Final Report for AmeriCorps Opportunity Youth Evaluation Bundling

Submitted: January 16, 2018

Submitted to:
Adrienne DiTommaso
Senior Evidence and Data Use Specialist
Department of Research and Evaluation
Corporation for National and Community Service
250 E St SW, Suite 300
Washington, DC 20024

Submitted by:
JBS International, Inc.
555 Airport Blvd, Suite 400
Burlingame, CA 94010-2002

*Value of **thought.***
*Value of **solution.***

JBS INTERNATIONAL

Final Report for AmeriCorps Opportunity Youth Evaluation Bundling

January 16, 2018

Authors

This report was prepared by JBS International, Inc., 555 Airport Blvd., Suite 400, Burlingame, CA 94010 pursuant to a contract with CNCS, a federal agency of the United States Government.

Rebecca Frazier, Ph.D., Research Associate
Donald J. Pratt, Ph.D., Research Associate
Gina Cardazone, Ph.D., Research Associate
Miriam Hartig, MPH, Research Associate
Jenny Liang, B.A., Research Assistant
Annie Georges, Ph.D., Senior Research Associate

Acknowledgements

We would like to thank the AmeriCorps grantees, program staff, and members that made this evaluation possible.

Citation

Frazier, R., Pratt, D., Cardazone, G., Hartig, M., Liang, J., Georges, A. (2018). Final Report for AmeriCorps Opportunity Youth Evaluation Bundling. North Bethesda, MD: JBS International, Inc.

Table of Contents

Table of Exhibits

Abstract

Each year the Corporation for National and Community Service's (CNCS) AmeriCorps program engages more than 80,000 members in service with non-profits and community groups nationwide. Recently, AmeriCorps has begun recruiting opportunity youth (16-24 year-olds who are disconnected from school or work) with the goal to increase these youth's employment, education, and civic engagement. However, it has been challenging to evaluate the impact of AmeriCorps service for opportunity youth due to small sample sizes within programs and limited staff capacity. To address this challenge, CNCS decided to bundle together 19 small AmeriCorps programs (each receiving less than $500,000 from CNCS) into a single quasi-experimental evaluation using a propensity-score matched comparison group. Surveys were administered to treatment and comparison youth at three time points – at the start, end, and six-months after their service with AmeriCorps. Results indicated that AmeriCorps opportunity youth showed improvements from pre-test to follow-up in high school completion, course completion, percentage currently employed, hourly wages, and longest employment-- but in the matched sample treatment youth were not more likely to show improvements than comparison youth. Youth who had a prior criminal history at the start of the program were significantly less likely to be facing criminal charges or to be on probation or parole at the six-month follow-up if they were in the treatment group. Additionally, treatment youth who had been employed at some point in the six months leading up to their AmeriCorps service were more likely to be employed at follow-up than comparison youth with a similar employment history. There was also a marginally significant positive effect of treatment on course completion which was largely driven by youth who were not high school graduates. Together, these results suggest that although treatment youth did not necessarily improve more than comparison youth, they did improve over time and AmeriCorps service seemed to be especially effective for opportunity youth with a prior criminal history.

Executive Summary

This report describes the results from an evaluation that examined the impact of participation in the Corporation for National and Community Service's (CNCS) AmeriCorps program on the educational, employment, and civic engagement of opportunity youth. Opportunity youth were defined as "people between the ages of 16 and 24 who are low-income and either homeless, in foster care, involved in the juvenile justice system, unemployed, or not enrolled in or at risk of dropping out of an educational institution." (Consolidated Appropriations Act, 2014[1]) The evaluation used an innovative "bundling" approach to combine 19 small programs with similar program models and intended outcomes into a single evaluation. The bundling approach is designed to meet the needs of small programs that lack the evaluation capacity and organizational resources to conduct an impact evaluation on their own. By participating in a bundle with other programs, we hypothesized that the members of a bundled evaluation could attain higher levels of evaluative evidence and support a more effective use of evaluation resources than would be possible for the individual programs on their own. The bundling process utilized participants across programs to attain a sample of sufficient size to support an impact evaluation with a comparison group to provide a rigorous assessment of evidence of effectiveness for their work with opportunity youth. The evaluation also had the explicit goal of enhancing the evaluation capacity of small grantees to gather and utilize outcome data.

The evaluation's primary objective was to determine whether opportunity youth who participated in AmeriCorps would have:

- Increased educational enrollment and attainment,
- Increases in the number of employment seeking activities,
- Increased job attainment and retention,
- Improved positive civic engagement, and
- Decreased likelihood of engagement with the criminal justice system.

The sample for this study consisted of a cohort of opportunity youth members who participated in one of 19 small AmeriCorps State and National (ASN) programs or program sites for one year, and a group of comparison participants who either applied to the program during the same year but declined to participate or were not selected for participation, or were selected from the geographic region of a corresponding AmeriCorps program and did not participate in an AmeriCorps program. The primary data sources for the evaluation were surveys collected from participant and comparison group members and data obtained from the programs' administrative records.

The evaluation design was a multi-program quasi-experimental propensity score matched study intended to draw causal conclusions about the overall impact of small ASN programs engaging opportunity youth as members. The propensity score method matched treatment and comparison

[1] This definition comes from the definition of disconnected youth in the authorizing legislation for the Performance Partnership Pilot program in the Consolidated Appropriations Act of 2014. The definition was modified slightly to include a narrower age range of 16-24 rather than the original range of 14-24 since youth must be at least 16 years old to participate in AmeriCorps. Additionally, the term "opportunity youth" was used instead of "disconnected youth" because of its more positive connotation and increasing use within the youth development field.

youth based on their responses at baseline. The final matched sample consisted of 124 treatment youth and 90 comparison youth from 15 sites with program sample sizes ranging from two to 40 youth. Multilevel multivariate linear and logistic regressions were then used to predict outcomes for the matched sample, while controlling for baseline characteristics and random effects due to the nesting of individuals within sites. Exploratory analyses for the treatment only sample were conducted using paired sample tests (Wilcoxon tests and McNemar tests) to examine differences from pre-test to follow-up among treatment youth.

Results indicated that AmeriCorps opportunity youth showed improvements from pre-test to follow-up in high school completion, course completion, percentage currently employed, hourly wages, and longest employment. However, analysis with the matched sample showed that treatment youth were not more likely to show improvements than comparison youth. Youth who had a prior criminal history at the start of their AmeriCorps service were significantly less likely to face charges or to be on probation or parole at the six-month follow-up if they were in the treatment group. Additionally, treatment youth who had been employed at some point in the six months leading up to their AmeriCorps service were more likely to be employed at follow-up than comparison youth with a similar employment history. There was also a marginally significant positive effect of treatment on course completion which was driven by youth who were not high school graduates. Together, these results suggest that although treatment youth did not necessarily improve more than comparison youth, they did improve over time. Furthermore, within the matched sample treatment seemed to be especially effective for opportunity youth who were not high school graduates and those with a prior criminal history.

Introduction

Evaluation Purpose and Implications for CNCS

According to the CNCS AmeriCorps authorizing legislation,[2] programs are required to conduct evaluations to assess the outcomes and impacts of their services to improve programming. Larger AmeriCorps programs (i.e., those with over $500,000 of annual funding) are better equipped to access evaluation resources to support rigorous outcome and impact studies than smaller programs (i.e., those with less than $500,000 of annual funding). The smaller programs face significant challenges with evaluation capacity. In the non-profit and social service world, the quality of evaluation studies is increasingly becoming a central consideration for funders, policy makers, and clients who seek to support and use programs with evidence of effectiveness. The emerging importance of acquiring sound evaluation evidence to make strategic decisions about programs, such as how and where to target improvements within a program and in funding programs, suggests that smaller programs would benefit from accessing evaluation expertise from external evaluators like those utilized by larger programs.

The AmeriCorps Opportunity Youth Evaluation Bundling project evolved as a potential solution to the challenge of funding and conducting rigorous evaluation studies among small AmeriCorps programs receiving less than $500,000 of annual funding. Instead of each program conducting its own evaluation using minimal resources, CNCS aimed to "bundle," or cluster a group of small AmeriCorps programs with similar activities, outputs, and outcomes. By pooling evaluation resources from each of the programs in the bundle, these programs could access high-quality evaluation expertise and increase the sample size necessary to conduct impact evaluations that include a comparison group. Involvement in an evaluation also gave the participating programs an opportunity to build evaluation capacity. Following a pilot effort in 2013, which assessed the feasibility of this approach, Year 1 activities consisted of screening and selecting AmeriCorps programs serving opportunity youth for participation in a bundled evaluation, and preparation for the current study. This preparation included the development of a collective logic model, the design and pilot testing of a survey instrument, the development of an evaluation plan, and the submission of a clearance package to the Office of Management and Budget. Year 2 activities focused on comparison group recruitment and administering pre-test and some post-test surveys to youth. Year 3 activities included the implementation of the post-test and the six month follow-up surveys and evaluation capacity building activities for participating programs.

This report provides an overview of these key components and results of the evaluation for audiences within the Federal Government, the program evaluation field, foundations, and non-profits. The report is intended to provide information useful to entities interested in developing evaluation plans that programs with similar resources, clients, activities, and outcomes can share.

The Bundled Evaluation Approach

The evaluation used an innovative "bundling" approach to combine a group of small programs with similar program models and intended outcomes into a single evaluation. The bundled

[2] The National and Community Service Act as Amended by The Edward M. Kennedy Serve America Act, 42 U.S.C. 12501 Sec.131(b)1A. See https://www.nationalservice.gov/sites/default/files/documents/1990_serviceact_as%20amended%20through%20pl%20111-13.pdf

evaluation approach is a hybrid of traditional multi-site and clustered evaluation approaches. While multi-site evaluations often focus on confirming the effects of a single well-established program model, clustered evaluations are typically more exploratory and focus on promoting learning across distinct program models (Barley & Jenness, 1993). A bundled evaluation merges these two approaches to allow for rigorous impact evaluation (in this case, by using propensity scores to match participants with comparison youth within each program and a hierarchical/multi-level model), while also helping diverse programs identify shared metrics, find relevant comparison groups, and build learning communities. The bundled approach to evaluating small programs also borrows heavily from participatory evaluation methods and builds on theory suggesting that evaluation capacity building (or evaluative inquiry) is dependent upon the development of a learning organization (Preskill & Torres, 1999) and requires stakeholders to participate in the evaluation process (Forss, Kruse, Taut, & Tenden, 2006). The collaborative nature of a bundled evaluation approach is uniquely positioned to promote organizational learning and build evaluation capacity—especially when programs are given frequent individualized feedback about their results and their progress toward long-term evaluation goals.

This approach has a higher likelihood of demonstrating strong results by pooling participants across the participating programs and increasing statistical power by creating a larger sample size. The bundling approach is especially designed to meet the needs of smaller programs that lack the evaluation capacity and organizational resources to conduct an impact evaluation on their own. By participating in a bundle with other programs, we hypothesized that the members of a bundled evaluation can attain higher levels of evaluative evidence and support a more effective use of evaluation resources than would be possible for the individual programs to attain on their own. The bundling process is intended to attain sufficient sample size to support an impact evaluation with a comparison group to assess the effectiveness for their work with opportunity youth. The evaluation also had the explicit goal of enhancing the evaluation capacity of small grantees to gather and utilize outcome data.

Bundling was achieved through a grantee screening and selection process that considered an array of program characteristics to assess each program's evaluability and compatibility with the overall bundle. The screening involved reviewing program documents and interviewing program staff to determine that each candidate program could contribute treatment and comparison group members and that the programs were sufficiently aligned in their programmatic activities and outcome focus to allow them to share a common set of research questions that could be addressed using a single evaluation approach, methodology, and set of measures.

Evaluating Opportunity Youth

Opportunity Youth (OY) are young adults between the ages of 16 and 24 who are neither employed in the labor market nor enrolled in school (Aspen Institute, 2014). One report estimated more than 6.7 million opportunity youth in the US (or roughly 17% of the total youth population between the ages of 16 and 24, see Belfield, Levin & Rosen, 2012). Opportunity youth are demographically diverse and include black and Hispanic youth, returning veterans, teen parents, immigrants, the homeless (including LGBT youth), those with mental and/or physical challenges, and those who are incarcerated (Belfield, Levin, & Rosen, 2012). These youth are called "opportunity youth" to reflect their perseverance in seeking viable pathways to economic stability and as a reminder that supporting their goals is an important step toward a

brighter future for employers and the nation (Aspen Institute, 2014; Belfield, Levin, & Rosen, 2012; Corcoran et al., 2012; Stuart Foundation, 2011). The Corporation for National and Community Service (CNCS), the parent agency overseeing the ASN program, defines opportunity youth as "economically disadvantaged individuals age 16-24 who are disconnected from school or work for at least six months prior to service," but the current evaluation expanded this definition to include "people between the ages of 16 and 24 who are low-income and either homeless, in foster care, involved in the juvenile justice system, unemployed, or not enrolled in or at risk of dropping out of an educational institution," (Congressional Appropriations Act, 2014).

The opportunity youth population was selected following a proof-of-concept test conducted in 2013 that indicated that grantees engaging opportunity youth as AmeriCorps members:

1. Demonstrated homogeneity in target outcomes, potential measures, and type of service experience or intervention,
2. Were likely to have sufficient numbers of eligible members within each participating program, and
3. Were likely to benefit from a multi-program evaluation plan.

Key Outcome Areas

The key outcome areas evaluated were education, career support, and community connection. These outcomes are particularly relevant to OY because achieving these results would represent success in overcoming the barriers these youth face in achieving economic self-sufficiency and reconnecting with their communities in positive ways.

Education

Developing pathways for opportunity youth who are having educational difficulties or dropped out of school is a core component of education support (Belfield, Levin, & Rosen, 2012; Bird et al., 2014; Corcoran et al., 2012; Gennetian, 2012; Gewertz, 2011; Julian & Kominski, 2011; Rutschow & Crary-Ross, 2014). Alternative schools and recuperative schools have the potential to support opportunity youth more effectively than traditional high schools (Lochner & Moretti, 2004). These nontraditional schools may also serve to connect opportunity youth to employment networks that could be valuable once their education and training is complete. Education support also encompasses building flexibility into education policy so that funding for the challenges faced by opportunity youth is more easily accessed (Belfield, Levin, & Rosen, 2012; Corcoran et al., 2012). In this way, improving the education system's flexibility can facilitate a more holistic approach to supporting opportunity youth.

Studies of low-income youth note that education supports can function to lift them and their families out of poverty and offer a buffer for the next generation against the threat of poverty (Bird et al., 2014; Martin & Broadus, 2013). Bird's work argues that educational advancement in one generation can provide momentum, knowledge, and resources for future generations to attain postsecondary degrees and succeed in the labor market (Bird et al., 2014). Kemple and Willner (2008) found strong evidence for the benefits of education support for students in low-performing school districts at risk of dropping out. Their randomized trial revealed that monthly earnings, number of months employed, hours worked per week, and hourly wages were higher

among those who received support and graduated compared to similar youth who did not graduate. Although numerous studies of youth in poverty suggest that education programs improve youths' life chances, research on such programs specific to opportunity youth are just beginning to surface. Martin and Broadus's (2013) randomized control trial of the GED Bridge program found that program participants were more likely to complete the GED course and pass the GED exam than students who receive treatment as usual. Bridgeland and Mason-Elder (2012) found similar results for youth in poverty who were receiving education support (many of whom were opportunity youth). They were more likely to attain their high school degree and earn college credits than youth who were not getting support.

Career Support

Career development among opportunity youth requires the assistance of local colleges and businesses, nonprofits, and other stakeholders in providing channels between education and labor market needs (Allen et al., 2014; Belfield, Levin, & Rosen, 2012; Mortimer, 2010; Snyder & Dillow, 2011; Vericker et al., 2009; Wallace, 2014). These entities need to develop in ways that facilitate smooth transitions from education to employment. One way can be through highlighting career pathways and through addressing business leaders' negative perceptions about hiring opportunity youth (e.g., that opportunity youth do not have skills to contribute to the workforce, that they are 'risky' employees) (Belfield, Levin, & Rosen, 2012; Corcoran et al., 2012; Sum et al., 2014). Creating internships and apprenticeships for opportunity youth can give employers the chance to learn about a youth's likelihood of contributing to their business or organization before offering permanent employment. Collective impact approaches to support opportunity youth have carved new grooves into local economies to bridge gaps between opportunity youth and the employment opportunities they seek. These approaches encourage businesses, government, nonprofits, and schools to work together in local communities to give opportunity youth access to credentials and connections so they can contribute to local economies and support themselves and their families (Allen et al., 2014; Corcoran et al., 2012; Wallace, 2014).

Opportunity youth often find themselves working in low-pay jobs with little room for advancement. Internship and training programs are avenues that have shown promise in supporting opportunity youth to improve their career prospects (Mortimer, 2010; Sum et al., 2014). The Urban Alliance has been particularly successful in supporting opportunity youth through internships (Theodos et al., 2014). A recent outcome study of the Urban Alliance's internship program for at-risk youth showed that more than 90 percent of participants reported positive feelings about working in professional office environments, noting that the internships provided both the hard and soft skill development necessary to succeed in these jobs (Theodos et al., 2014). An experimental study of the Job Corps program, an initiative that provides education and employment internships to disadvantaged youth (ages 16 to 24) found that participants had higher annual earnings than the control participants three and four years after completing the program (Schochet et al., 2008). Other programs have addressed those in need, some of whom are opportunity youth, through collaborations between local governments and businesses. The Transitional Work Corporation (TWC) participated in a randomized control trial that showed statistically significant improvement in employment in unsubsidized jobs, increases in earnings, and less reliance on food stamps for the treatment groups versus control participants who were receiving government assistance (Bloom et al., 2009).

Community Connection

A third factor related to opportunity youth is encouraging commitment to community connection (Corcoran et al., 2012; Wahl et al., 2012). Some of the research posits that supports to assist opportunity youth in knowing what resources are available are keys to supporting connections to the community, such as financial support and literacy and transitional education services (Stuart Foundation, 2011). Community reengagement success often hinges on the ways in which communities organize resources and systems (Bridgeland & Milano, 2012; Hanleybrown & Kramer, 2012). A study by the Center for Law and Social Policy (2003) found that urban centers with established youth resource delivery systems had positive effects on the uptake of community services among opportunity youth. Particularly for opportunity youth who have been incarcerated, having support in community reentry and assistance with housing and other basic needs may provide exposure to tangential education and employment supports (Allen et al., 2014; Corcoran et al., 2012; Osypuk et al., 2013; Wallace, 2014).

Research suggests that opportunity youth who succeed in connecting with their communities can become key resources for other opportunity youth trying to find their own way (Corcoran et al., 2012; Wallace, 2014). Thus, educators and others who support opportunity youth can encourage them to become role models and offer instructive advice for new cohorts of youth, including how to navigate the education system (Hanleybrown, Kania, & Kramer, 2012; Wahl et al., 2012). Supporting opportunity youth can inspire self-confidence and instill a sense of leadership in reconnected youth. More established opportunity youth can take their experience and inform their communities about best practices and areas of improvement for opportunity youth (Bird, 2013; Corcoran et al., 2012).

Research Questions

This AmeriCorps Opportunity Youth Bundled Evaluation was guided by the following overarching confirmatory research question:

> Following program participation, do opportunity youth participating in AmeriCorps programs as members significantly improve on various measures of educational and career attainment, and become more connected to their communities, as compared to matched opportunity youth who do not participate in national service programs?

This overarching question can be broken down into the following sub-questions: At follow-up, compared to matched opportunity youth who are not AmeriCorps members, do opportunity youth AmeriCorps members have:

- Increased educational enrollment and attainment?
- Increases in the number of employment seeking activities?
- Increased job attainment and retention?
- Improved positive civic engagement?
- Decreased likelihood of engagement with the criminal justice system?

