impact of the program on the average value of the outcome; α = the intercept of the regression; and δ = the set of regression coefficients for the background characteristics.

Since the DOL-funded programs were selected using probability-proportional-to-size sampling, each youth served by the program was given an equal chance of being selected for the study. For this reason, the (unweighted) findings will provide an estimate of the effect of the program on the average youth served by DOL-funded sites. Similarly, findings from the CNCS only-funded sites will be representative of the effects on youth served by sites that received at least $95,000 in CNCS funding in 2011 but not DOL funding. In addition, findings from all 83 sites will be representative of the larger universe from which they were drawn. However, we will not attempt to generalize the findings to youth served by programs beyond our sampling universe, that is, all YouthBuild programs.

Table B.1 presents the Minimum Detectable Effects (MDEs) for several key outcomes, or the size of program impacts that are likely to be observed or detected for a set of outcomes and a given sample size. MDEs are shown

for the survey sample of 2,772, assuming an 80 percent response rate from a fielded sample of 3,465.

The MDE for having a high school diploma or GED is 4.6 percentage points. The MDE for a subgroup comparison, assuming the subgroup comprises half of the sample, is 6.5 percentage points. These effects are policy relevant and well below the effects on educational attainment found from several other evaluations of programs for youth. For example, impacts on GED or high school diploma attainment were 15 percentage points and 25 percentage points in the Job Corps and ChalleNGe programs, respectively.

MDEs for short-term employment rates are shown in the second column. These MDEs are similarly 4.6 percentage points for a full sample comparison and 6.5 percentage points for a subgroup comparison. Several evaluations of programs for disadvantaged youth have found substantial increases in short-term employment rates, although these gains tended to diminish in size over time. The Job Corps evaluation, for example, found impacts on annual employment rates of 10-to-16 percentage points early in the follow-up period using Social Security Administration records. However, Job Corps' effects on employment when measured with survey data were

Table B.1
Minimum Detectable Effects for Key Survey Outcomes

	Has high school diploma/GED	Employed since random assignment	Annual earnings
Survey			
Full sample (2,772)	0.046	0.046	$1,016

6

Key subgroup (1,360)   0.065   0.065 $1,436

Notes: Assumes an 80 percent response rate to the follow-up surveys and a random assignment ratio of 60 program to 40 control.

Average rates for high school diploma/GED receipt (45 percent) and employment (52 percent) are based on data from the YouthBuild Youth Offender Grantees. For annual earnings, the average ($11,000) and standard deviation ($11,000) are based on data for youth with a high school diploma or GED from the CET replication study. Calculations assume that the R-squared for each impact equation is .10.

substantially smaller, ranging from 3to-5 percentage points per quarter.

Other effects on employment include an impact of 4.9 percentage points after 21 months in ChalleNGe and an impact of 9.3 percentage points after 30 months for women in the high fidelity CET replication sites.

MDEs for earnings are shown in the final column. With the survey sample size, we would be able to detect as statistically significant an impact of at least $1,016 on annual earnings during a given follow-up year. Assuming a control group average of $11,000 in annual earnings, for example, this impact would represent a nine percent increase in earnings. MDEs for a subgroup comparison are larger, at $1,436, or a 13 percent increase. These effects are large, even relative to the significant effects generated by the Job Corps program. Career Academies, however, did lead to substantially larger earnings gains than these, particularly for the men in the sample.

2.   Procedures for the Collection of Information

The follow-up surveys are the primary source of outcome data to assess program impacts. These data will inform DOL and CNCS on how YouthBuild programs affect program participants' educational attainment, postsecondary planning, employment, earnings, delinquency and involvement with the criminal justice system, and youth social and emotional development.

For the YouthBuild evaluation, we will use a multi-mode survey that begins on the web, and then moves to more intensive methods – Computer Assisted Telephone Interviewing (CATI) and in-person locating – as part of our non-response follow-up strategy. To encourage youth to complete the survey early and online, we will conduct an experiment during the 12-month follow-up survey that will inform our data collection strategy for the subsequent rounds. Survey sample members will be randomly assigned to one of two incentive conditions: 1) "the early bird" and 2) the control condition. The "early bird" group will receive a higher incentive of $40 for completing their survey online within the first four weeks of the field period, or $25 for completing the survey thereafter regardless of mode of completion. The control group (which will consist of members of the YouthBuild evaluation's treatment and control group) will receive $25 regardless of when they complete the survey. We will share the findings, along with other findings from the 12-month follow-up survey, with DOL shortly after completion of the survey. For those sample members who do not complete the survey on the web, evaluation team members will attempt to complete the survey over the phone using a CATI survey. For those

sample members who cannot be located by telephone, evaluation team members will use custom database searches and telephone contacts to locate sample members. Field locators will use the last known address of residency and field locating techniques to attempt to find sample members who cannot be found using electronic locating methods. Field locators will be equipped with cell phones and will be able to dial into Mathematica's Survey Operations Center before giving the phone to the sample group member to complete the CATI survey. We expect that each of the follow-up surveys will take approximately 40 minutes to complete.

Earlier Paperwork Reduction Act clearance packages requested Office of Management and Budget (OMB) clearance for a site selection questionnaire, the baseline forms and participant service data, the grantee survey, site visits protocols and the cost of data collection. This OMB package pertains only to the follow-up surveys to be administered to study participants.

3.   Methods to Maximize Response Rates and Data Reliability

We expect to achieve an 80 percent response rate for each round of the follow-up surveys. The YouthBuild data collection plan incorporates strategies that have proven successful in our other random assignment evaluations of disadvantaged youth. The key components of the plan include:

- Selecting data collection modes that maximize data quality and are appropriate for the population under study;

- Designing questionnaires that are easily understood by people with low-literacy levels and/or limited English-language proficiency;

- Implementing innovative outreach and locating strategies through social media outlets such as Facebook, MySpace and Twitter; and

- Achieving high response rates efficiently and cost effectively by identifying and overcoming possible barriers to participation.[2]

As mentioned earlier, we will use a multi-mode approach that will begin on the web and then move to CATI and in-person locating. The advantage of this multi-mode approach is that it allows participation by people who do not have listed telephone numbers, telephone land lines, or access to a computer or the internet; it also accommodates those who have a preference for self-administration. The use of field locators instead of in-person interviewing minimizes possible mode effects by limiting the number of modes to two: (1) self-administered via the web, and (2) interviewer-administered over the telephone. We gave careful consideration to offering a mail option; however, youth's high mobility rates and increasing use of the

---

[2] These procedures have proven highly effective in other Mathematica studies, such as the evaluation of AT&T Foundation's Aspire Program. For this study, we implemented a web-only survey of 172 grantees: school districts, school district foundations, and non-profit organizations that provided support to at-risk high school students. During a one-month field period, we achieved a 97 percent response rate. While this high rate demonstrates the potential of web-based surveys, our experience with other projects indicates that more than one month is often needed to achieve a high response rate with web surveys. For example, in an evaluation of Early Head Start Programs that serve low-income pregnant woman and children aged 0-3, we achieved an 89 percent response rate using web surveys, with telephone and mail follow-up over a five-month period. Based on our past experience, we propose conducting the web-based grantee survey for this study across a three-month period.

internet suggest that a mail option would not be cost-effective. [3,4] We will

provide a hardcopy questionnaire to any sample member who requests one

but, increasingly, studies of youth find a preference for completing surveys

online.[5,6,7]

The questionnaire itself will be designed to be accessible at the 6th

grade reading level. This will ensure that the questionnaire is appropriate for

respondents with low literacy levels and/or limited English proficiency.

Throughout the data collection process, we will implement a variety of

outreach strategies to encourage respondent participation and facilitate

locating for the 12-, 30- and 48-month follow-up surveys. A key component

of this outreach strategy will be the development of a YouthBuild Evaluation

presence on social media outlets, such as Facebook, MySpace and Twitter.

Because youth are more likely to have a stable electronic presence than

---

[3] The Census Bureau estimates that about 14 percent of the overall population moves within a one-year period. However, the highest mobility rates are found among young adults in their twenties. Thirty percent of young adults between the ages of 20-24 and twenty-eight percent of those between the ages of 25-29 moved between 2004 and 2005, the most recent years for which these data are available (.http://www.census.gov/population/www/pop-profile/files/dynamic/Mobility.pdf) For the YouthBuild evaluation, this means that approximately one-third of our sample is expected to move between the time of random assignment and the first follow-up survey; and a higher percentage is likely to relocate prior to the second and third follow-up surveys.