Evaluation Methods

This section of the report discusses the key technical details concerning the implementation of the evaluation. It provides an overview of the evaluation timeline, evaluation design, survey development and testing, survey administration, response rates, and cleaning and analysis procedures.

Evaluation Timeline

The evaluation was conducted over a four-year period beginning in September 2014 and finishing in Spring 2018. The first year was predominately a planning year, while the second year marked the start of pre-test and some post-test data collection, and technical assistance to programs, and the third year encompassed completion of the post-test and data collection for the third and final survey as well as evaluation capacity building activities. A final fourth year was included for cleaning and analyzing the data and writing the final report. A summary of the evaluation timeline and activities is shown in Table 1 below.

Table 1. Primary Evaluation Activities by Program Year

Program Year	Primary Activities
Year 1: August 2014-August 2015	• Screening and recruiting AmeriCorps programs for the bundle • Identifying sources of comparison group individuals • Selecting and refining the quasi-experimental evaluation design • Developing an evaluation plan • Refining grantee logic models and theories of change • Developing a group logic model and theory of change • Assisting programs to review and revise their data collection systems to support the evaluation • Designing a survey instrument to collect pre-test, post-test, and follow-up data from study participants • Securing OMB clearance for study instruments • Piloting the survey and other instruments
Year 2: August 2015-August 2016	• Developing and implementing individualized recruitment plans (IRP) to support treatment and comparison group recruitment • Designing program-specific survey implementation plans (SIP) to support survey data collection at each program • Implementing the survey pre-test • Implementing the survey post-test with programs that were ready to collect post-test data • Conducting a small-scale qualitative study to understand the experiences of comparison group youth and the alternative services to which they may have access

Program Year	Primary Activities
Year 3: August 2016-August 2017	• Completing the post-test and the six-month follow-up surveys • Providing group and individual TA to support evaluation capacity building (ECB) and ongoing use of the results • Merging, cleaning and analyzing data from pre-test and post-test surveys to provide programs preliminary results based on an unmatched sample of treatment youth
Year 4: August 2017-February 2018	• Cleaning and analysis of survey data • Writing the final report

Evaluation Design

The evaluation uses a quasi-experimental design to provide rigorous estimates of the impact of AmeriCorps program participation on educational, employment, and community connection outcomes for opportunity youth. Using propensity score matching (PSM), program impacts were estimated by comparing the outcomes of program members (treatment group) to the outcomes of non-members who are observationally equivalent to program participants (comparison group). Comparison group members were drawn from two sources:

- Eligible applicants to the program who did not become members due to the limited number of positions available or because they declined to participate; and,
- Potential eligible applicants to the program who were recruited from community partners that frequently referred youth to the program.

Data collected from treatment and comparison youth were collapsed into a single dataset for the impact analysis. A dichotomous program participation indicator (1=participated in an AmeriCorps opportunity youth program as a member; 0=did not participate in an AmeriCorps program) served as the variable used to evaluate the program impact.

Propensity score matching (PSM) was used to match treatment and comparison youth within each program. PSM has been shown to be a reliable way to minimize selection bias in circumstances where random assignment is not feasible and/or ethical (Rosenbaum & Rubin, 1983; Rubin & Thomas, 1996). Furthermore, when propensity scores are correctly specified and result in groups that are equivalent on baseline characteristics, PSM can substantially reduce threats to the internal validity of the study (Stuart, 2010; Olmos & Govindasamy, 2015).

In brief, the PSM method involved the following steps:

1. Merge all study participant data along with the dichotomous program participation variable;
2. Generate the propensity scores;
3. Use propensity scores to match treatment with comparison; and
4. Test the resulting matched sample for baseline equivalency and modify the matching specifications as necessary until equivalency is attained.

Appendix G describes the propensity score matching methods and results.

Survey Development and Testing

The survey was developed collaboratively with CNCS staff and leadership from nine of the sites participating in the evaluation. Survey items were pulled from the current AmeriCorps application and exit surveys as well as from well-validated surveys such as the National Longitudinal Survey of Youth and the DC Alliance of Youth Advocates survey. Additionally, scale items came from the general self-efficacy scale (Schwarzer & Jerusalem, 1995), the competence for civic action scale (Flanagan, Syvertsen, & Stout, 2007; Kehane, Middaugh, & Schutjer-Mance, 2005), the career decisions self-efficacy scale (Betz, Klein, & Taylor, 1996), the sense of community scale (Furco, Muller, & Ammon, 1998), and the career competencies indicator scale (Francis-Smythe, Haase, Thomas, & Steele, 2013). The survey was pilot tested with 37 opportunity youth from participating sites in March 2015. After pilot testing, all scales demonstrated satisfactory internal consistency with Cronbach's alpha ranges between 0.74 and 0.90 (updated survey validation results are presented for the final analysis sample in Appendix C, Section ii).

Survey Administration

Planning and Technical Support for Survey Administration

Identifying comparison group youth. Each program site received technical assistance to help identify potential comparison group youth. This process, described in greater detail in the Individual Recruitment Plan (IRP) Memorandum, typically involved helping sites recruit comparison group youth through a combination of over-recruitment[3] of program applicants and recruitment of additional youth from community partners. Recruitment of youth from community partners involved identifying community partners that serve a similar target demographic and enlisting these partners' support with the evaluation, outlining their prospective role, responding to questions, and providing them with relevant materials (e.g., copies of the surveys, tracking sheets).

Reviewing survey administration procedures with program staff. Each program site received one-on-one technical assistance (TA) approximately one month prior to the start of each in-person survey administration to review survey procedures and to respond to staff questions. The TA entailed reviewing procedures for distributing and recording surveys and gift cards on the baseline survey tracking sheet (BSTS) or post-test survey tracking sheet, ensuring that each youth could be tracked longitudinally. Following survey administration, the JBS TA staff scheduled at least one call with each program to discuss how the survey administration process unfolded and to identify and address any potential challenges.

Group technical assistance. Throughout the evaluation, the sites also met as a group. In the planning phase, the meetings oriented programs to the evaluation, reviewed and refined an overarching logic model for the evaluation, obtained input on the proposed evaluation design, sought feedback on the draft survey instrument, and pilot tested the survey. During the implementation phase, the meetings served as a forum for evaluators and programs to problem-

[3] Over-recruitment refers to sites recruiting comparison youth from individuals in the program's applicant pool who did not join the program, but otherwise met the program's eligibility requirements and target OY characteristics. The number of applicants and the number of positions can fluctuate from year to year. Of the nine programs included in the PSM results for this report, six had at least some over-recruitment from their applicant pool.

solve around treatment and comparison group recruitment and retention, and share data collection best practices. As results became available at each stage of data collection, the meetings provided an opportunity to discuss interim results and next steps.

Details of Survey Implementation

Pre-test survey administration. Pre-test survey administration began in September 2015 and was completed in July 2016. Administration of pre-test surveys varied by program as a function of the start and stop dates of their cohorts (see Appendix D, Figure D1 for a more detailed description of the exact dates of survey administration for each program). The aim was to survey all treatment and comparison groups within one month of the program start date. Most pre-test surveys (93%) were administered as paper surveys by program staff either during pre-service orientation or during the program's application process (see Appendix H, Table H5 for a full listing of the number of surveys completed by modality and site). Staff members at partner agencies also administered paper surveys to comparison group youth. Each youth who completed the pre-test survey received a $10 Visa gift card. Program staff members recorded the survey number, gift card number, and whether each participant was in the treatment group or comparison group on the BSTS, which was then sent to JBS regularly during the pre-test data collection period. Surveys were also administered online via SurveyMonkey®. JBS mostly administered the online survey to comparison youth who had applied and not been accepted to the program or were referred by the partner agencies. Each program received weekly status updates via e-mail to track progress toward its target recruitment goals for the evaluation. Monthly status updates were also e-mailed to all sites, showing collective progress towards the evaluation's goals and status by program.

Post-test survey administration. Post-test survey administration began in March 2016 and was completed in April 2017. The timing of post-test survey administration varied across each program and coincided with each member's start and stop date (see Appendix D, Figure D1 for a more detailed description of the exact dates of survey administration for each program). The post-test surveys for treatment and comparison group youth were administered within one month of the AmeriCorps member's end of service. Youth received a $10 Visa gift card for completing the post-test survey. Most post-test surveys were completed in person as a paper survey (57%, see Appendix H, Table H5), with a smaller percentage completed online (37%), and only a very small percent over the phone (6%). At sites, the survey was usually administered during member exit interviews or during other program exiting activities. Some treatment members completed the post-test survey online or over the phone if they exited the program before a paper survey could be administered to them, or if the logistics of administering paper surveys caused undue burden to the program staff[4]. Programs that administered the post-test survey in person used a Post-test Survey Tracking Sheet (PSTS), which included the names and contact information of all eligible members who completed the pre-test survey. The PSTS also included space for the staff to record each youth's survey number for the post-test survey, gift card number given to the

[4] Sites were given the option to administer post-test surveys to members in person themselves or to have JBS conduct post-test outreach with the youth online or over the phone. The majority of treatment sites chose to administer paper surveys to youth as they were exiting the program, but three sites noted that they did not have the capacity to administer the post-test themselves and thus elected to have JBS conduct post-test outreach for them. These sites exclusively completed the post-test online or over the phone. See Table H3 for a more detailed listing of the number of surveys completed by survey wave, modality, and site.

youth, and any changes in the youth's contact information. Each program received biweekly status updates via e-mail during the period the program was administering the post-test survey.

Post-test surveys were administered online or over the phone to comparison group youth at the same time as they were administered to treatment members. JBS sent postcards to comparison group participants prior to post-test data collection, reminding them of the study's importance and about the gift card incentive for study participation, and requesting any updates to their contact information.

Follow-up survey administration. Follow-up survey administration began in September 2016 and was completed in August 2017. The timing of follow-up survey administration varied across each program and was linked to each member's start and stop dates. JBS aimed to administer all follow-up surveys for treatment and comparison group youth six months after the end of the AmeriCorps member's term of service[5]. Youth received a $20 Visa gift card for completing the follow-up survey. Most follow-up surveys were completed online (81%) except one program, which administered paper surveys at alumni events (15% of the total sample, see Appendix H, Table H5 for modality information by site). Follow-up surveys were also administered online or over the phone to comparison group youth at the same time as they were administered to treatment members. Each program site received biweekly status updates via email during the period the program site was actively administering the follow-up survey to track progress.

Procedures for collecting Participation Data. In addition to survey data, each program provided information about members' participation in various training activities (MPTA)[6], start and stop dates, target number of hours, and number of hours completed (see Appendix C for a copy of the MPTA data collection tool used with programs). The MPTA was completed by program staff

[5] Although all surveys were designed to be administered at the start and stop of the program and six months after the end of the program, there were several key challenges encountered in implementing data collection that may have impacted the timing of survey administration. In particular, sites sometimes had difficulty administering the survey within the agreed upon timeframe and occasionally administered the survey to ineligible individuals. During pretest data collection, widely varying enrollment schedules—and the fact that grantees sometimes adjusted their enrollment schedule midstream in response to difficulties with enrollment—meant that the evaluation team had to pay close attention to enrollment activity at each program. In particular, it became necessary to adjust comparison group data collection to keep it in sync with changes to member recruitment and enrollment. Also, for programs that recruited members throughout the program year ("rolling enrollment") it was necessary to identify a cutoff date for collecting pre-test data from treatment and comparison group youth in order to maintain aligned timeframes for baseline and follow-up data collection from these youth. A somewhat less common challenge arose when programs inadvertently administered the pre-test survey to individuals (either treatment or comparison) who did not fit the selection criteria for opportunity youth as defined in this study, such as being outside the specified age range or showing signs of successful educational attainment, such as having completed a bachelors' degree. The evaluation team carefully screened incoming surveys for signs of ineligibility and provided feedback to programs on a case-by-case basis to minimize these occurrences. In cases where the first batch of surveys took several weeks to reach JBS, this sometimes resulted in ineligible surveys being administered before corrective action could take place.

[6] Many programs noted that data systems were not advanced enough to comprehensively report on the specific training activities completed by each youth so results in the current report focus exclusively on the number of hours served overall relative to their target number of hours. Additionally, JBS also conducted a brief qualitative study among comparison youth to determine the extent to which youth were accessing additional services. Results from this study were presented in a separate memo and are available upon request.

after youth had completed their term of service and used program data management systems and paper records to document youth's participation information.

Response Rates

Overall Sample Sizes and Response Rates

Response rates overall were 41 percent at post-test and 33 percent at follow-up. Response rates across both time points were higher for treatment youth than comparison youth, but response rates for treatment youth declined more from post-test to follow-up than for comparison youth (see Table 2 below). This decline in the response rates may be because most treatment youth completed the post-test survey in person as they were exiting the program, but had to be contacted online or over the phone to complete the follow-up survey (See Appendix H, Table H2 and Table H5 for response rates by program, survey wave for treatment and comparison groups).

Table 2. Response Rates by Survey Wave for Treatment and Comparison Youth

Sample	Pre-test Survey # Completed	Post-test Survey # Completed	Post-test Response Rate	Follow-up Survey # Completed	Follow-up Survey Response Rate
Treatment	482	294	61%	216	45%
Comparison	590	141	24%	139	24%
Overall	**1,072**[7]	**435**[8]	**41%**	**355**[9]	**33%**

Survey Non-Response Bias

The overall follow-up response rate was 33 percent, so a survey non-response bias analysis was conducted to determine which factors (if any) were significant predictors of non-response. Overall, results indicated that treatment youth had 2.2 times greater odds of participating in the follow-up survey than comparison youth. Additionally, women, non-Hispanic whites, high school graduates, and those with no criminal history were also significantly more likely to participate in the follow-up survey. When these effects were examined separately for treatment and comparison youth, non-Hispanic whites were significantly more likely to complete the follow-up survey across both treatment and comparison youth. However, within treatment, women and high school graduates were more likely to participate in the follow-up survey, and within comparison youth, those with no criminal history were more likely to participate. An exploratory analysis within the treatment group only indicated members who completed their

[7] A total of 1,348 pre-test surveys were originally administered to youth; 1,072 (79.5%) of these were ultimately included in the study sample. A total of 276 surveys were excluded with the most common reasons for exclusion being age ineligibility (youth too old or too young for the study, N=125), enrolling in another program geared towards OY (e.g., YouthBuild, N=62), refusing the survey or not providing consent (N=22), taking the pre-test survey twice (N=20), not completing the survey or failing to provide contact information for follow-up (N=17), not actually being an OY (N=14), or never being contacted for follow-up (N=3).

[8] 436 youth completed the post-test survey and one youth was excluded for having incomplete data (more than 50% of survey items missing).

[9] 358 youth completed the follow-up survey and three youth were excluded for having incomplete data (more than 50% of survey items missing).

target number of AmeriCorps hours had five times greater odds of completing the follow-up survey than members who did not meet this target (See Appendix J).

Item Non-Response Bias

Multivariate logistic regressions predicting item non-response were to be conducted for any items with greater than 30 percent missing values. However, for the variables used in the matching process and the primary outcome analysis, non-response was low (less than 5%) for all member and comparison youth. As such, no item-level non-response analyses were needed.

Cleaning and Analyses

Data cleaning and preparation steps. All data from the paper, online, and phone surveys were entered into SurveyMonkey®, then cleaned and analyzed using SPSS and R. Aggregate variables were also created to indicate sum and mean scores for each survey scale and subscale. Data cleaning also included constructing several recoded versions of survey items and excluding any participants with more than 50 percent of survey questions unanswered. Cleaned data files for each program were then merged into a single aggregate data file with both treatment and comparison youth at each time point. Aggregate files from the pre-test, post-test, and follow-up surveys were then merged for analysis. Raw data files, cleaning scripts, and final cleaned datasets for each program, and datasets in aggregate were submitted to CNCS.

Overview of analyses. Analyses consisted of descriptive statistics of pre-test and follow-up data, propensity score matching, confirmatory intent to treat analyses with the matched sample, and exploratory treatment on treated analyses with a subset of the matched sample and with the full treatment only sample.

> *Descriptive analysis.* The descriptive analysis included the presentation of the mean and standard deviation for each of the outcome variables at pre-test and follow-up for treatment and comparison youth separately. Descriptive results also included the presentation of the change from pre-test to follow-up within each group and the difference in the percentage change between treatment and comparison youth.
>
> *Propensity score matching (PSM).* PSM methods were used to create a sample of comparison youth with characteristics that closely corresponded to those of AmeriCorps program participants (treatment youth). The comparison youth included both applicants to each program and potential applicants who were receiving services from the program's referral partner. The PSM was implemented using three steps:
>
> > *Step 1: Merge data*. Data from each program were aggregated into a master pre-test data file including all available characteristics and outcomes of participants and non-participants. A dichotomous program participant variable was included to indicate whether the participant was an AmeriCorps member identified as an opportunity youth (with a score of 1) or was in the comparison group (with a score of 0).
> >
> > *Step 2: Impute missing values*. To conduct propensity score matching, it is necessary to have complete data available for all matching variables. To do this, the aggregated dataset for all youth who had completed the follow-up survey was

used to perform multiple imputation on any cases with missing values on the matching variables. Overall, the amount of missing data was extremely low with only three variables being imputed (gender, high school graduate, and criminal history) and only two to three cases being imputed for each of these variables. This imputation process was performed five times and then the imputed datasets were averaged to create a single dataset.

Step 3: Produce propensity score. Using the average imputed dataset a multivariable logistic regression model was then used to estimate the likelihood of program participation based on available baseline control variables. The baseline variables used in the logistics regression were:

- Number of months from the start of the evaluation to pre-test survey (study timing);
- Number of days (study duration) between pre-test survey (survey 1) and follow-up survey (survey 3);
- Non-Hispanic White (vs. not);
- Gender (female vs. male);
- Disability (limited due to physical, mental or emotional problems vs. not);
- Parent or primary caregiver of a child;
- Primary caregiver of a parent or other adult;
- Served on active duty in the military;
- High school graduate or equivalent;
- Employed in the last six months;
- Criminal history (convicted of a crime or currently facing charges or probation or parole); and
- Received any state or federal government support (e.g., – food assistance, housing assistance, healthcare assistance, or other financial or practical support).

The coefficient output from the regression was then used to produce a propensity score for each member and comparison youth (non-member) in the imputed dataset (see Appendix G for a more detailed description of the process for creating propensity scores and the results of matching).

Step 4: Use propensity score to match members with non-members. Using the propensity scores, we conducted 2:1 treatment to comparison nearest neighbor matching within each program with a caliper of 0.5 standard deviations and no replacement. This method allowed us to match participant youth to the comparison youth that had an identical or nearly identical propensity score.[10]

[10] We compared different matching methods including 1:1 and 2:1 nearest neighbor with a 0.25 standard deviation caliper and a 0.5 standard deviation caliper, 1:1 and 2:1 nearest neighbor matching with dummy coding for sites, nearest neighbor matching with replacement, and full matching. Full matching and 2:1 matching using the program dummy codes with a 0.5 standard deviation caliper did not completely erase baseline differences, so these methods

Step 6: Test for baseline equivalence. Once matching was completed, we tested whether members and comparison youth shared similar baseline characteristics before and after the match. To do this, mean standardized differences and Mann-Whitney or Chi-squared tests were used to compare treatment and comparison youth on each of the matching variables. If matching was successful, any baseline differences in the matching variables would be eliminated after the match.

Confirmatory ITT Impact Analyses. Multi-level multivariate mixed effects linear and logistic regression models were used to evaluate the impact of program participation on educational, employment, and civic engagement outcomes with the matched sample[11]. Since results included all treatment and comparison youth regardless of program completion, this analysis can be considered an Intent to Treat (ITT) analysis. ITT models were run for each of the nine confirmatory outcome variables including high school graduation status, course completion, college or trade school enrollment, current employment status, hourly wage, longest period of employment, number of job search activities completed, current involvement with the criminal justice system (facing charges or being on probation or parole), and level of civic engagement. Models for each outcome included fixed effects for treatment (1=Treatment, 0=Comparison) and baseline values of demographic variables (criminal history, employed in the last six months, age, gender (1=Female), parent or primary caregiver of a child, ethnicity (Non-Hispanic White=1), high school graduation status), the number of days between pre-test and follow-up, and the pre-test value of the outcome variable when relevant. Fixed effects also included specific interactions for each outcome area by relevant matching variables-- namely interactions between treatment and high school graduation status (for education), treatment and employment history (for employment), and treatment and criminal history (for civic engagement). Additionally, models also included random intercept effects for each program and random slope effects to model the impact of treatment within program.[12]

Exploratory TOT Analyses. In addition to the confirmatory analyses with the matched sample, three additional exploratory analyses were conducted to assess the impact of treatment on the treated. First, a subset of matched sample was created consisting only of

---

were rejected. We then decided that exact matching on program was necessary to reduce the influence of any unmeasured differences between sites so dummy coding of program was not used. Matching with replacement was also discarded because several comparison youth appeared to be dramatically over-weighted and were matched with many treatment youth. Ultimately, the 2:1 matching with a 0.5 caliper was chosen because it was the method resulting in the largest observed sample size of the remaining methods.

[11] All models tested in R 3.4.2 using the "lme4" package. Models predicting dichotomous outcomes used the "glmer" function which uses a generalized linear mixed model fit by maximum likelihood (Laplace approximation) to predict outcomes from a binomial (logit) family. Within this function we also included the bobyqa optimizer from the "minqa" package to implement derivative-free optimization as advocated by Powell (2009). All models for continuous outcomes were tested using the "lmer" function from the "lme4" package. This function uses a linear mixed model fit by restricted maximum likelihood to predict outcomes (e.g.- hourly wage).

[12] Although all models were run as multi-level models, one model did not fully converge (the model predicting high school graduation rates at follow-up) and so results for that model are based on a logistic regression model predicting high school graduation status at post-test among youth who were not high school graduates at pre-test. Results from this model should not be compared with the results from the other models given that is not a multi-level model.

the treatment youth who met their target number of AmeriCorps hours and the matched comparison group for those youth. Using this sub-sample, multi-level multivariate mixed effects linear and logistic regressions were used to predict differences among youth who had completed treatment and their matched comparison group (see Appendix G3). Second, linear and logistic regressions were used with the full unmatched treatment-only dataset to examine dosage effects based on program completion (defined as meeting the target number of hours) and the total number of hours served (see Appendix G4). Finally, paired samples Wilcoxon and McNemar tests were used to determine if there were statistically significant differences from pre-test to post-test and from pre-test to follow-up within program participants only (see Appendix G5).

Overall Findings

Presentation of Primary Outcome Results: Intent to Treat Analysis

Education Results

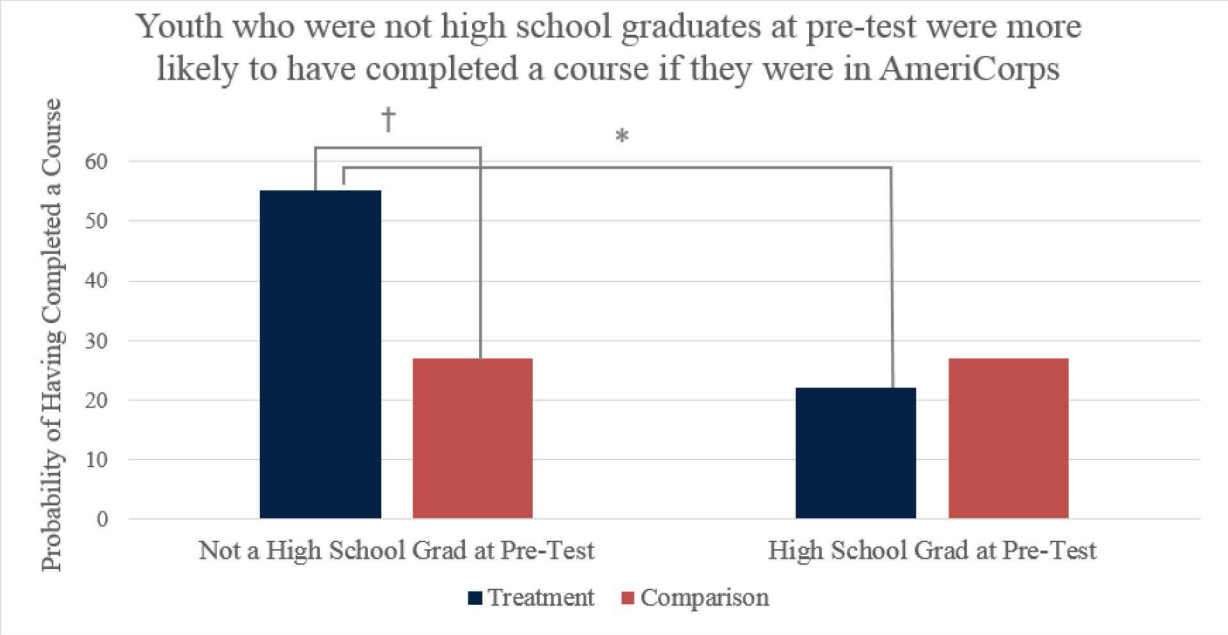
Within the matched sample, both treatment and comparison youth showed improvements in educational attainment from pre-test to follow-up with the percentage of high school graduates increasing by 13 percent within the treatment and 16 percent within the comparison group. Youth in both groups were less likely at follow-up to have completed a course within the last six months, but slightly more likely to have enrolled in college or trade school. Appendix G1 provides more details of the changes from pre-test to follow-up for the treatment and comparison youth in the final matched sample.

Multi-level multivariate mixed effects logistic regression models were used to predict educational outcomes at follow-up for treatment and comparison youth[13]. The results showed there were no statistically significant main effects of treatment (see Appendix G, Table G2). However, there was a marginally significant main effect of treatment on course completion such that youth who participated in AmeriCorps were more likely to have recently completed a course at follow-up than comparison youth (*b*=1.21, *SE*=0.70, *OR*=3.35, *p*=0.08)[14]. There was also a marginally significant interaction between high school graduation status at pre-test and program participation on course completion (*b*=-1.48, *SE*=0.80, *p*=0.07, *OR*=0.23; see Figure 1 below). Post-hoc pairwise comparisons indicated that youth who had high school diploma at pre-test were less likely to complete a course if they were in the comparison group (*OR*=0.30, *SE*=0.21, *p*=0.08) [15]. Additionally, treatment youth who did not have a high school diploma at pre-test were more likely to have recently completed a course than comparison youth who did not have a diploma (*OR*=4.24, *SE*=2.77, *p*=0.03). However, there were no statistically significant

[13] For more detail on the exact models tested please see the Cleaning and Analyses section of this report and Appendix G: Impact Analyses and Results. Also, again please note that a logistic regression model was used to predict high school graduation status at post-test among youth who were not high school graduates at pre-test after the MLM would not converge.
[14] Throughout the current report we use the value "*b*" to represent the unstandardized coefficients from each model.
[15] Post-hoc pairwise comparisons were conducted in R 3.4.2 using the lsmeans package to compare the effects of treatment by employment history and the effects of employment history by treatment within the context of the mixed effects model. Estimates, standard errors, and p-values were based on Tukey's HSD test results.

differences in course completion within comparison youth or within youth who had a high school diploma or GED at pre-test.

Figure 1. Predicted Probability of Having Completed a Course[16]



Employment Results

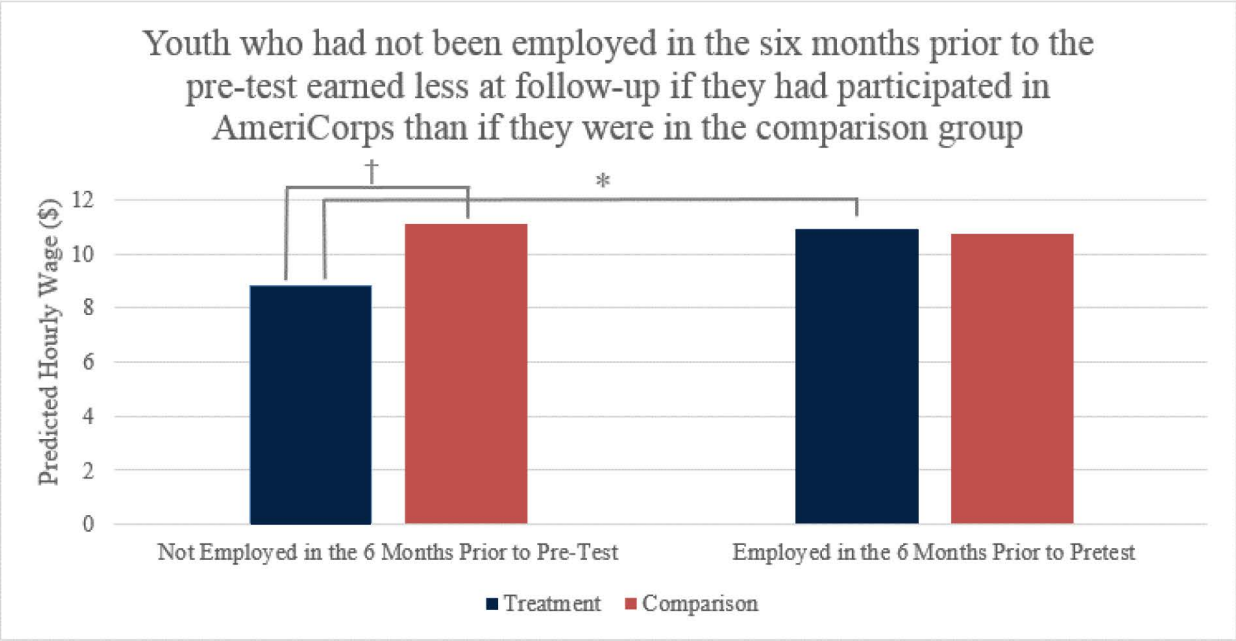
Within the matched sample, both treatment and comparison youth showed improvements in current employment from pre-test to follow-up; the percentage of treatment youth who were employed increased by 54 percent, while the percentage of currently employed comparison youth increased by 20 percent. Additionally, youth in both groups showed slight improvements in their hourly wage (increasingly roughly $1-$1.50/per hour) and longest employment at follow-up (number of months employed increased by about 2.5 months). However, youth in both groups showed very small declines in the number of job search activities they had completed in the last six months (decreasing by 0.4 job search activities). Appendix G1 shows the changes from pre-test to follow-up for the treatment and comparison youth in the final matched sample.

Multi-level multivariate logistic and linear regressions were used to predict employment outcomes at follow-up for treatment and comparison youth. The results showed there were no statistically significant main effects of treatment, although there was one marginally significant main effect of treatment on hourly wage with treatment youth earning slightly less per hour than comparison youth ($10.88 vs. $10.72, b= -2.31, SE=1.21, p=.06) (see Appendix G, Table G3). There was also a marginally significant interaction between employment history (employed in the last 6 months) at pre-test and program participation on hourly wage (b=2.47, SE=1.35, p=0.07). Pairwise comparisons indicated that youth who had not been employed in the six months prior to the pre-test earned less at follow-up if they were in the treatment condition

[16] Please note that all figures presented here are based on the estimated marginal means from the mixed effects models. Significance values are from the results of Tukey HSD post-hoc tests. † = p<.10, * = p<.05, ** = p<.01, *** = p<.001.

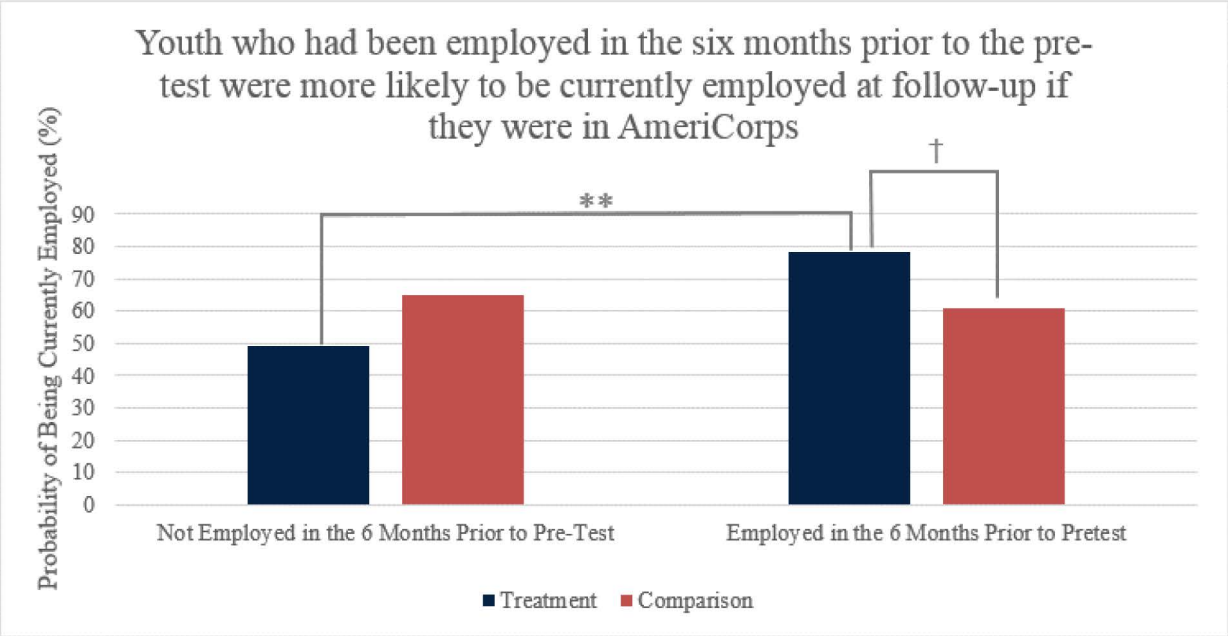
instead of the comparison condition (*b*=-2.31, *SE*=1.21, *p*=0.06; see Figure 2 below). Additionally, within the treatment condition, youth who had been employed in the six months prior to pre-test had higher wages at follow-up than treatment youth who had not been employed (*b*=2.09, *SE*=0.89, *p*=0.02). There were no statistically significant differences in wages among youth in the comparison condition or youth who had been employed in the six months prior to pre-test.

Figure 2. Predicted Hourly Wage



There was also a significant interaction effect of treatment and employment in the six months prior to pre-test on current employment (*b*=1.44, *SE*=0.69, *p*=0.04; see Figure 3 below). Pairwise comparison indicated that youth who had been employed in the six months prior to pre-test were less likely to be currently employed if they were in the comparison group rather than the treatment group (*OR*=0.45, *SE*=0.19, *p*=0.06), but there was no statistically significant difference in current employment by condition within youth who had not been employed in the six months prior to the pre-test. Additionally, AmeriCorps members who had not been employed at any point in the six months prior to the pre-test were less likely to be currently employed at follow-up than youth who had been working (*OR*=0.27, *SE*=0.13, *p*=0.005). There was no statistically significant difference in the probability of current employment at follow-up based on previous employment history among comparison youth.

Figure 3. Predicted Probability of Being Currently Employed

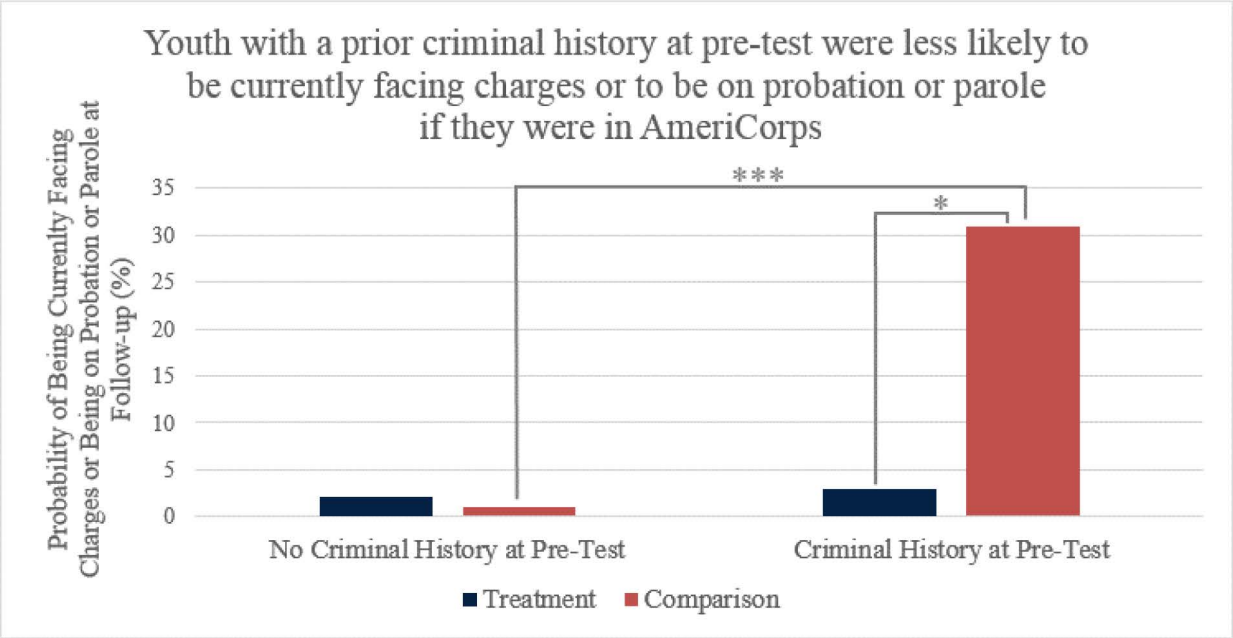


Civic Engagement Results

Within the matched sample, there was a five percent decline in the number of treatment youth who were currently facing charges or on probation and parole, while there was a three percent increase among the comparison group. Although treatment youth reported higher levels of civic engagement activities overall at both pre-test and follow-up, their self-reported frequency of civic engagement declined slightly from pre-test to follow-up, while comparison youth increased slightly. See Appendix G1 for a more detailed listing of the changes from pre-test to follow-up for the matched treatment and comparison sample.

The results from multi-level multivariate logistic and linear regressions showed no statistically significant main effects of treatment for civic engagement outcomes at follow-up (see Table G4 in Appendix G). However, there was a statistically significant interaction between criminal history at pre-test and program participation on the likelihood of currently facing charges or being on probation or parole (*b*=-3.06, *SE*=1.51, *p*=0.04, *OR*=0.05). Post-hoc pairwise comparisons showed that youth who had a criminal history at pre-test were more likely to be currently facing charges or to be on probation or parole at follow-up if they were in the comparison condition instead of the treatment condition (*OR*=14.04, *SE*=16.92, *p*=0.03). Additionally, within the comparison group, youth with no prior criminal history at pre-test were less likely to be currently facing charges or to be on probation or parole at follow-up than youth with a criminal history at pre-test (*OR*=0.02, *SE*=0.03, *p*=0.0006). However, there were no significant differences within treatment youth, or among treatment and comparison youth who did not have a criminal history.

Figure 4. Predicted Probability of Currently Facing Charges or Being on Probation or Parole



Presentation of Exploratory Outcome Results: Treatment on Treated

Differences from Pre-test to Post-test for the Full Unmatched Treatment Only Sample. Exploratory analyses looking at paired differences between the pre-test and post-test surveys within program participants showed promising improvements in many of the key outcome domains (see Appendix G a full listing of results). In particular, statistically significant improvements were seen in the number of youth who had a high school diploma or equivalent, completed a course, and were currently employed. Additionally, youth reported significantly longer periods of employment and more frequent civic engagement.