[4] Lenhart, Amanda and Kristen Purcell, Aaron Smith and Kathryn Zickuhr. "Social Media and Young Adults." February 2010. http://www.pewinternet.org/Reports/2010/Social-Media-and-Young-Adults.aspx

[5] Lygidakis, Charilaos, Sara Rigon, Silvio Cambiaso, Elena Bottoli, Federica Cuozzo, Silvia Bonetti, Cristina Della Bella, and Carla Marzo. "A Web-Based Versus Paper Questionnaire on Alcohol and Tobacco in Adolescents." *Telemedicine Journal & E-Health,* vol. 16, no. 9, 2010, pp. 925-930.

[6] McCabe, Sean E. "Comparison of Web and Mail Surveys in Collecting Illicit Drug Use Data: A Randomized Experiment." *Journal of Drug Education,* vol. 34, no. 1, 2004, pp. 61-72.

[7] McCabe, Sean E., Carol J. Boyd, Mick P. Couper, Scott Crawford, and Hannah D'Arcy. "Mode Effects for Collecting Alcohol and Other Drug Use Data: Web and U.S. Mail." *Journal of Studies on Alcohol,* vol. 63, no. 6, 2002, pp. 755.

physical address, we will use social media platforms to communicate information about the study, including the timing of follow-up surveys, and assist in locating efforts as necessary. We will also use a participant advance letter (Appendix B) to introduce each round of data collection, and will send a round of interim letters (Appendix C) that include YouthBuild evaluation-branded items, such as USB drives or pens.

Mathematica will begin making phone calls approximately four weeks after the launch of the web survey. The web survey will remain active during this time. Mathematica's sample management system allows for concurrent multi-mode administrations and prevents surveys from being completed in more than one mode. Bilingual interviewers conduct interviews in languages other than English.

Sample members will be assigned to field locators if they do not have a published telephone number, cannot be contacted after multiple attempts, or refuse to complete the survey. Field locators will be informed of previous locating efforts, including efforts conducted through social media outlets, as well as interviewing attempts. When field locators find a sample member, they will gain cooperation, offer the sample member the use of a cell phone to call into Mathematica's Survey Operations Center and wait to ensure the completion of the interview. Table B.2 shows the expected response rates by mode for each follow-up survey. Because the results of the incentive experiment are unknown at this time, Table B.2 assumes that the higher incentive payment did not have a significant effect on response rates or costs. If the "early bird" incentive proves effective, we would expect to see a

larger proportion of cases completed on the web over time. This table will be

updated as part of our report on the findings from 12-month follow-up

survey's incentive experiment. As shown in Table B.2, our expected

response rates are based on our previous experiences with similarly aged

populations. On the 2008 NSRCG study, Mathematica completed 66 percent

of the surveys via a web instrument, and the College Experiences Survey

was able to complete 80 percent of its surveys by web.[8, 9] We reduced our

expected proportion of web completes for the YouthBuild survey because

this is a more disadvantaged population, and we anticipate needing more

interaction with sample members in order to have them complete the

survey. Our experience with longitudinal samples indicates that, over time,

we will need to do more locating and require more direct interaction with

sample members to get them to complete a survey, therefore our

assumptions for the percentage of completes from field staff were increased

for the second and third follow-ups.

Table B.2

Anticipated Response Rates by Mode of Administration

	Completion Rates (Percents)				Nonresponse (Percents)		
Survey Round	**Web**	**Telephone (CATI)**	**Field/ Cell Phone**	**Total Completion Rate**	**Refusals**	**Unlocatable**	**Total Nonresponse**
12-Month Survey	20	40	20	80	5	15	20

---

[8] Mooney, Geraldine M. "National Survey of Recent College Graduates (NSRCG) 2008: Impact of the 2008 Incentive Experiment on 2008 NSRCG Costs." Report submitted to the National Science Foundation. Princeton, NJ: Mathematica Policy Research, October 2011.

[9] DesRoches, David, John Hall and Betsy Santos. "College Experience Survey: Methodological Summary." Draft Report. Princeton, NJ: Mathematica Policy Research. November 2009.

30-Month Survey	25	30	25	80	5	15	20
48- Month Survey	25	25	30	80	5	15	20

We will use three key steps to ensure the quality of the survey data: 1) a systematic approach to developing the instrument; 2) in-depth interviewer training; and 3) data reviews, interviewer monitoring and field validations. The follow-up surveys include questions and scales with known psychometric properties used in other national studies with similar populations. We will use Computer Assisted Interviewing (CAI) to: manage respondent's interaction across modes so that respondents cannot complete the questionnaire on both the web and through CATI; ensure that respondents can exit the survey in either mode and return later to the last unanswered question, even if they decide to complete in a different mode; integrate the information from both of the data collection modes into a single database; and build in data validation checks within the surveys to maintain consistency.