Differences from Pre-test to Follow-up for the Full Unmatched Treatment Only Sample. Exploratory analyses looking at paired differences between the pre-test and follow-up surveys within program participants showed additional improvements in many of the key outcome domains (see Appendix G for a full listing of results). In particular, statistically significant improvements were seen in the number of youth who had a high school diploma or equivalent, completed a course, and were currently employed. Additionally, youth reported significantly higher wages, longer periods of employment, and marginally higher numbers of job search activities completed[17].

[17] Appendix G also includes the results of two additional exploratory treatment on treated analyses including one analysis using a subset of the matched sample and another looking at dosage effects within the full treatment sample. Given that most treatment youth in the final sample had completed their term of service, results from the subsetted analysis of the matched sample closely resembled those from the Intent to Treat analysis and were thus moved to an appendix. Results from the dosage analysis indicated that there was not a statistically significant impact of dosage on follow-up outcomes.

Summary of Key Evaluation Results

On average, opportunity youth who participated in the AmeriCorps program were neither more nor less likely to experience increases in educational enrollment and attainment, employment, or civic engagement at follow-up than matched comparison youth. However, participation in an AmeriCorps program did show positive impacts for youth with certain characteristics prior to the start of their service with AmeriCorps. For example, youth who had a prior criminal history at baseline were significantly less likely to be currently facing charges or to be on probation or parole if they were in the treatment group. Additionally, treatment youth who had been employed at some point in the six months leading up to their AmeriCorps participation were more likely to be employed at follow-up than comparison youth with a similar employment history. Furthermore, exploratory analyses within the unmatched treatment only group suggested that program participation significantly improved high school completion, course completion, hourly wage, and employment experience. Together these results suggest that although treatment youth did not necessarily improve compared to comparison youth, they improved over time, and treatment may be especially effective for opportunity youth with certain characteristics such as prior criminal history.

Evaluation Limitations

Overall, the results provide promising moderate evidence of the effectiveness of AmeriCorps programming for certain populations of OY. However, the evaluation did not find statistically significant results across all the predicted outcomes and populations. This lack of significant findings suggests that either the current programming is not having the predicted impacts on these populations, or that limitations of the current evaluation may be impairing our ability to detect impacts that may be occurring. In particular, the limitations of the propensity score matching design, attrition and non-response bias, restricted sample sizes, and unmeasured differences across sites could indicate that future research is needed to unpack the impacts of AmeriCorps participation within opportunity youth.

Limitations of the Propensity Score Matching Design. When propensity scores are correctly specified and result in groups that are equivalent on baseline characteristics, PSM can substantially reduce threats to the internal validity of the study (Stuart, 2010; Olmos & Govindasamy, 2015). However, it is often difficult to determine if all relevant differences between treatment and comparison youth have been measured and correctly specified. In the current study, although twelve covariates were used in the calculation of the propensity score, additional unmeasured covariates could contribute to selection bias[18]. As such, even if baseline equivalency is obtained on the observed covariates, without the use of a randomized design it's always possible that unmeasured factors could influence both selection into treatment and outcomes. As just one example, although the current evaluation included most major demographic characteristics in the matching procedure, it's possible that additional individual-level psychological variables such as grit (defined as trait-level perseverance and passion for long-term goals; see Duckworth & Quinn, 2009) or conscientiousness could also be impacting

[18] Although unmeasured covariates are always a potential threat to the validity of any PSM design, the current study selected the core set of matching variables based on the literature on opportunity youth and discussions with program staff in advance of the analysis. Covariates were selected to ensure that most major differences between treatment and comparison youth were controlled for in the matching and then again in the analysis.

results. To address this, it would be useful for future designs to consider utilizing additional administrative and other existing data sources to include additional matching variables[19] and to assess the feasibility of implementing a randomized control trial.

Attrition and Non-response bias. The original evaluation plan had anticipated 20 percent attrition over the course of the study, however attrition rates were substantially higher than originally expected (roughly 66 percent at follow-up) due to challenges in maintaining contact with the youth. Given the high levels of attrition, a non-response bias analysis was conducted to determine what factors (if any) contributed to non-response (see Appendix E). Results indicated that there was differential attrition, with treatment youth being more likely to complete the follow-up survey than comparison youth. Additionally, within the treatment condition, youth who successfully completed their target number of service hours were also more likely to have completed the follow-up survey (making the original intent to treat analysis more closely resemble a treatment on treated analysis). Finally, women, whites, high school graduates, and youth with no prior criminal history were also more likely to complete the follow-up. To address these differences, gender, race/ethnicity, education status, and criminal history were all included as covariates in the final impact analysis. However, it is possible that even with these controls non-response bias could still be impacting the results. In future evaluations it would be helpful to take additional steps to ensure that comparison youth and program dropouts are retained in the study sample. For example, it may be helpful to consider providing higher incentives to youth and having more frequent contact with them between survey administration.

Restricted Sample Size. At the start of the evaluation it was anticipated that each program site would have an average of 32 youth (16 treatment and 16 comparison) in the final sample, for a total of 640 youth across 19 sites. With this sample size it was anticipated that the evaluation would be able to detect a standardized effect size in the range of 0.24-0.36 (a small to medium sized effect according to Cohen, 1988)[20]. However, the final sample sizes for the PSM sample ranged from 2 to 40 youth per site, with a harmonic mean of approximately 7 youth per site[21]. Post-hoc power analyses revealed that with this sample size the minimum detectable effect size was approximately 0.49 (a medium effect size)[22]. These results suggest that the study may have been underpowered (particularly within smaller sites) to detect the smaller effects often shown in youth development interventions. Additional post-hoc power analyses revealed that to detect a small effect size (for example d=0.2) the study would need to have about 40 youth per site (ideally 20 treatment and 20 comparison). Furthermore, since the power of the study depends on the average sample size in each site it would be helpful for future bundled evaluations to ensure

[19] We recommend the use of administrative or existing data sources in lieu of additional primary data collection as surveys were already very long and should be reduced in length to increase youth participation.

[20] More information on the original power estimates are included in the Part B of the OMB package and can be accessed here: https://www.reginfo.gov/public/do/PRAViewDocument?ref_nbr=201504-3045-001

[21] The harmonic mean is the reciprocal of the arithmetic mean of the given set of observations and is calculated by dividing the total number of groups by the sum of the inverses of the sample size in each group. It is typically lower than a traditional arithmetic mean and represents a more conservative estimate of the sample size in each group. For this reason, the harmonic mean is a more common and more sensitive estimate to use for power calculations.

[22] Post-hoc power analyses were conducted using Dong and Maynard's (2013) PowerUp! Tool using the MDES calculator for a 2-level fixed effects blocked individual random assignment design (according to the authors propensity score matched designs can be analyzes in the same way as a RCT design). Information from the mixed effects model for income was used to calculate the MDES and included assumptions of a two-tailed test with alpha =0.05, power = 0.80, P=0.58, R_1^2= 0.19, 11 level one covariates, average block size = 7, and 15 blocks (sites).

that sites all have the capacity to attain at least some minimum threshold sample size across both their treatment and comparison groups. This is especially important since unequal sample sizes between treatment and comparison youth within a site can dramatically limit that site's ability to be represented in the final sample. For example, although Site 4 had 37 members complete the follow-up survey they only had 2 comparison youth who completed the follow-up. Therefore, even with a 2:1 matching of treatment to comparison youth only 6 youth from Site 4 could be included in the final matched sample. These results suggest that it may be helpful for future bundled evaluations to have a minimum sample size for each participating treatment and comparison site and to ensure that comparison sites have the ability to maintain contact with youth throughout the duration of the study.

Unmeasured differences across sites. Finally, to control for differences across sites the current study matched youth within each site and then included random effects of site and interactions between site and treatment in the mixed effects model. However, it is still possible that unmeasured differences between sites could affect the results, especially since sample sizes were larger at some sites than others. In future bundled evaluations it would be helpful to consider quantifying differences between sites (e.g., program size, number of years working with OY, urban vs. rural, annual funding amounts) and then controlling for site-level characteristics in the analysis as level 2 covariates. Additionally, it may be helpful to focus future evaluations on sites that are very similar not only in terms of their program focus and population, but also in their program size.

Recommendations for ASN programs serving OY and for CNCS

The evaluation's findings indicate that ASN programs may benefit from targeting youth with specific baseline characteristics. ASN programs serving OY may be particularly effective in preventing recidivism among youth with a prior criminal history. Young adults in the U.S. (aged 18 to 24) are responsible for a disproportionately high percentage of crime, have increased rates of recidivism compared to other age groups, and may be particularly underserved by programs that are focused on juvenile offenders or tailored to older adults.[23] Involvement with the criminal justice system may have long term implications for young adults' future employment and other outcomes. Therefore, there are significant individual and societal benefits for an intervention that demonstrates some success in preventing recidivism for this population. Youth with a criminal history who completed more hours of treatment were also slightly more likely to be engaged in their communities, indicating that it may be worthwhile to recommend increased hours for members of this population.

Additional findings indicate that ASN programs may have increased benefits in maintaining school involvement among youth who have not completed high school, and in promoting employment among those who have been recently employed. While the latter finding is contrary to expectation, it does not necessarily indicate that youth without a recent employment history do not benefit from the program in other ways. For example, interaction effects between educational and employment characteristics with treatment condition were not included in this model, and it is conceivable, especially given the findings above, that youth without a high school degree and no recent employment may be more likely to be engaged in school or in other programs that have

[23] *Reducing Recidivism and Improving Other Outcomes for Young Adults in the Juvenile and Adult Criminal Justice System* (2015). Retrieved from: https://csgjusticecenter.org/wp-content/uploads/2015/11/Transitional-Age-Brief.pdf

long-term benefits at the expense of short term gains in employment outcomes. Nevertheless, the findings indicate that programs with a primary focus on increasing near term employment outcomes may see greater gains in maintaining workforce involvement among those with some employment experience, while programs with a primary focus on education could benefit youth without a high school diploma.

Key Lessons Learned about the Bundled Evaluation Methodology

The following additional lessons and recommendations emerge from the bundled OY evaluation. The list begins with advice pertaining to methodology and study implementation followed by items pertaining to the capacity of the participating programs.

1. **Examine seemingly similar programs more closely during the screening process to understand how they may differ and to ensure a good fit with the bundle.** While OY programs are similar in many ways, arguably each program in the bundle is "unique" in one way or another. While these differences did not render the bundle nonviable, they do point to gaining a more detailed understanding of the OY population the program recruits, how each program works in practice, and its intended outcomes. For example, one subtle difference between the programs in the bundle may be between those that focus exclusively on serving OY versus those programs with an OY component as an add-on to an existing AmeriCorps program that has historically served a more conventional ASN population.

2. **Provide adequate survey incentives to promote participant recruitment and retention.** Programs consistently stated that the gift card incentive provided to participants for completing each survey was essential to successful data collection. We also observed a modest uptick in response rates for the final survey (including from participants that had not participated in the second survey) when the value of the incentive increased from $10 to $20. This suggests that even small increases in the value of the incentive can boost retention.

3. **Provide monetary incentives to community partners.** Programs were crucially dependent upon the goodwill of community partners to find, recruit, and retain a substantial portion of OY in the comparison group. Yet, community partners lacked any meaningful near-term incentive to divert scarce staff time and resources to such tasks. Future bundled evaluations could benefit greatly by providing a monetary incentive to community partners as a lure to participation and to partially defray the costs of their participation. The need to incentivize programs is already accounted for by the bundled evaluation design since it gives programs the opportunity to participate in an impact evaluation at little or no cost to themselves while helping them build evaluation capacity.

4. **Engage the programs in the bundle in frequent, meaningful communication.** Programs highly valued frequent communication from the evaluation team to feel engaged in a multi-year effort and to maintain a sense of progress over time. Frequent communication was also useful for keeping programs on schedule with key data collection tasks and to detect and address challenges early on. Program staff felt frequent communication also compensated somewhat for the challenges and burdens of participation.

5. **Provide the programs concise versions of the IRP and SIP.** The Individualized Recruitment Plan (IRP) and Survey Implementation Plan (SIP) proved to be useful tools for planning and implementing a bundled evaluation. Evaluators purposely developed these program-specific documents in considerable detail with a view to their possible reuse in future evaluations. At the same time, programs may have found concise versions of the IRP and SIP more user-friendly and actionable.

6. **Adhere tightly to survey administration windows.** As data collection progressed at each program, evaluators sometimes found it practical or even necessary to allow some flexibility in how long a program could continue administering the survey. However, for better consistency in the data collected across programs, and ultimately higher validity, it would be advisable to follow more strict timelines for survey administration. Ensuring greater consistency in the selection of programs for a bundled evaluation—including regarding program timelines—would also facilitate tighter adherence to survey administration windows.

7. **Revise the Member Participation in Training Activities (MPTA) document to provide more precise operational definitions of data elements.** The MPTA held promise as a tool for collecting administrative data from the programs to shed more light on variation in program implementation across the bundle. The value of data collected via the MPTA could be enhanced by providing more precise operational definitions of the data elements to ensure comparability of information across programs.

8. **Continue to develop planning pieces in a highly collaborative manner.** The collaborative processes for planning the evaluation appear to have worked well, including developing the theory of change and logic model and developing the survey instrument. Collaboration in developing the theory of change resulted in stronger and more complete articulations of these conceptual pieces while building buy-in to the evaluation at the start. Developing the instrument collaboratively proved essential to ensure the instrument would work for the respondent population. It also gave programs an opportunity to invest in a tool for which they would bear the main burden to implement in the first two waves of data collection. The added value of a collaborative approach in these tasks suggests it is highly advisable to continue doing them in a highly collaborative fashion in future bundled evaluations. An emphasis on collaboration also extends to working with more than one point-of-contact at each program so that inevitable staff turnover does not pose a threat to the viability of a longitudinal study.

9. **Provide programs with a clear picture of level-of-effort requirements for participating in a bundled evaluation at the outset.** Programs consistently reported that participating in the evaluation was more time-intensive than they anticipated, although the program understood the relevance of evaluation tasks and did not consider them to be "busy work". Still, to help programs decide if they can make the required time commitment for an evaluation, and to facilitate transparency and trust at the outset of the process, it is wise to give programs a clear picture of time level-of-effort requirements. This may even take the form of providing programs with an estimate of the typical weekly staff hours expected at each phase of the evaluation.

10. **Anticipate and plan for challenges with participant recruitment and retention challenges.** Especially when dealing with marginalized populations like OY, it is wise to build infrastructure at the local level to promote comparison group recruitment and retention. Cost savings stemming from conducting an evaluation remotely do not always translate into efficiency. One option would be to recruit a local staff person or perhaps an ASN OY member or recent program alumnus, to serve as a "retention liaison" (preferably with a modest stipend) to engage in regular outreach to comparison group youth to remind them about the study and keep their contact information up to date. An on-site presence with an AmeriCorps member or temporary hire to support the evaluation might address concerns related to conducting the evaluation remotely. This individual could advertise the study to partner agencies, collect contact information from participants, and collect data, essentially undertaking evaluation processes that partner organizations do not have the time or resources to conduct. If this is not feasible, then another option is to plan on recruiting more comparison group youth to offset anticipated high attrition.

11. **Provide programs more intensive support for recruiting and screening study participants (treatment and comparison)**. During the early phases of the evaluation the programs, supported with TA, took the lead in recruiting members and enrolling them in the study, and identifying and utilizing internal and external sources for comparison group recruitment. This critical activity also included screening enrollees for eligibility to participate in the study. Programs were largely successful in these activities, in part due to the provision of intensive TA to support these efforts. However, most programs faced challenges in meeting their recruitment targets, particularly for comparison group members. In addition, there were instances in which inadequate screening led to ineligible youth completing the pre-test survey who had to be removed from the analysis. Future bundled evaluations would benefit from greater anticipation of these challenges and closer attention to them during the early phases of the study. In particular, the process of developing recruitment plans (the IRP) collaboratively with programs could benefit from extensive discussion of member recruitment procedures (including recruitment sources and outreach modalities) and address recurrent challenges in this area of program operations. Regarding proper screening, evaluators can be more vigilant during the first several weeks of recruitment and work with programs to continually reinforce understanding about the eligibility criteria (i.e., what they are and why they are important) and to monitor incoming data, detect ineligible participants early on, and provide immediate feedback to programs.

12. **Account for potentially complex mix of member recruitment patterns and terms of service when working with bundled evaluations focusing on AmeriCorps members**. The group of programs in the bundled exhibited wide variety in their schedules for member recruitment and in the terms of service for which members signed up (e.g.- full-time, half-time). Some programs recruited all members at once during a relatively brief recruitment window, while others recruited multiple cohorts of members during the program year, or even recruited continuously throughout the year (rolling enrollment). Program often contained a mix of members with different terms of service, which may have implications for the "intensity" of exposure to the treatment. (Consider, for example, two members serving for a full year while some serve full-time while others serve half-time.) In recruiting programs for an evaluation bundle it is worth considering

how individual programs differ on these factors and looking for opportunities to assemble a bundle that is more homogeneous regarding these factors. Alternatively, evaluators need to capture relatively detailed data on these characteristics at the site and individual levels to take them into account during the analysis.

Recommendations for Improvement for Future Bundling Projects

This section offers recommendations to inform future bundling projects.

***Recommendation #1*: Future bundled evaluations can benefit from more intensive efforts to maintain contact with treatment and comparison group members between surveys.** It may also be beneficial to offer small incentive payments to study participants who respond to requests to verify or update their contact information during these intervals, and to offer incentives to community partner agencies to enlist their help in maintaining a connection with comparison group members in their program. Retention depends upon keeping study participants connected to the study through ongoing contact and frequent use of incentives. Retention can be improved by enlisting the ongoing cooperation of local entities (including not only AmeriCorps programs, but also their community partners) and by staying in direct contact with study participants. Additional steps that can be taken to bolster retention include integrating surveys into the exit process well in advance of the typical exit period, and handing members who exit early a postcard containing a brief message inviting them to participate in upcoming surveys and reminding them about any incentive payment associated with said participation.

***Recommendation #2:* Future bundled evaluations can benefit from working with sites prior to study implementation to help them build or improve systems for remaining in contact with recent program graduates.** This would not only boost efforts to retain study participants during the post-service follow-up phase, but would also generate organizational infrastructure that has lasting value for programs.

***Recommendation #3:* Future bundled evaluations should look for ways to increase sample size without excluding the smaller programs for which the bundled approach is specifically designed.** The discussion of evaluation limitations notes that substantial attrition of study participants (on the order of 66 percent at follow-up) negatively affected the PSM analysis by restricting sample size. In addition to more aggressive efforts to reduce attrition, moderately larger sample sizes would facilitate the detection of more subtle program impacts (operationalized in terms of smaller effect sizes). This seems particularly appropriate in the context of the AmeriCorps OY evaluation bundle given anecdotal evidence that differences between AmeriCorps programming and treatment-as-usual as experienced by comparison group members are sometimes minor and thus the evaluation may be trying to detect very small effect sizes. The recommendation to increase sample size must be counterbalanced by the attendant requirement to recruit an even larger comparison group in light of the decreased retention observed among comparison youth. This recommendation would entail greater reliance on larger program sites in the bundle and increased sample size commitments from comparison sites. For example, given the attrition rates observed in the present study, in order to retain at least 40 matched participants per site (i.e., 20 treatment and 20 comparison group members), it would have been necessary to favor AmeriCorps programs that typically recruit at least 60 members for inclusion in the bundle. That said, as long as the average sample size remains high and sites are operating similar program models with similar populations, it may still be possible to include a

combination of larger and smaller programs in the bundle to achieve the desired final sample size.