We will conduct in-depth interview training with select telephone interviewers and locators who have the best track records on prior surveys in gaining cooperation and collecting data from youth or vulnerable populations. The training will include identifying reasons for non-response and strategies for gaining cooperation among those who have decided not to participate by web. We will train interviewers to address more common reasons youth do not respond to surveys by using role-playing scenarios. A key component of our training will focus on gaining cooperation from control

group members, who may feel less inclined to participate in a YouthBuild-related study because they were not invited to participate in the program itself.

Data reviews will be conducted throughout the field period to assess the quality of the data. Mathematica will use several standard techniques to monitor the quality of data captured and the performance of both web and CATI instruments. We will review frequencies and cross tabulations for inconsistencies a few days after the start of data collection, as well as throughout the data collection period. It is standard Mathematica practice to monitor ten percent of the hours each interviewer spends contacting sample members and administering interviewers ensuring that interviewer performance and survey administration are carefully monitored throughout data collection. Interviewers will be trained to enter comments when respondents provide responses that do not match any of the response categories or when respondents insist on providing out-of-range responses. Ten percent of the field-generated completes will be validated through a combination of mail and telephone validation techniques that ensure that the proper sample member was reached, the locator behaved professionally, and the date of the interview was reported correctly.

The follow-up questionnaire was pretested, as discussed in section 4 below.

Finally, we will conduct a response analysis for each survey prior to using the data to estimate program impacts. First steps in this analysis are to: 1) compare respondents to non-respondents on a range of background

characteristics; 2) compare program group respondents with control group respondents on a range of background characteristics; and 3) compare impacts calculated using administrative records data for respondents versus non-respondents. If these analyses suggest that the findings from the respondent sample cannot be generalized to the full sample, we will consider weighting (using the inverse predicted probability of response) or multiple imputation. However, the combined adjustment methods are not a complete fix, since both assume that respondents and non-respondents are similar on unobservable characteristics. For this reason, results from this adjustment will be presented with the appropriate caveats.

4. Test of Procedures or Methods

Pretesting all surveys is vital to the integrity of data collection. We reviewed previously used questions and developed new questions for the evaluation according to the following guidelines:

- Questions will be worded simply, clearly, and briefly, as well as in an unbiased manner, so that respondents can readily understand key terms and concepts.
- Questions will request information that can reasonably be expected of youth.
- Question response categories will be appropriate, mutually exclusive, and reasonably exhaustive, given the intent of the questions.
- Questions will be accompanied by clear, concise instructions and probes so that respondents will know exactly what is expected of them.
- For the YouthBuild population, all questions will be easily understood by someone with a sixth-grade reading level.

Prior to formal pretesting, the evaluation team completed internal testing of the instrument to estimate the amount of time it will take respondents to complete each follow-up survey. We found that the survey took approximately forty minutes when averaged across four internal tests.

During formal pretesting, we pretested each of the follow-up surveys with nine respondents some of whom are participating in YouthBuild programs that are not in the study and others who are similar to the YouthBuild target population but not participating in the program. Cognitive interviews administered as either think-aloud protocols or retrospective protocols were conducted as part of our pretest efforts. We pretested in each survey mode by having five participants complete a self-administered version of the questionnaire and four complete an interviewer-administered version. Each interview included a respondent debriefing, administered by senior project staff that assessed respondent comprehension, clarity of instruction, question flow, and skip logic. We also conducted debriefings with interviewers to assess the survey timing and ease of administration. Detailed findings from the cognitive interview pretest are included in Appendix G.

5.   Individuals Consulted on Statistical Methods

There were no consultations on the statistical aspects of the design for the youth follow-up surveys.

Contact information for key individuals responsible for collecting and analyzing the data:

Lisa Schwartz, Task Leader for the Surveys
Mathematic Policy Research
P.O. Box 2393

Princeton, NJ 08543
609-945-3386
lschwartz@mathematica-mpr.com

Cynthia Miller, Project Director and Task Leader for the Impact Component
MDRC
16 East 34th Street
New York, NY 10016
Cynthia.miller@mdrc.org
212-532-3200