***Recommendation #4*: Future bundled evaluations should treat the inclusion of a "backup" staff member as a prerequisite for participation in the evaluation.** The backup person need not be actively engaged in TA, although this may be preferable whenever possible. However, as the experience of some programs shows, simply ensuring that the backup staff person is regularly included in TA communication means they can be quickly "brought up to speed" if it becomes necessary for them to serve as the program's designated point of contact for the evaluation.

***Recommendation #5:* Future bundled evaluations should continue to look for ways to streamline information requests to bundled sites and other community-level organizations while offering flexible deadlines.** Sites appreciated steps JBS took to streamline information requests, including prepopulating the MPTA with available data. Sites also appreciated being given generous lead time and flexible deadlines to fulfill information requests. This may involve asking programs for information well before it is needed to build in deadline flexibility.

***Recommendation #6:* Future bundled evaluations should continue to promote the use of paper surveys administered in person by program staff members.** Experience through all three phases of data collection demonstrates higher response rates for paper surveys than for online or phone surveys. As a matter of necessity, the online survey was the primary modality for most data collection during the follow-up survey phase since most study participants were no longer accessible to sites and community partners. However, even in the few instances when multiple survey modalities were available to a program conducting the follow-up survey, the paper survey often proved most efficacious.

Appendix A: Evaluation Report Handout

Evaluation Report Brief

AmeriCorps Opportunity Youth Evaluation Bundling Project

What are the goals of the program?

According to recent estimates there are more than 6.7 million youth between ages of 16 and 24 who are low-income and either homeless, in foster care, involved in the juvenile justice system, unemployed, or not enrolled in or at risk of dropping out of an educational institution. AmeriCorps has recently begun recruiting these "opportunity youth" (OY) as members and engaging them in national service projects designed to improve local education, public safety, health, and the environment. These OY programs provide services for local communities, and aim to advance the education, employment, and civic engagement of the youth who participate.

Program At-a-Glance

CNCS Program: AmeriCorps State and National

Focus Population: Opportunity Youth

Focus Areas: Youth Development and Economic Opportunity

Communities Served: Austin, TX; Flagstaff, AZ; Tucson, AZ; Davenport, IA; Baltimore, MD; Batavia, NY; Miami, FL; Brooklyn, NY; Hartford, CT; New Orleans, LA; Eugene, OR; Madison, WI; Philadelphia, PA; Portage, WI; North Charleston, SC; Richmond, VT; Grand Junction, CO; and Independence, WI

What was the purpose of evaluation?

The evaluation of the AmeriCorps Opportunity Youth Program ran from 2014-2018. The evaluation bundled 19 small AmeriCorps programs (receiving less than $500,000/year) into a single quasi-experimental evaluation using a propensity-score matched comparison group. The study assessed whether OY who participated in AmeriCorps showed greater improvements in education, employment, and civic engagement outcomes than comparison youth. Surveys were administered to youth at three time points-- the start, end, and six months after their service with AmeriCorps. The final matched sample consisted of 124 treatment youth and 90 comparison youth from 15 sites-- with program sample sizes ranging from two to 40 youth. Additional exploratory analyses examined changes in outcomes for a sample of 216 treatment youth within all 19 sites.

Evaluation At-a-Glance

Evaluation Design(s): Propensity-Score Matched Quasi-Experimental Impact Evaluation

Study Population: Opportunity Youth (16-24)

Independent Evaluator: JBS International

This Evaluation's Level of Evidence*: Strong

*SIF and AmeriCorps currently use different definitions of levels of evidence.

What did we learn from the evaluation?

The CNCS office of research and evaluation engaged JBS International as an independent evaluator to conduct an impact evaluation of the AmeriCorps opportunity youth program. Key findings indicated that:

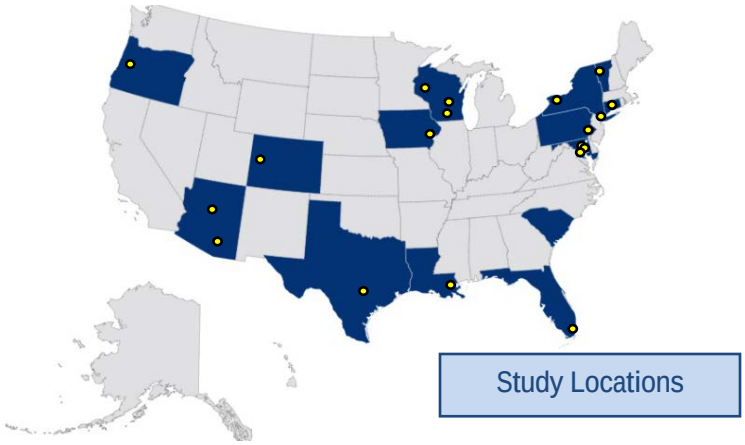
- AmeriCorps OY showed **improvements from pre-test to follow-up in high school completion, course completion, percentage currently employed, hourly wages, and longest employment**.
- AmeriCorps OY were not more likely to show improvements in key outcomes than comparison youth.
- Youth with a prior criminal history were **less likely to face charges or be on probation or parole** at follow-up if they had participated in AmeriCorps.
- Youth who had been employed at some point in the six months before the evaluation were **more likely to be currently employed** at follow-up if they had participated in AmeriCorps.
- Youth who had participated in AmeriCorps were **more likely to have recently completed a course** than comparison youth, especially if they were not high school graduates at pre-test.

AmeriCorps Opportunity Youth Evaluation Bundling Project

Notes on the evaluation

Evaluation findings do provide promising evidence of AmeriCorps' effectiveness with specific populations of youth, however overall main effects of program participation were not found in the current evaluation. Challenges with study recruitment and retention (especially among comparison youth and program dropouts) suggest the need for larger sample sizes (ideally 20 treatment and 20 comparison youth per site) and intensified retention efforts to detect smaller effect sizes in future studies.

Vermont Youth Conservation Corps members participating in an environmental conservation project.



19 programs in 16 states participated in the evaluation.

How are the AmeriCorps OY programs using the evaluation experience and findings?

AmeriCorps programs in the bundled evaluation received technical assistance (TA) both to support their implementation of the evaluation and to promote the sustainability of evaluation at their site. Evaluation sustainability TA was geared to help the participating programs to identify and pursue their evaluation goals. TA included support and training around developing and maintaining data collection systems, improving recruitment and retention for OY, identifying external resources to support evaluation, and developing and implementing future impact studies. At the study's conclusion, several programs commented on how much they had learned about evaluation and their desire to conduct future evaluations internally.

Additionally, within CNCS, these findings have been used as an example of a promising new methodology to help small grantees participate in an impact evaluation and may be used with future AmeriCorps grantees. The findings themselves also suggest the importance of continuing to provide AmeriCorps services to the most at-risk populations including those with a prior criminal history and those without a high school diploma.

The content of this brief was drawn from the full evaluation report submitted to CNCS by JBS International. The section of the brief that discusses evaluation use includes contribution of the AmeriCorps grantees who participated in the evaluation. All original content from the report is attributable to its authors.

To access the full evaluation report and learn more about CNCS, please visit http://www.nationalservice.gov/research.

Appendix B: Opportunity Youth Program Documents

B1. Opportunity Youth Bundled Evaluation: Theory of Change

Education

The comprehensive and simplified versions of the logic model indicate that remedial support, and other instruction support of members with educational deficits, will lead to improved educational outcomes for members. Interventions and outcomes are as follows:

1. For AmeriCorps members who lack a high school diploma or are at-risk of not graduating from high school, receiving individualized instruction and support in the form of classes and/or mentoring leads to successful completion of their education requirements. With member preparation and mastery, programs anticipate there will be increased high school graduation rates or GED certificates awarded to these youth.
2. For AmeriCorps members without any postsecondary education, participation in postsecondary preparation activities (including exploration of colleges or trade schools, assistance with applications and financial resources, and remedial education) leads to increased knowledge about postsecondary education (e.g., mastery of specific areas of training, college/trade applications, enrollment process, financial aid), and increased expectations regarding postsecondary education. This fuller understanding of the postsecondary education process and increase in expectations by members increases the completion of preparation activities, including completed college/trade school applications and other milestones in the post-application process (e.g., securing financial aid). This is expected to lead to increases in college or other postsecondary institutional enrollment.

Employment

The comprehensive and simplified versions of the logic model indicate that when members with poor employment prospects participate in employment training and career exploration, they improve their employment preparation, expand their career opportunities and are more likely to obtain and keep jobs. The outcomes and areas are as follows:

1. By providing AmeriCorps members currently lacking adequate employment opportunities with training in the area of “soft skills” (defined as skills that increase the ability of a person to search for, obtain, and maintain a job) members gain knowledge of how to find and keep a job. This includes skills such as resume writing, interviewing techniques, communication, and how to conduct a job search. With soft skill mastery, programs anticipate that members will find and maintain employment or secure other career opportunities, such as internships.
2. When AmeriCorps members currently lacking adequate employment opportunities receive training in “hard skills” (e.g., landscaping, construction, farming), they improve in these skill areas. They are then able to obtain certifications or licenses, pass tests, gain relevant work experience, and ultimately secure and maintain employment.

3. With counseling, and work/career exploration activities (e.g., exploring work/career options, developing realistic expectations, learning how to navigate the job environment despite disabilities/barriers), members develop strategies for obtaining employment and increase positive attitudes and expectations for obtaining and maintaining employment. Through these activities, members are better able to obtain and keep jobs, internships, or apprenticeships.

Community Connection

The comprehensive and simplified versions of the logic model indicate that when members are given opportunities to connect and integrate into the community through community service and information sessions, they acquire knowledge and experience about their community that make them better able to access community resources, become better integrated into the life of the local community, and decrease the likelihood that they will be involved in the criminal justice system. The outcomes and areas are as follows:

1. Informational sessions will increase members' knowledge of community resources. This will lead to an increased ability to successfully access community resource "products" such as enrollment in health care or completed housing applications.
2. Participation in community service activities will increase members' self-efficacy and sense of community, and offer experience that can be applied to later jobs. Increased employment opportunities, along with increased self-efficacy and sense of community, will lead to decreased involvement with the criminal justice system, including decreased rates of recidivism among those with prior arrests.

Though these groups of activities and expected outcomes have been described separately, in reality, they are expected to work in tandem to improve overall outcomes for opportunity youth. For example, hard skill development will generally occur in the context of community service, and may be combined with classroom-based activities in a service-learning model. Service learning may increase opportunity youth's sense of community while also improving the likelihood of academic achievement. Some programs may tailor activities or combine them to suit individual goals. Despite differences within and between programs, all share a common basis of activities that are expected to lead to improved outcomes in education, employment, and community connection for AmeriCorps opportunity youth members.

B2. Collective Logic Model for Opportunity Youth Programs

Inputs	Activities	Outputs	Outcomes	Impacts
Program structure and opportunity to serve Recruitment and outreach to target populations AmeriCorps funding and education awards Program staff experienced in working with at-risk populations Positive role models / mentors Program partnerships with community agencies, employers, and educational institutions	High school completion or GED support (e.g., mentors, classes) Postsecondary education preparation Soft skills instruction and support (e.g., resume writing) Hard skills instruction and support (e.g., construction) Counseling and work/career exploration Connect and integrate into the community; show members where to go for services and resources Community service activities Leadership activities (e.g., leadership skills training)	Members without diplomas or GEDs receive high school completion or GED support Members receive postsecondary education preparation Members receive training on job search (soft) skills Members receive training in trade/job (hard) skills Members receive work and career exploration support or counseling Members receive information on community resources and how to access them Members engage in service activities Members engage in leadership activities Members gain practical job experience	*During the program* Members increase desire/expectations for postsecondary education Members gain knowledge about postsecondary education Members gain knowledge of job search skills Members increase positive attitudes about obtaining and maintaining employment Members gain knowledge of community resources Members increase positive attitudes / sense of community Members increase self-efficacy *Up to three months after the program* Members complete job search components *Up to 12 months after the program* Members gain job experience Members complete coursework and/or take GED test.	*During the program through six or more months afterward* Members or alumni increase in GED certificates/H.S. diplomas Members or alumni increase access to (or utilization rate of) community products (e.g., housing applications, TANF) *By three to six months after the program* Members or alumni obtain a job/internship/ apprenticeship Members increase civic engagement (registered to vote, engaged in the community.) *During the program or up to 12 or more months after the program* Members or alumni increase in completed college/trade school applications Members or alumni decrease recidivism/ interaction with the criminal justice system *Three to 12 or more months after the program* Members or alumni increase in college enrollment *Six to 12 or more months after the program* Members or alumni maintain employment

Appendix C: Data Collection Tools

C1. Youth Employment and Education Survey

You are invited to participate in a research study that will examine youth employment and education outcomes. The study will be conducted by our partners at JBS International, a research organization, on behalf of the Corporation for National and Community Service. We are asking for your consent to participate in this project.

Do you agree to participate in this study?

- O Yes, I agree. Please indicate your agreement by signing below.
- O No, I do not agree. → STOP. Do not complete the survey.

What activities will you do in the study and how long will the activities last? If you decide to participate in the study, you will be asked to complete a survey in one of three formats: online, on paper, or over the phone. You will be asked to complete the survey a total of three times over a two-year period. The survey should take about 20 minutes to complete. The survey will include questions such as "What school did you most recently attend?" and "How much confidence do you have that you could successfully manage the job interview process?"

Benefits and Risks: We believe there is little or no risk to you in participating in this project. You may be contacted by email or phone to complete follow-up surveys. There is a possibility that you may become uncomfortable or stressed when answering an interview question or questions. If that happens, you may skip the question, take a break, or you may stop the survey or interview. You may also withdraw from the project altogether. There may be no direct benefits to you for participating in this research project. The results of this project may contribute to knowledge about factors affecting youth outcomes.

Confidentiality and Privacy: We will protect all the information from the surveys and phone interviews, and keep them in a safe place. Only our researchers will have access to the information. Your responses will be combined with the responses of other participants. No identifying information about you will be shared when the results of the survey are presented.

Voluntary Participation: Participation in this research project is voluntary. You can choose freely to participate or not to participate. At any point during this project, you can stop participating without any adverse consequences. You will receive a $10 gift card after completing the first survey, and again after completing the second survey; you will receive a $20 gift card for completing the third survey.

Please indicate your agreement to participate in today's study by signing below…

Printed Name: _____

Signature: _____ Date: _____

Youth Employment and Education Study

Thank you for your willingness to complete this survey. Your responses to this survey will be kept confidential and your responses will only be reported summarized with other responses. This is NOT a test. There are no right or wrong answers to the questions, so please choose the responses that best apply to you. This survey will take approximately 20 minutes to complete. Upon completion of this survey, you will receive a $10 gift card.

This study will take place over the course of two years. In order to track your responses confidentially over time, we are asking that you create a code that will enable us to link your surveys to each other.

a. First three letters of the city or town in which you were born: ____ ____ ____ [For example: S P R for Springfield]
b. Number of letters in your last name: ____ [For example 5 for Smith]
c. First 2 letters of your mother's first name: ___ ___ [For example M A for Mary. If unknown, enter "AA."]

1. Date of birth [Month/day/year]: _____

2. Gender

- O Male
- O Female
- O Other (specify): _____

3a. Are you a parent or primary caregiver of a child?

- O No
- O Yes, and all of my children live with me
- O Yes, and some of my children live with me
- O Yes, and none of my children live with me

3b. Are you a primary caregiver of a parent or other adult (e.g., disabled or sick relative)?

- O No
- O Yes

4. Have you served on active duty in the military?

- O No
- O Yes

5a. Do you consider yourself:

- O Hispanic or Latino origin
- O Not Hispanic or Latino origin

5b. What is your race? Please select one or more.

- ☐ American Indian or Alaskan Native
- ☐ Asian
- ☐ Black or African American
- ☐ Native Hawaiian or Other Pacific Islander
- ☐ White

6. Where do you currently live?

a. City or town _____

b. Zip code _____

7. Check the highest level of education that you have completed:

- ☐ Middle school → Skip to question 8.
- ☐ Some high school → Skip to question 8.
- ☐ High school diploma or GED
- ☐ Technical school / Apprenticeship
- ☐ Some college
- ☐ Associate's degree
- ☐ Bachelor's degree
- ☐ Graduate degree
- ☐ Other (please specify): _____

7a. When did you receive a high school diploma or GED? What month and year?

_____

8. When were you last enrolled in school – What was the month and year?

_____

9. What school did you most recently attend?

- ☐ Middle school
- ☐ High school
- ☐ Vocational/technical/alternative school (e.g., online school, trade school)
- ☐ Community college
- ☐ 4-year college

10. Were you employed at any point in the last 6 months?

- O No
- O Yes

11. Are you currently employed (AmeriCorps service does not qualify as employment)?

- O No → Skip to 11c.
- O Yes

11a. Please describe your current employment status:

- ☐ Part-time 1-20 hours per week
- ☐ Part time 21-39 hours per week
- ☐ Full-time 40 or more hours per week

11b. How long have you held your current position?

- ☐ Less than 1 month
- ☐ 1 to 3 months
- ☐ 4 to 6 months
- ☐ More than 6 months

11c. Please describe your current employment status:

- ☐ Looking for work
- ☐ Not looking for work
- ☐ Disabled, not able to work
- ☐ Pursuing school or training instead of work
- ☐ Engaged in part-time volunteer position, internship, or apprenticeship
- ☐ Engaged in full-time volunteer position, internship, or apprenticeship

12. In your most recent job, what was your hourly pay rate (in dollars)? _____

13. Have you ever done any of the following? Select all that apply

- ☐ A paid part-time job
- ☐ A paid full-time job
- ☐ A paid internship
- ☐ An unpaid internship
- ☐ Volunteer work that was not part of a requirement for high school graduation
- ☐ Received a stipend for attending a program (note: does not include AmeriCorps)
- ☐ Baby-sitting, yard-work or chores that you were paid for by a friend or neighbor

14. What is the longest amount of time you have been at a single job?

- O I have never been employed
- O Less than a month
- O 1 to 3 months
- O 4 to 6 months
- O 7 to 11 months
- O 1 to 2 years
- O More than 2 years

15. How much do you agree or disagree that each of the following statements describes you?

Statement	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
I can always manage to solve difficult problems if I try hard enough.					
If someone opposes me, I can find the means and ways to get what I want.					
It is easy for me to stick to my aims and accomplish my goals.					
I am confident that I could deal efficiently with unexpected events.					
Thanks to my resourcefulness, I know how to handle unforeseen situations.					
I can solve most problems if I invest the necessary effort.					
I can remain calm when facing difficulties because I can rely on my coping abilities.					
When I am confronted with a problem, I can usually find several solutions.					
If I am in trouble, I can usually think of a solution.					
I can usually handle whatever comes my way.					

16. If you found out about a problem in your community that you wanted to do something about, how well do you think you would be able to do each of the following.

Problem	I definitely could do this	I probably could do this	Not sure	I could not do this	I definitely could not do this
Create a plan to address the problem					
Get other people to care about the problem					
Organize and run a meeting					
Express your views in front of a group of people					
Identify individuals or groups who could help you with the problem					
Express your views on the Internet or through social media					
Call someone on the phone you had never met before to get their help with the problem					
Contact an elected official about the problem					

17a. How much confidence do you have that you could:

Statement	No confidence at all	Very little confidence	Moderate confidence	Much confidence	Complete confidence
Use the internet to find information about occupations that interest you					
Select one occupation from a list of potential occupations you are considering					
Determine what your ideal job would be					
Prepare a good resume					
Decide what you value most in an occupation					
Find out about the average yearly earnings of people in an occupation					
Identify employers, firms, and institutions relevant to your career possibilities					
Successfully manage the job interview process					
Identify some reasonable career alternatives if you are unable to get your first choice					

17b (Continued). How much confidence do you have that you could:

Statement	No confidence at all	Very little confidence	Moderate confidence	Much confidence	Complete confidence
Determine the steps to take if you are having academic trouble					
Complete a college or trade school application					
Apply for financial aid to further your educational goals					
Obtain formal training needed to support your career goals					
Pass a college course					
Obtain certification in a technical or vocational field (e.g., construction, landscaping, health)					
Sign up for health care					
Obtain housing vouchers or other housing assistance					
Find community resources that address your needs					

18. How much do you agree or disagree with the following.

Statement	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
I have a strong and personal attachment to a particular community					
I am aware of the important needs in the community					
I feel a personal obligation to contribute in some way to the community					
I am or plan to become actively involved in issues that positively affect the community					
I believe that voting in elections is a very important obligation that a citizen owes to the country					

19. Generally speaking, would you say that you can trust *none of the people, some of the people, most of the people,* or *all of the people* in your neighborhood?

- O All of the people
- O Most of the people
- O Some of the people
- O None of the people

20. Please indicate how much you agree or disagree with the following statements.

Statement	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
I have a clear idea of what my career goals are					
I have a plan for my career					
I intend to pursue education beyond high school (e.g., college, trade school)					
I know what to seek and what to avoid in developing my career path					

21. In the last 6 months, have you done any of the following? Check all that apply.

- ☐ Sent in a resume or completed a job application
- ☐ Written or revised your resume
- ☐ Interviewed for a job
- ☐ Contacted a potential employer
- ☐ Talked with a person employed in a field you are interested in
- ☐ Taken a GED test
- ☐ Completed a course in high school, college, or an alternative school
- ☐ Completed a college or trade school application
- ☐ Completed a financial aid application (e.g., FAFSA – Free Application for Federal Student Aid)
- ☐ Enrolled in a college, trade school, or a certification course

22. In the last 6 months, have you looked for any of the following? Check all that apply.

- ☐ Full-time work
- ☐ Part-time work
- ☐ Internship or apprenticeship
- ☐ Volunteer position

23. Please assess if the following factors are barriers to employment for you personally:

Factor	Not a barrier to employment	A barrier but can be overcome	Large barrier to employment
No jobs available where I live			
Do not have enough work experience for the job I want			
Do not have enough education for the job I want			
Have family or other responsibilities which interfere			
Do not have transportation			
Not good at interviews or do not know how to create a resume			
Can make more money not in an “official” job			
Criminal record makes it difficult to find a job			
Credit issues make it hard to find a job			
Illness or injury makes it challenging to find a job			
Do not wish to work			

24. Are you currently using or visiting any of the following? Check all that apply.

- ☐ Local employment development division (for unemployment insurance or for help with finding a job)
- ☐ Housing center (for help with finding housing)
- ☐ Job center
- ☐ Crisis center
- ☐ Homeless shelter
- ☐ Food bank
- ☐ Community health clinic
- ☐ Adult school / community college extension programs
- ☐ Mutual support or other assistance programs (e.g., AA, NA, AlAnon, grief support groups)

25. Are you currently accessing any of the following federal or state government supports? Check all that apply.

- ☐ Food assistance (e.g., WIC, SNAP)
- ☐ Health care assistance (e.g., Medicaid or other health insurance)
- ☐ Housing assistance (e.g., housing vouchers)
- ☐ Other financial or practical assistance (e.g., TANF, child care assistance programs)

26. Have you ever been convicted as an adult, or adjudicated as a juvenile offender, of any offense by either a civilian or military court, other than minor traffic violations?

- O No
- O Yes

27. Are you currently facing charges for any offense or on probation or parole?

- O No
- O Yes

28. Are you limited in any way in any activities because of physical, mental, or emotional problems?

- O No
- O Yes

29. How many times have you moved in the last 12 months?

- O I have not moved
- O Once
- O Two or more times

30. Were you registered to vote in the last presidential election?

- O Yes
- O No
- O No I was not eligible to vote
- O Don't know

31. Did you vote in the last presidential election?

- O Yes
- O No
- O Don't know

32. In the last 12 months, how often did you participate in the following activities?

Activity	Not at all	Less than once a month	Once a month	A few times a month	A few times a week	Basically every day
Participate in community organizations (school, religious, issue-based, recreational)						
Keep informed about news and public issues						
Help to keep the community safe and clean						
Volunteer for a cause or issue that I care about						
Donate money or goods to a cause or issue that I care about						

33. What programs are you participating in or services are you receiving? Check all that apply.

- ☐ AmeriCorps or similar national or community service program (e.g., Job Corps, YouthBuild, City Year, Public Allies, Year Up)
- ☐ Employment supports, other than AmeriCorps (e.g., job training)
- ☐ Educational supports, other than AmeriCorps (e.g., tutoring, GED classes, college enrollment assistance)

C2. Survey and Scale Construction, Reliability, and Validity Results

Survey items collected two categories of data: (1) demographic characteristics to be used in propensity score matching and/or entered as covariates in the final analysis; and (2) data regarding outcomes and impacts shared by the bundle. In developing the survey, a number of existing tools were identified and examined for potential use in collecting data on demographics, outcomes, and impacts.

The AmeriCorps application, Exit survey, and Alumni survey were used as the basis of several questions. In addition, items from the following instruments were used or adapted:

- Behavioral Risk Factor Surveillance System survey (BRFSS). http://www.cdc.gov/brfss/
- Career Decision Self-Efficacy scale (CDSE). Betz, N. E., Klein, K. L., & Taylor, K. M. (1996). Evaluation of a short form of the career decision-making self-efficacy scale. *Journal of Career Assessment*, 4(1), 47-57.
- Career Competencies Indicator (CCI). Francis-Smythe, J., Haase, S., Thomas, E., & Steele, C. (2013). Development and validation of the career competencies indicator (CCI). *Journal of Career Assessment*, 21(2), 227-248.
- Competencies for Civic Action scale. Flanagan, C. A., Syvertsen, A. K., & Stout, M. D. (2007). *Civic Measurement Models: Tapping Adolescents' Civic Engagement*. CIRCLE Working Paper 55. Center for Information and Research on Civic Learning and Engagement (CIRCLE).
- DC Alliance of Youth Advocates survey (DCAYA). http://www.dc-aya.org/
- Gates Foundation survey. http://www.gatesfoundation.org/
- National Longitudinal Study of Youth (NLSY). https://nlsinfo.org/

The survey instrument was pilot tested with 36 youth in treatment in the year prior to the evaluation, and cognitive interviews were conducted with a subset of nine survey completers. In several cases, minor adjustments were made to questions based on participant feedback. During pilot testing, the survey was assessed and found to have sufficient item variability and scale reliability. Scale reliability was confirmed with the complete pre-test responses, with Cronbach's alpha ranging from .78 to .90.

Scale	Items	Cronbach's alpha
General self-efficacy	10	0.90
Community self-efficacy	8	0.88
Work self-efficacy	9	0.90
Academic self-efficacy	6	0.88
Support self-efficacy	3	0.86
Sense of community	5	0.80
Goals - work	3	0.86
Community engagement	5	0.78

C3. Member Participation in Training Activities Document

Youth Employment and Education Evaluation: Member Participation in Training Activities

Program or Site Name:

INSTRUCTIONS

1. In the first column, enter each member's name. In the example in the table, we entered John Doe and Jane Doe.
2. In the second column under Returning Member, please put a "Y" if the member is in her/his second term as a member and a "N" if the member is a new member.
3. Next, for each member, use the Target # of AC Hours column to record the member's anticipated term-length (e.g., 900 hours for a half-time slot) and then use the Total # of AC Hours Served column to record the number of AmeriCorps hours the member ultimately served during their term.
4. Then use the Term Start Date and Term Stop Date columns to record the dates at which each member began and ended their current term of service, and if a member dropped out of the program please include your best estimate of their drop-out date under the Term Stop Date.
5. Additionally, if any member did not complete their service please use the Reason for Not Completing Term column to briefly explain (to the best of your knowledge) why they dropped out of the program.
6. For each member please record which training activities (Education Preparation, Employment Preparation, and Community Connection) the member participated in. Please use the Training Activities Descriptions at the bottom of this table to determine which activities are included under each category.
7. For each member, enter a "Y" if the member was eligible and participated in the activity, enter "N" if the member was eligible but did not participate in the activity, or enter "N/A" if the member was ineligible for the activity or the activity was not relevant to the member's service.
8. There is an "Other" option included for each category (e.g., Education Preparation). If there is a training activity that your program provides that is not listed in the table, please rename the "Other" column with the name of the activity, and repeat instruction #4. Please do this for any activities involving education preparation, employment preparation, and community connection that are not included in the table. Then, in the section below the table, please define the newly added activity in the row labelled "Other" and define the activity.

	AC Service Information						AC Outcomes			Education Preparation				Employment Preparation				Community Connection				
Member Name or ID	Returning member (Y/N)	Target # of AC Hours	Total # of AC Hours Served	Term Start Date	Term Stop Date	Reason for not completing term?	Received Workforce Development Services	Received GED/Education Services	Received Industry Recognized Cert/Credential	GED or HS Diploma Support	HS Completion Support	Post-Secondary Ed Preparation	Other	Soft Skill Instruction	Hard Skill Instruction	Career Counseling	Other	Service Referrals	Community Service	Civic Literacy	Leadership Development	Other
Typical Member:	*N*	*900*	*900*	*Rolling*	*9-11 mon.*	*N/A*	*Y*	*N*	*Y*	*Y*	*Y*	*Y*	*N/A*	*Y*	*Y*	*Y*	*N/A*	*Y*	*N/A*	*N/A*	*Y*	*N/A*
John Doe	*Y*	*900*	*652*	*8/15/2015*	*10/15/2015*	*Accepted another job*	*Y*	*N*	*N/A*	*N*	*N*	*N/A*	*N*	*Y*	*N/A*	*N*	*N/A*	*N*	*Y*	*N/A*	*N/A*	*N*
Jane Doe	*N*	*900*	*900*	*8/15/2015*	*12/15/2015*	*N/A*	*Y*	*Y*	*Y*	*Y*	*N/A*	*N/A*	*N*	*N*	*N/A*	*N*	*N/A*	*N/A*	*Y*	*N/A*	*N/A*	*N*

Training Activity	*Description*
Education Preparation	
GED or HS Diploma Support	Participated in GED or HSED classes (if member already has a HSD or is enrolled in high school please mark this as "N/A")
HS Completion Support	Participated in mentoring, classes, or other programs to encourage youth to enroll or remain enrolled and complete HS (if member has a GED or HSD please mark this as "N/A")
Post-secondary Ed Exploration	Participated in trade school/college exploration, received application assistance, received information on financial resources for education
Other	
Employment Preparation	
Soft Skills Instruction & Support	Participated in resume writing, interviewing, networking, connecting with potential employers, or other soft skills trainings
Hard Skill Instruction & Support	Participated in hard skill training (e.g., in construction, landscaping, chainsaw use, farming, environmental preservation)
Career Counseling	Received counseling and support to explore realistic work/career options, develop realistic expectations, learn how to navigate the job environment despite barriers
Other	
Community Connection	
Service Referrals	Received information and/or referrals to community resources and services (e.g., healthcare enrollment supports, housing assistance, employment programs)
Community Service Activities	Participated in community service activities (e.g., community beautification, environmental conservation, work with disadvantaged populations)
Civic Literacy Activities	Participated in civic literacy activities (e.g., information sessions on voter registration, strategies for enacting community change)
Leadership Development	Participated in leadership development activities (e.g., mentoring youth, leadership skills training)
Other	

Appendix D: Survey Administration Timing and Response Rates

Table D1. Survey Administration Periods by Program, September 2015 – August 2017

Program	Survey	2015 Sep	Oct	Nov	Dec	2016 Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	2017 Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
CW	pre post follow																								
LV	pre post follow																								
SI	pre post follow																								
AY*	pre post follow																								
DC*	pre post follow																								
RU*	pre post follow																								
WD*	pre post follow																								

Note: Sites marked with an asterisk have a rolling enrollment period or extended their enrollment period due to slower than anticipated recruitment. These sites also administered the post-test survey on a rolling basis.

Table D2. Survey Completion by Modality and Condition for Each Survey Wave

Survey	Paper Treatment	Paper Comparison	Online Treatment	Online Comparison	Phone Treatment	Phone Comparison
Survey 1	447	516	5	65	0	9
Survey 2	245	7	48	118	7	17
Survey 3	45	0	165	137	9	6

Appendix E: Survey Non-Response Bias Analysis

Significant Differences between Respondents and Non-Respondents.

JBS conducted a non-response bias analysis to determine if the non-response was random, or if there were systematic differences in characteristics between respondents and non-respondents. In particular, the non-response bias analysis examined whether respondents to the follow-up survey differed systematically from the target population. To examine non-response bias across key variables, JBS used a relative non-response bias measure based on the formula for non-response bias defined by the Office of Management and Budget (OMB). This non-response measure is derived by first dividing the difference in respondent and non-respondent means by the mean of the entire sample (in this case, the complete sample of pre-test respondents), and then multiplying this by the non-response rate (defined as the number of non-respondents divided by the number of pre-test respondents). JBS calculated non-response for all respondents (Table E1), then separately for treatment members (Table E2) and comparison group members (Table E3). This result is then divided by the mean for all respondents.

Non-Response Bias in the Full Sample. First, JBS assessed differences in response patterns among treatment and comparison group members (Table E1), and found that selective attrition of comparison group members was relatively high, with a relative bias of 35 percent. Among the variables examined, the relative bias was as low as 0.2 percent for disability, 0.9 percent for age, and 10 percent for employment history. At the high end, relative bias ranged from 19 percent to 47 percent for education (high school graduate), gender, military status, parental status, caregiver status (for an adult), criminal history, and ethnicity (Non-Hispanic White vs. Not).

Table E1. Non-Response Bias—All Respondents

Variable	Mean - all	Mean - respondent	Mean - non-respondents	# cases	# respondents	# non-respondents	Difference in means	Bias
Type - Comparison	.45	0.61	0.37	1,072	355	717	0.24	35.3%
Female	.37	0.45	0.33	1,067	352	715	0.12	21.7%
Age	20.43	20.61	20.35	1,072	355	717	0.26	0.9%
Parent	.21	0.16	0.24	1,071	355	716	-0.08	-26.2%
Caregiver	.05	0.03	0.06	1,068	355	713	-0.03	-31.9%
Military	.02	0.03	0.02	1,069	355	714	0.01	23.2%
Non-Hispanic White	.28	0.41	0.22	1,070	355	715	0.19	46.7%
HS Graduate	.61	0.72	0.55	1,049	350	699	0.17	19.0%
Recently employed	.56	0.61	0.53	1,069	354	715	0.08	10.3%
Received government support	.48	0.42	0.51	1,072	355	717	-0.09	-13.1%
Criminal history	.30	0.20	0.35	1,057	352	705	-0.15	-34.2%
Disability	.10	0.10	0.10	1,063	355	708	0.00	-0.2%

To help contextualize these differences, JBS also conducted a logistic regression examining the effects of each variable on non-response when all other variables tested were held constant (see Table E2). When examining all respondents, comparison group members, males, non-Whites, youth who had not graduated from high school at baseline, and youth with a criminal history were all significantly less likely to have responded. No other variables exhibited statistically significant differences. This suggests that treatment, gender, high school graduation status at pre-test and criminal history at pre-test are all covariates that could bias results. To address this JBS controlled for each of these variables in the final impact analysis.

Table E2. Likelihood of Non-Response Based on Participant Characteristics—All Respondents

Variable	B	S.E.	Wald	df	Sig.	Exp(B)
Type - Treatment	0.842	0.146	33.470	1	0.000	2.322
Female	0.443	0.150	8.690	1	0.003	1.557
Age	0.023	0.037	0.387	1	0.534	1.023
Parent	-0.082	0.200	0.167	1	0.683	0.922
Caregiver	0.002	0.364	0.000	1	0.995	1.002
Military	0.470	0.476	0.976	1	0.323	1.600
Non-Hispanic White	0.570	0.161	12.583	1	0.000	1.769
HS Graduate	0.474	0.174	7.385	1	0.007	1.606
Employed in past 6 months	0.236	0.146	2.636	1	0.104	1.267
Received government support	-0.108	0.153	0.496	1	0.481	0.898
Criminal history	-0.596	0.174	11.700	1	0.001	0.551
Disability	0.093	0.236	0.156	1	0.692	1.098

Non-Response Bias in the Treatment-Only Sample. Next, JBS performed non-response bias analysis on the treatment group only (Table E3). Relative bias was lower in this group, in part because of the substantially higher response rate in this group. The relative bias was as low as 1.5 percent for age and 2.1 percent for parental status. At the high end, relative bias was as high as 26 percent for caregiver status and 67 percent for military status.

Table E3. Non-Response Bias—Treatment Members

Variable	Mean - all	Mean - respondent	Mean - non-respondents	# cases	# respondents	# non-respondents	Difference in means	Bias
Female	0.352	0.419	0.298	480	215	265	0.121	18.9%
Age	20.746	21.060	20.491	482	216	266	0.568	1.5%
Parent	0.150	0.153	0.147	481	216	265	0.006	2.1%
Caregiver	0.044	0.032	0.053	480	216	264	-0.021	-25.9%
Military	0.017	0.028	0.008	481	216	265	0.020	67.0%
Non-Hispanic White	0.349	0.426	0.287	481	216	265	0.139	21.9%
HS Graduate	0.719	0.818	0.639	474	214	260	0.179	13.7%
Recently employed	0.570	0.634	0.517	481	216	265	0.117	11.3%
Received government support	0.461	0.421	0.493	482	216	266	-0.072	-8.5%
Criminal history	0.285	0.235	0.326	474	213	261	-0.091	-17.6%
Disability	0.090	0.102	0.080	479	216	263	0.022	13.5%

To help contextualize the relative bias results, a logistic regression was also conducted to predict non-response within the treatment only group. When examining treatment youth only, women, non-Hispanic whites, and high school graduates were all significantly more likely to have responded (Table E4). Furthermore, an exploratory model was also run with an additional covariate indicating whether treatment youth successfully completed their target number of AmeriCorps hours. Results indicated that treatment youth who completed their target number of AmeriCorps hours had five times greater odds of completing the follow-up survey than those who did not meet this target (Table E5). This suggests that gender, high school graduation status at pre-test, and race/ethnicity are all covariates that could bias results. To address this JBS controlled for each of these variables in the final impact analysis and in exploratory treatment-only analyses.

Table E4. Likelihood of Non-Response Based on Participant Characteristics—Treatment

	B	S.E.	Wald	df	Sig.	Exp(B)
Female	.449	.211	4.548	1	.033	1.567
Age	.072	.052	1.916	1	.166	1.075
Parent	.337	.296	1.289	1	.256	1.400
Caregiver	-.315	.497	.403	1	.526	.730
Military	1.230	.857	2.059	1	.151	3.421
Non-Hispanic White	.470	.222	4.486	1	.034	1.600
HS Graduate	.654	.270	5.889	1	.015	1.924
Employed in past 6 months	.382	.202	3.580	1	.058	1.465
Received government support	-.194	.210	.857	1	.355	.824
Criminal history	-.396	.234	2.861	1	.091	.673
Disability	.399	.345	1.336	1	.248	1.490

Table E5. Likelihood of Non-Response Based on Participant Characteristics—Treatment with Completion of Target AmeriCorps Hours Added

	B	S.E.	Wald	df	Sig.	Exp(B)
Completed Target # of AC Hours	1.667	0.248	45.368	1	.000	5.297
Female	.600	.226	7.037	1	.008	1.823
Age	.074	.056	1.769	1	.184	1.077
Parent	.442	.315	1.965	1	.161	1.555
Caregiver	-.293	.524	.312	1	.576	.746
Military	.958	.953	1.011	1	.315	2.608
Non-Hispanic White	.471	.236	3.978	1	.046	1.601
HS Graduate	.744	.284	6.845	1	.009	2.104
Employed in past 6 months	.531	.215	6.104	1	.013	1.701
Received government support	-.245	.225	1.184	1	.277	.783
Criminal history	-.243	.252	.933	1	.334	.784
Disability	.561	.371	2.295	1	.130	1.753

Non-Response Bias in the Comparison-Only Sample. Finally, JBS performed non-response bias analysis on the comparison group only. Relative bias was substantially higher in this group, in part due to the lower response rate among comparison group members. The relative bias was as low as 1.4 percent for age and 6.5 percent for recent employment (see Table E6). Many variables had relative bias at the high end, with relative bias of 38 percent to 74 percent for parental status, criminal history, and ethnicity (Non-Hispanic white vs. not).

Table E6. Non-Response Bias—Comparison Members

Variable	Mean - all	Mean - respondent	Mean - non-respondents	# cases	# respondents	# non-respondents	Difference in means	Bias
Female	0.387	0.504	0.351	587	137	450	0.153	30.2%
Age	20.177	19.905	20.261	590	139	451	-0.356	-1.4%
Parent	0.266	0.166	0.297	590	139	451	-0.131	-37.8%
Caregiver	0.054	0.036	0.060	588	139	449	-0.024	-33.9%
Military	0.024	0.022	0.025	588	139	449	-0.003	-9.4%
Non-Hispanic White	0.228	0.396	0.176	589	139	450	0.220	73.9%
HS Graduate	0.515	0.574	0.497	575	136	439	0.077	11.4%
Recently employed	0.544	0.580	0.533	588	138	450	0.047	6.5%
Received government support	0.502	0.417	0.528	590	139	451	-0.110	-16.8%
Criminal history	0.309	0.137	0.363	583	139	444	-0.226	-55.7%
Disability	0.106	0.094	0.110	584	139	445	-0.017	-11.9%

Again, JBS performed a logistic regression to determine which variables were significant predictors of non-response within the comparison-only sample. Within the comparison-only group, non-Hispanic whites and individuals with no criminal history were more likely to have responded (Table E7). To address this, JBS controlled for each of these variables in the final impact analyses.

Table E7. Likelihood of Non-Response Based on Participant Characteristics—Comparison

	B	S.E.	Wald	df	Sig.	Exp(B)
Female	.411	.220	3.485	1	.062	1.509
Age	-.028	.056	.246	1	.620	.973
Parent	-.409	.296	1.919	1	.166	.664
Caregiver	.435	.543	.639	1	.424	1.544
Military	-.023	.702	.001	1	.974	.977
Non-Hispanic White	.800	.239	11.223	1	.001	2.225
HS Graduate	.236	.237	.989	1	.320	1.266
Employed in past 6 months	.067	.218	.095	1	.758	1.070
Received government support	-.011	.233	.002	1	.964	.990
Criminal history	-.963	.287	11.278	1	.001	.382
Disability	-.180	.352	.263	1	.608	.835

Appendix F: Propensity Score Matching Methods and Results

Characteristics of the pre-match sample. A total of 355 youth completed the follow-up survey and were included in the propensity score matching analysis (216 treatment and 139 comparison). The first section of Table F1 (pre-match sample) shows the characteristics of this sample before matching the two groups. Prior to matching, treatment and comparison youth had similar characteristics on all matching variables except for age, status as a high school graduate, criminal history, and month that Survey 1 was completed. On average, treatment youth were significantly older (M= 21.1, SD=2.3) than comparison youth (M=19.9, SD=2.1) and a higher percentage of treatment youth were high school graduates at pre-test (82%) than comparison youth (57%). Additionally, a higher percentage of treatment youth had a criminal history at pre-test (23%) than comparison youth (14%), and, on average, treatment youth took the pre-test about one month earlier (M=6.6, SD=3.0) than comparison youth (M=7.6, SD=2.1). On all other matching variables, treatment and comparison youth were very similar. Most participants were not the parent or primary caregiver of a child, were not a part of the military, did not have a disability, and did not receive any government support.

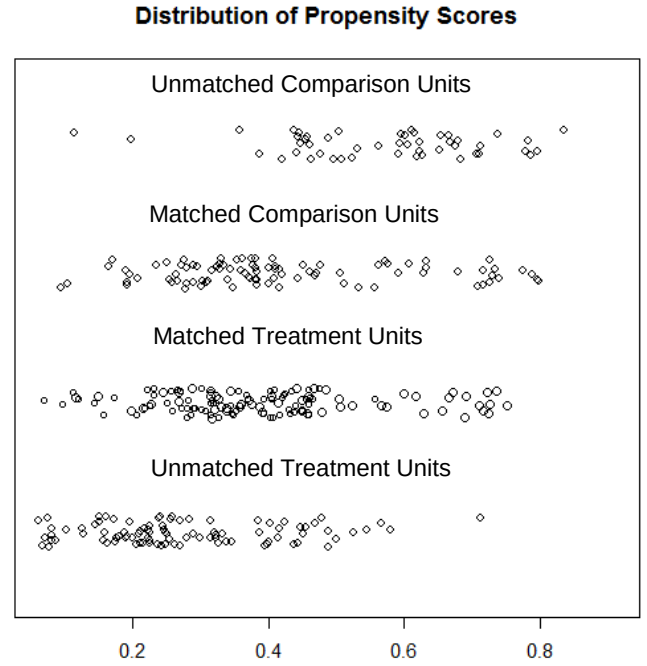
Table F1. Baseline Equivalence Before and After the Propensity Score Matching

Variables	**Pre-match Sample Treat. (N=216)** Mean (SD) / Percent	**Pre-match Sample Comp. (N=139)** Mean (SD) / Percent	**Pre-match Mean Std. Diff**	**Pre-match Sig. Diff?**	**Post-Match Sample (2:1) Treat. (N=124)** Mean (SD) / Percent	**Post-Match Sample (2:1) Comp. (N=90)** Mean (SD) / Percent	**Post-match Mean Std. Diff**	**Post-match Sig. Diff?**
Month Survey 1 was completed	6.6 (3.0)	7.6 (2.1)	-0.38	**	7.4 (2.8)	7.7 (2.2)	-0.12	n.s.
Number of days between Survey 1 and Survey 3	385.2 (90.9)	379.7 (82.9)	0.06	n.s	375.7 (97.6)	376.5 (86.89)	-0.01	n.s.
Age	21.1 (2.3)	19.9 (2.1)	0.52	***	20.5 (2.3)	20.3 (2.2)	0.08	n.s.
Gender (Female)	42%	50%	-0.17	n.s.	47%	48%	-0.01	n.s.
Non-Hispanic White	43%	40%	0.06	n.s.	48%	44%	0.08	n.s.
Parent or primary caregiver of a child	15%	17%	-0.03	n.s.	12%	9%	0.10	n.s.
Primary caregiver of a parent or other adult	3%	4%	-0.02	n.s.	4%	2%	0.10	n.s.
Served on active duty in the military	3%	2%	0.04	n.s.	2%	1%	0.10	n.s.
High school graduate	82%	57%	0.55	***	77%	72%	0.12	n.s.
Employed in the last 6 months	63%	58%	0.11	n.s.	62%	67%	-0.10	n.s.
Criminal history (charged or convicted)	23%	14%	0.25	*	14%	13%	0.02	n.s.
Received any government support	42%	42%	0.01	n.s.	33%	32%	0.02	n.s.
Has a disability	10%	9%	0.03	n.s.	14%	9%	0.15	n.s.

Note. Significance testing is based on two-tailed independent samples Mann-Whitney tests or chi-squared tests comparing treatment and comparison youth within the pre-match and post-match samples. * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Matching results. Propensity score matching was conducted using 2:1 treatment to comparison nearest neighbor matching with a 0.5 standard deviation caliper and exact matching on program.[24] Overall, the final matched dataset retained 60 percent of the follow-up survey sample of youth who were eligible for matching (N=355), 57 percent of eligible treatment youth (N=124), and 65 percent of eligible comparison youth (N=90). Within sites, the match rate ranges from 15-100 percent (see Appendix H for more detailed information on the success of matching within each program).

Figure F1. Distribution of Propensity Scores



Characteristics of the matched sample. Baseline equivalence findings for the full pre-match and post-match datasets are presented in the second section of Table F1 above ("Post-match Sample"). As noted previously, prior to the match, treatment youth were significantly older than comparison youth and a higher percentage of treatment youth were high school graduates than comparison youth. Additionally, a higher percentage of treatment youth had a criminal history at pre-test than comparison youth, and, on average, treatment youth took the pre-test about one month earlier than comparison youth. After the match, these differences were no longer statistically significant, indicating that the match had successfully equalized the two groups on the baseline matching characteristics.

[24] To conduct 2:1 matching in SPSS, the smaller group must be coded as the reference/treatment group. To accommodate this, the treatment variable was reverse coded so that propensity scores reflected the likelihood of being in the comparison group rather than the treatment group (see Figure 1 above). However, mean standardized differences as presented in Table F1 above were based on the correct treatment versus comparison groupings using the formulas specified in Austin (2011).

Appendix G: Impact Analyses and Results

G1. Outcome Differences for the Matched Sample

Table G1. Mean Differences in Outcomes from Pre-Test to Follow-up for Treatment (N=124) and Comparison (N=90) Youth

Outcome Area	Outcome Variables	Treatment Pre-Test Mean	Treatment Post-Test Mean	Treatment Change	Comparison Pre-Test Mean	Comparison Post-Test Mean	Comparison Change	Difference-in-Change
	Primary Outcome Variables							
Education	Has a high school diploma or GED	77%	90%	13%	72%	88%	16%	-3%
Education	Completed a course in high school, college, or alternative school	39%	33%	-6%	48%	33%	-14%	9%
Education	Enrolled in college, trade school, or certification course	26%	27%	1%	28%	32%	4%	-4%
Employment	Employed currently	11%	66%	54%	40%	60%	20%	35%
Employment	Hourly income (in dollars)	9.61	10.72	1.11	9.42	10.88	1.45	-0.34
Employment	Longest employment experience (in months)	9.82	12.51	2.68	11.97	14.47	2.50	0.18
Employment	Number of job search activities completed in the last 6 months (0-9)[1]	5.51	5.10	-0.40	5.02	4.61	-0.41	0.01
Civic Engagement	Currently facing charges or on probation or parole	10%	5%	-5%	8%	11%	3%	-8%
Civic Engagement	Mean frequency of civic engagement activities (1-6)[2]	3.20	3.11	-0.10	2.89	2.93	0.04	-0.14
Outcome Area	**Secondary Outcome Variables**							
Education	Took a GED test	4%	5%	1%	3%	2%	-1%	2%
Education	Intends to pursue education beyond high school (1-5)[3]	4.37	4.17	-0.20	4.27	4.18	-0.09	-0.11
Education	Confidence in the ability to obtain postsecondary education (1-5)[4]	4.15	4.09	-0.06	3.96	4.05	0.09	-0.16
Education	Applied to college or trade school	10%	16%	6%	14%	16%	1%	5%
Employment	Employed in last 6 months	62%	75%	13%	67%	74%	7%	6%
Employment	Number of types of employment experiences (0-7)[5]	2.80	3.15	0.35	2.91	3.31	0.40	-0.05
Employment	Confidence in job search skills (1-5)[4]	3.96	4.00	0.04	3.89	3.97	0.08	-0.04

Outcome Area	Outcome Variables	Treatment Pre-Test Mean	Treatment Post-Test Mean	Treatment Change	Comparison Pre-Test Mean	Comparison Post-Test Mean	Comparison Change	Difference-in-Change
Employment	Average perceptions of barriers to employment (1-3)[6]	1.51	1.45	-0.06	1.48	1.47	-0.01	-0.05
Civic Engagement	Ever been convicted of a crime	10%	15%	5%	10%	18%	8%	-3%
Civic Engagement	Confidence in ability to access community resources (1-5)[4]	3.62	3.63	0.01	3.55	3.74	0.19	-0.18
Civic Engagement	Average sense of community (1-5)[3]	3.89	3.77	-0.12	3.72	3.64	-0.08	-0.04
Civic Engagement	Trust in community members (1-4)[7]	2.36	2.48	0.11	2.32	2.51	0.19	-0.08
Civic Engagement	Average self-efficacy to solve personal problems (1-5)[3]	3.96	3.96	0.01	3.97	3.96	-0.01	0.02
Civic Engagement	Average self-efficacy to solve community problems (1-5)[8]	3.93	3.96	0.03	3.84	3.91	0.07	-0.05
Civic Engagement	Received any form of government support[9]	33%	40%	7%	32%	34%	2%	5%
Civic Engagement	Received any form of community support[10]	26%	31%	5%	28%	28%	0%	5%

Table Notes.

[1] Up to 9 of the following: Submitted job application, wrote/revised resume, interviewed, contacted employer, had informational interview, looked for full-time work, looked for part-time work, looked for internship/apprenticeship, looked for volunteer position.

[2] 1=Not at all, 2= Less than once a month, 3= Once a month, 4=A few times a month, 5=A few times a week, 6=Basically every day.

[3] 1=Strongly Agree, 2= Disagree, 3=Neither Agree nor Disagree, 4=Agree, 5=Strongly Agree

[4] 1= No confidence at all, 2=Very little confidence, 3=Moderate confidence, 4=Much confidence, 5=Complete Confidence

[5] Up to 7 of the following: paid part-time job, paid full-time job, paid internship, unpaid internship, volunteer work, program stipend, paid babysitting/yardwork

[6] 1=Not barriers, 2=Barriers that can be overcome, 3=Large barriers

[7] 4 = Can trust all of the people in my neighborhood, 3 = Most, 2 = Some, 1 = None of the people

[8] 1 = I definitely could not do this , 2 = I could not do this, 3 = Not sure, 4 = I probably could do this, 5 = I definitely could do this.

[9] Government supports included currently accessing: food assistance (such as WIC or SNAP), health care assistance (such as Medicaid or other health insurance), housing assistance (such as housing vouchers), or other financial or practical assistance (such as TANF or child care assistance programs).

[10] Community supports included using or visiting any of the following: local unemployment division, housing center (for help with finding housing), job center, crisis center, homeless shelter, food bank, community health clinic, adult school/community college extension programs, mutual support or other assistance programs (e.g., Alcoholics Anonymous, Narcotics Anonymous, AlAnon Family Groups, grief support groups).

G2. Multivariate Mixed Effects Models Predicting Outcomes

Table G2. Educational Outcomes Predicted by Baseline Characteristics, Treatment, And Random Effects of Treatment by Program and Site (N=214)

Student Level Fixed Effects (from Pre-test Characteristics)	***Obtained HS Diploma or GED***[25]	b [26]	SE	Exp(B)	Sig?	***Completed a Course in High School or College in the Last 6 months***	b	SE	Exp(B)	Sig?	***Enrolled in college, trade school, or certification course***	b	SE	Exp(B)	Sig?
Treatment		0.15	0.69	1.17	n.s.		1.21	0.70	3.35	†		-0.55	1.80	0.58	n.s.
Criminal history (1=Yes)		-0.73	0.92	0.48	n.s.		-0.75	0.60	0.47	n.s.		-0.10	0.70	0.91	n.s.
Employed in the last 6 months (1=Yes)		-0.01	0.67	0.99	n.s.		0.64	0.39	1.90	†		0.10	0.51	1.10	n.s.
Age		0.39	0.19	1.48	*		-0.19	0.10	0.83	†		-0.21	0.36	0.81	*
Female (1=Yes)		0.39	0.72	1.47	n.s.		0.15	0.35	1.16	n.s.		0.22	0.10	1.24	n.s.
Parent or Primary Caregiver of a child (1=Yes)		-2.31	1.18	0.10	*		-0.41	0.72	0.66	n.s.		0.53	0.34	1.70	n.s.
Non-Hispanic White (1=Yes)		0.74	0.77	2.09	n.s.		0.13	0.38	1.14	n.s.		-0.07	0.58	0.93	n.s.
Days between pre-test and follow-up[27]		0.21	0.39	1.24	n.s.		-0.58	0.20	0.56	**		-0.30	0.37	0.74	n.s.
High school graduate or GED (1=Yes)		N.A.	N.A.	N.A.	N.A.		0.04	0.65	1.04	n.s.		0.32	0.19	1.38	n.s.
Pre-test value of outcome		N.A.	N.A.	N.A.	N.A.		1.05	0.37	2.86	**		0.86	0.61	2.35	*
HSGrad*Treatment		N.A.	N.A.	N.A.	N.A.		-1.48	0.81	0.23	†		0.32	0.37	1.37	n.s.

*** p<.001, ** p<.01, * p<.05, † p<.10, n.s. p>.10

[25] The multi-level model predicting high school graduation status at follow-up did not fully converge (most likely because of the small number of youth who were not high school graduates at pre-test, N=53), Convergence error messages indicated that the model was nearly unidentifiable with a very large eigenvalue. To address this, a logistic regression model was run instead of the originally proposed multi-level model and predicted high school graduation status at post-test strictly among youth who were not high school graduates at pre-test.

[26] In this table and throughout this report, values labeled as "*b*" represent the unstandardized coefficients for each term in each model.

[27] The variable measuring the number of days between pre-test and follow-up was standardized for all mixed effects models to improve model fit.

Table G3. Employment Outcomes Predicted by Baseline Characteristics, Treatment, And Random Effects of Treatment by Program and Site (N=214)

Student Level Fixed Effects (from Pre-test Characteristics)	***Employed Currently***	b	SE	Exp(B)	Sig?	***Hourly Wage***	b	SE	Sig?	***Longest Employment***	b	SE	Sig?	***Number of Job Search Activities Completed***	b	SE	Sig?
Treatment		-0.64	0.53	0.53	n.s.		-2.31	1.21	†		0.84	1.52	n.s.		0.84	1.52	n.s.
Criminal history (1=Yes)		-0.74	0.44	0.48			-0.37	0.85	n.s.		0.75	1.16	n.s.		0.75	1.16	n.s.
Employed in the last 6 months (1=Yes)		-0.13	0.50	0.88	n.s.		-0.38	1.01	n.s.		3.00	1.32	*		3.00	1.32	*
Age		0.12	0.09	1.12	n.s.		0.24	0.16	n.s.		0.29	0.21	n.s.		0.29	0.21	n.s.
Female (1=Yes)		0.09	0.33	1.10	n.s.		0.78	0.60	n.s.		1.37	0.78	†		1.37	0.78	†
Parent or Primary Caregiver of a child (1=Yes)		-0.26	0.53	0.77	n.s.		-1.32	1.04	n.s.		0.74	1.33	n.s.		0.74	1.33	n.s.
Non-Hispanic White (1=Yes)		1.03	0.36	2.81	**		-0.82	0.66	n.s.		-1.15	0.86	n.s.		-1.15	0.86	n.s.
Days between pre-test and follow-up[28]		0.15	0.18	1.17	n.s.		0.29	0.32	n.s.		1.46	0.42	***		1.46	0.42	***
High school graduate or GED (1=Yes)		0.02	0.45	1.02	n.s.		0.92	0.98	n.s.		-0.12	1.14	n.s.		-0.12	1.14	n.s.
Pre-test value of outcome		N.A.	N.A.	N.A.	N.A.		0.46	0.14	**		0.65	0.05	***		0.65	0.05	***
Employed in the Last 6 months*Treatment		1.44	0.69	4.24	*		2.47	1.35	†		-1.62	1.65	n.s.		-1.62	1.65	n.s.

*** p<.001, ** p<.01, * p<.05, † p<.10, n.s. p>.10

[28] The variable measuring the number of days between pre-test and follow-up was standardized for all mixed effects models to improve model fit.

Table G4. Mean Civic Engagement Outcomes Predicted by Baseline Characteristics, Treatment, and Random Effects of Treatment by Program and Site. (N=214)

Student Level Fixed Effects (from Pre-test Characteristics)	***Currently Facing Charges or on Probation or Parole***	b	SE	Exp(B)	Sig?	***Mean Civic Engagement***	b	SE	Sig?
Treatment		0.42	0.92	1.52	n.s.		-1.43E-04	2.01E-01	n.s.
Criminal history (1=Yes)		3.72	1.09	41.22	***		3.14E-03	3.31E-01	n.s.
Employed in the last 6 months (1=Yes)		1.62	0.96	5.03	†		-6.05E-02	1.53E-01	n.s.
Age		0.11	0.18	1.12	n.s.		-1.11E-02	3.89E-02	n.s.
Female (1=Yes)		-0.46	0.75	0.63	n.s.		-8.37E-02	1.43E-01	n.s.
Parent or Primary Caregiver of a child (1=Yes)		-0.16	1.00	0.85	n.s.		6.53E-01	2.49E-01	n.s.
Non-Hispanic White (1=Yes)		-1.91	0.91	0.15	*		3.75E-01	1.54E-01	n.s.
Days between pre-test and follow-up[29]		-0.74	0.45	0.48	n.s.		6.82E-02	7.82E-02	n.s.
High school graduate or GED (1=Yes)		0.18	0.97	1.19	n.s.		1.79E-01	2.03E-01	n.s.
Pre-test value of outcome		N.A.	N.A.	N.A.	N.A.		5.19E-01	6.51E-02	***
Criminal History*Treatment		-3.06	1.51	0.05	*		-1.13E-01	4.28E-01	n.s.

*** p<.001, ** p<.01, * p<.05, † p<.10, n.s. p>.10

[29] The variable measuring the number of days between pre-test and follow-up was standardized for all mixed effects models to improve model fit.

G3. Exploratory TOT Analyses with a Subset of the Matched Sample.

An exploratory TOT dataset was created from the final matched sample by eliminating all treatment individuals who did not reach their target number of hours and any comparison youth who were matched exclusively to them. This process only eliminated 19 treatment members, so the composition of the ITT and TOT samples was very similar. The TOT sample demonstrated baseline equivalency for treatment and comparison youth on all variables except for disability with comparison youth being more likely to report having limitations due to mental, physical, or emotional problems (9% vs. 0%). See Table G5 below for more detailed information on the composition of the exploratory TOT sample.

Table G5. Baseline Equivalency (Unmatched Time 3 Sample and TOT Sample)

Variables	**Pre-match Sample Treat. (N=216)** Mean (SD) / Percent	**Pre-match Sample Comp. (N=139)** Mean (SD) / Percent	**Pre-match Mean Std. Diff**	**Pre-match Sig. Diff?**	**TOT Sample (2:1) Treat. (N=105)** Mean (SD) / Percent	**TOT Sample Comp. (N=79)** Mean (SD) / Percent	**TOT Mean Std. Diff**	**TOT Post-match Sig. Diff?**
Month Survey 1 was completed	6.6 (3.0)	7.6 (2.1)	-0.38	**	7.3 (2.8)	7.7 (2.2)	-0.16	n.s.
Number of days between Survey 1 and Survey 3	385.2 (90.9)	379.7 (82.9)	0.06	n.s	370.5 (97.6)	376.5 (86.9)	0.02	n.s.
Age	21.1 (2.3)	19.9 (2.1)	0.52	***	20.4 (2.3)	20.3 (2.2)	0.08	n.s.
Gender (Female)	42%	50%	-0.17	n.s.	46%	48%	-0.16	n.s.
Non-Hispanic White	43%	40%	0.06	n.s.	48%	44%	0.11	n.s.
Parent or primary caregiver of a child	15%	17%	-0.03	n.s.	11%	9%	0.11	n.s.
Primary caregiver of a parent or other adult	3%	4%	-0.02	n.s.	5%	2%	0.07	n.s.
Served on active duty in the military	3%	2%	0.04	n.s.	3%	1%	0.07	n.s.
High school graduate	82%	57%	0.55	***	77%	72%	-0.05	n.s.
Employed in the last 6 months	63%	58%	0.11	n.s.	60%	67%	-0.19	n.s.
Criminal history (charged or convicted)	23%	14%	0.25	*	12%	13%	0.04	n.s.
Received any government support	42%	42%	0.01	n.s.	34%	32%	-0.02	n.s.
Has a disability	10%	9%	0.03	n.s.	0%	9%	0.15	n.s.

*** p<.001, ** p<.01, * p<.05, † p<.10, n.s. p>.10

Overall, results with the treatment on treated sample generally resembled those in the ITT model, see Tables G6-G8 below. Overall, there were no statistically significant main effects of treatment. However, there were marginally significant main effects suggesting that treatment youth were more likely to complete a course (*b*=1.27, *SE*=0.76, *p*=0.09, *OR*=3.57) than comparison youth, but less likely to enroll in college or trade school (*b*=-1.39, *SE*=0.83, *p*=.09, *OR*=0.25).

Furthermore, two interaction effects from the ITT model were replicated with the TOT sample. In particular, the interaction between treatment and employment history on current employment was marginally significant (*b*=1.46, *SE*=0.76, *p*=.06, *OR*=4.32) in the same direction as before, as was the effect of treatment and criminal history on current involvement with the criminal justice system (*b*=-3.31, *SE*=1.95, *p*=.09, *OR*=0.04; See Figures G2 and G3 below). Additionally, one new interaction was statistically significant within the TOT sample— the interaction of treatment and employment history on longest period of employment (*b*=-3.56, *SE*=1.72, *p*=.04). Specifically, youth who had not been employed in the six months prior to the pre-test had longer periods of employment if they were in the treatment group than comparison group, while those who had been employed showed the opposite effect (see Figure G4 below).

Figure G1. Interaction of Treatment and Employment History on Current Employment in the TOT Sample

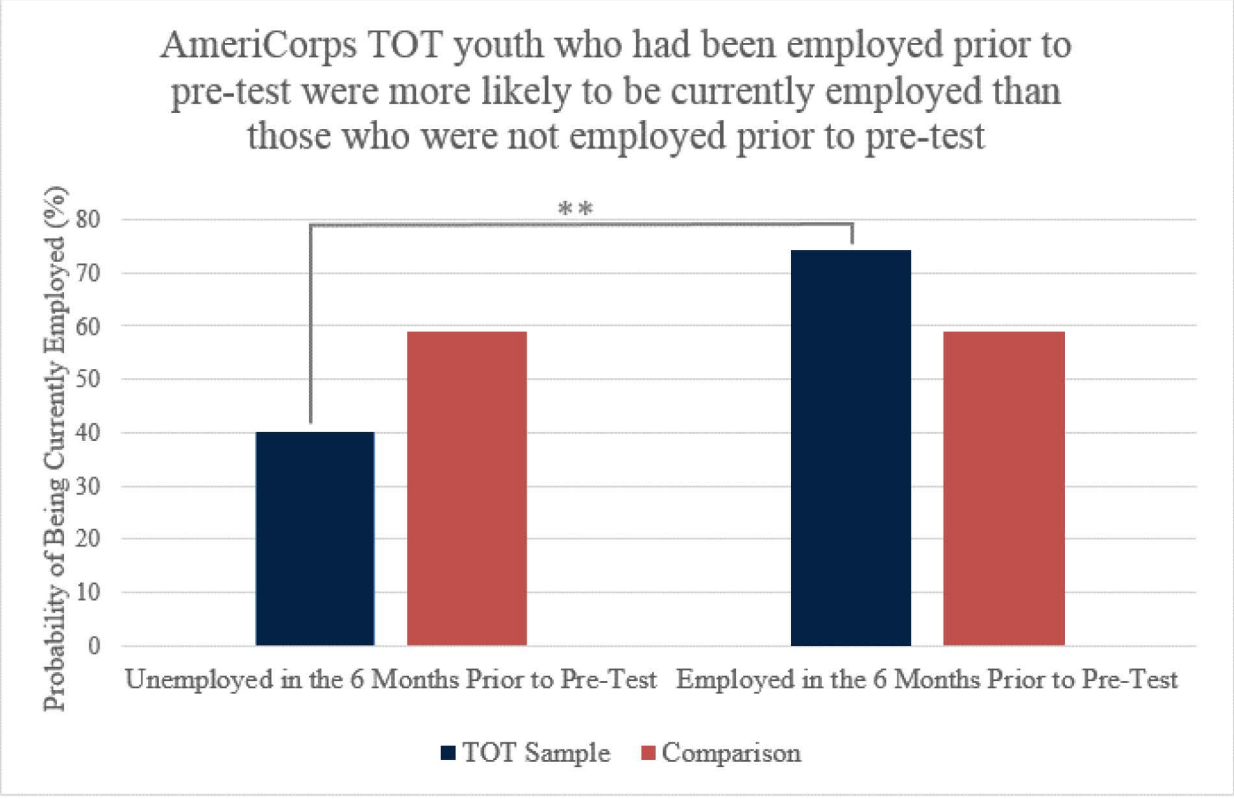


Figure G2. Interaction of Treatment and Employment History on Currently Facing Charges or on Probation or Parole in the TOT Sample

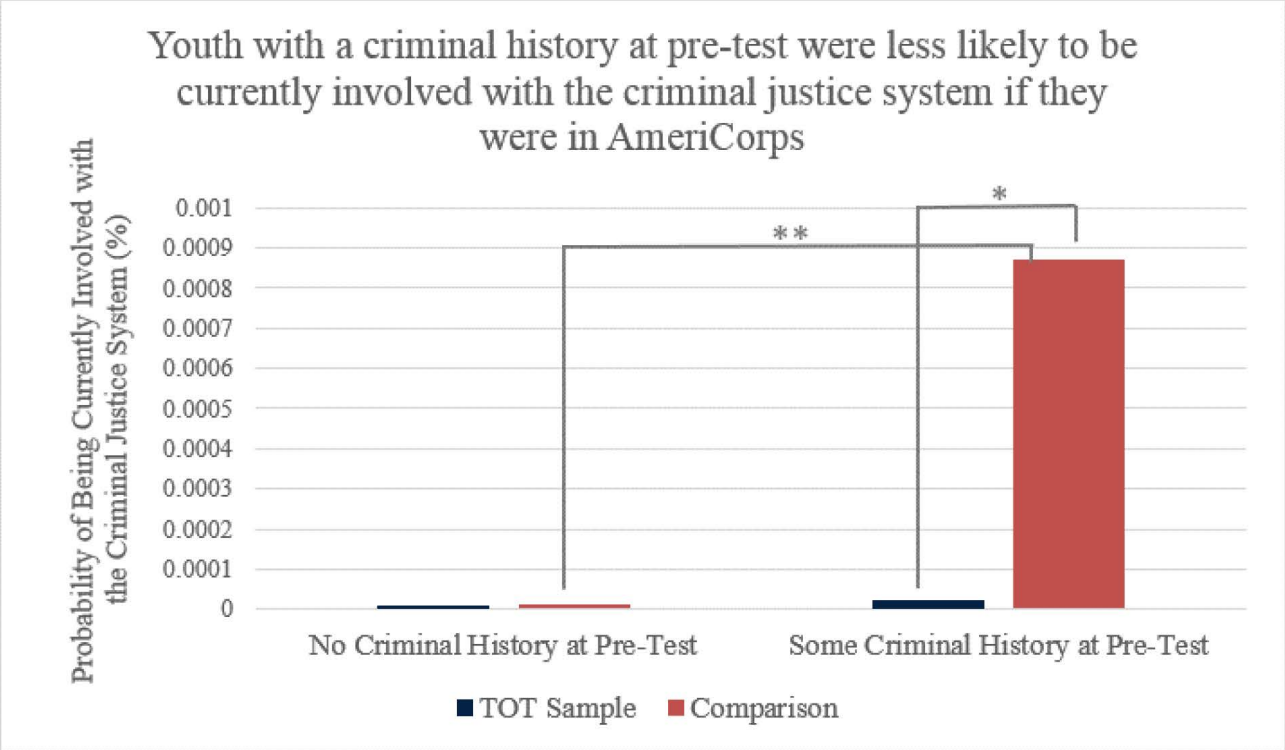


Figure G3. Interaction of Treatment and Employment History on Longest Period of Employment (in Months) in the TOT Sample

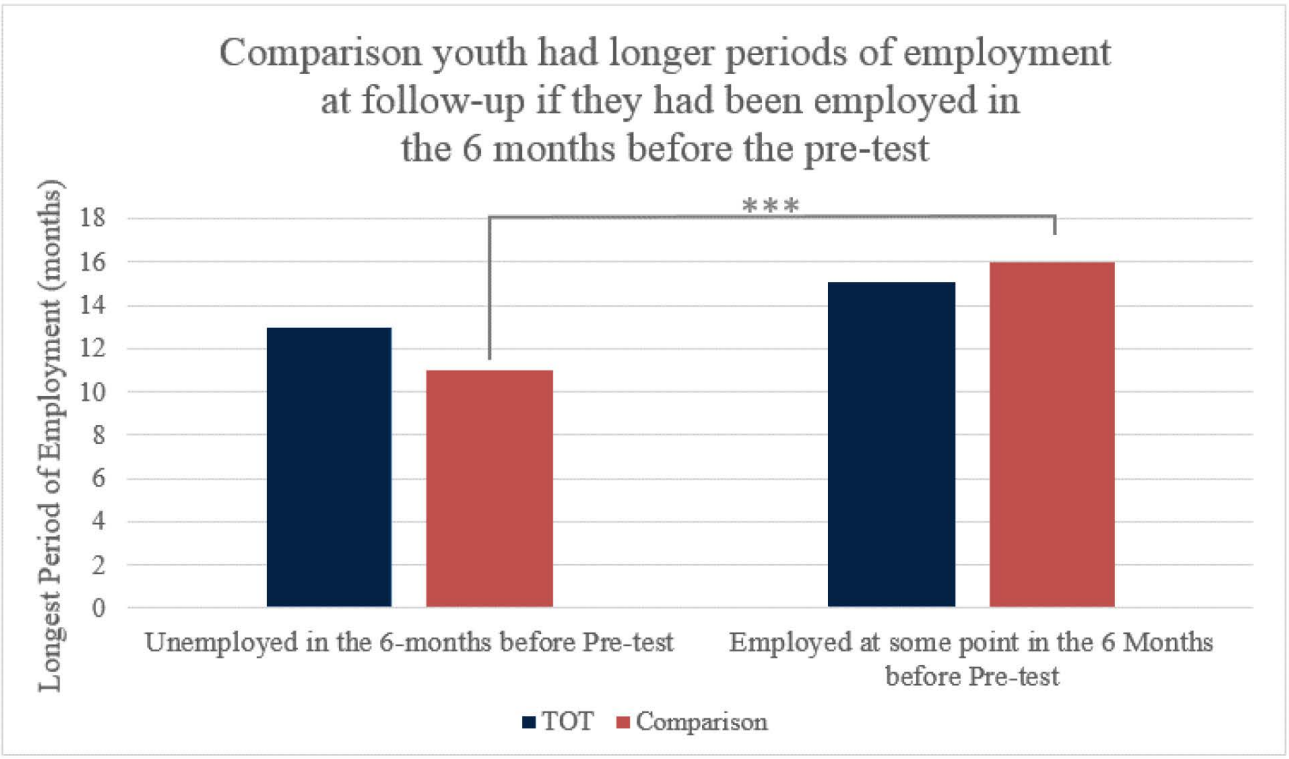


Table G6. Exploratory Educational Outcomes Predicted by Baseline Characteristics, Treatment, and Random Effects of Treatment by Program and Site for Treatment Participants Who Completed Their Target Number of Hours and Their Matches in The Comparison Group (N=184)

Student Level Fixed Effects (from Pre-test Characteristics)	***Obtained HS Diploma or GED*** [30]	b	SE	Odds Ratio (Exp(B))	Sig?	***Completed a Course in High School or College in the Last 6 months***	b	SE	Odds Ratio (Exp(B))	Sig?	***Enrolled in college, trade school, or certification course***	b	SE	Odds Ratio (Exp(B))	Sig?
Treatment		-0.19	0.76	0.83	n.s.		1.27	0.76	3.57	†		-1.39	0.82	0.25	†
Criminal history (1=Yes)		-1.04	1.09	0.35	n.s.		-0.56	0.64	0.57	n.s.		0.26	0.59	1.29	n.s.
Employed in the last 6 months (1=Yes)		0.29	0.75	1.34	n.s.		0.84	0.43	2.32	†		0.18	0.39	1.19	n.s.
Age		0.44	0.22	1.55	†		-0.24	0.11	0.79	*		-0.32	0.11	0.73	**
Female (1=Yes)		0.33	0.84	1.38	n.s.		0.03	0.39	1.03	n.s.		0.44	0.38	1.55	n.s.
Parent or Primary Caregiver of a child (1=Yes)		-3.21	1.56	0.04	*		-0.02	0.77	0.98	n.s.		0.87	0.64	2.38	n.s.
Non-Hispanic White (1=Yes)		0.60	0.83	1.83	n.s.		0.32	0.41	1.37	n.s.		-0.13	0.41	0.88	n.s.
Days between pre-test and follow-up[31]		0.12	0.46	1.12	n.s.		-0.68	0.22	0.51	**		-0.35	0.21	0.71	†
High school graduate or GED (1=Yes)		N.A.	N.A.	N.A.	N.A.		0.08	0.73	1.08	n.s.		0.66	0.77	1.94	n.s.
Pre-test value of outcome		N.A.	N.A.	N.A.	N.A.		1.15	0.40	3.14	**		N.A.	N.A.	N.A.	N.A.
HSGrad*Treatment		N.A.	N.A.	N.A.	N.A.		-1.40	0.89	0.25	0		1.13	0.97	3.08	n.s.

*** p<.001, ** p<.01, * p<.05, † p<.10, n.s. p>.10

[30] The multi-level model predicting high school graduation status at follow-up did not fully converge (most likely because of the small number of youth who were not high school graduates at pre-test, N=46), To address this, a logistic regression model was run instead of the originally proposed multi-level model and predicted high school graduation status at post-test strictly among youth who were not high school graduates at pre-test.

[31] The variable measuring the number of days between pre-test and follow-up was standardized for all mixed effects models to improve model fit.

Table G7. Exploratory Employment Outcomes Predicted by Baseline Characteristics, Treatment, and Random Effects of Treatment by Program and Site for Treatment Participants Who Completed Their Target Number of Hours and Their Matches in The Comparison Group (N=184)

Student Level Fixed Effects (from Pre-test Characteristics)	***Employed Currently***	b	SE	Odds Ratio (Exp(B))	Sig?	***Hourly Wage***	b	SE	Sig?	***Longest Employment***	b	SE	Sig?	***Number of Job Search Activities Completed***	b	SE	Sig?
Treatment		-0.75	0.59	0.47	n.s.		-2.16	1.42	n.s.		1.61	1.56	n.s		0.06	0.59	n.s.
Criminal history (1=Yes)		-0.83	0.51	0.44	n.s.		-0.24	1.04	n.s.		1.69	1.27	n.s.		-0.77	0.55	n.s.
Employed in the last 6 months (1=Yes)		-0.01	0.55	0.99	n.s.		-0.19	1.13	n.s.		4.98	1.37	***		-0.09	0.57	n.s.
Age		0.04	0.10	1.04	n.s.		0.15	0.19	n.s.		0.09	0.21	n.s.		0.06	0.10	n.s.
Female (1=Yes)		0.28	0.37	1.32	n.s.		0.81	0.70	n.s.		1.67	0.81	*		0.21	0.35	n.s.
Parent or Primary Caregiver of a child (1=Yes)		-0.46	0.60	0.63	n.s.		-0.94	1.31	n.s.		0.16	1.43	n.s.		-0.52	0.61	n.s.
Non-Hispanic White (1=Yes)		1.09	0.41	2.98	**		-0.82	0.75	n.s.		-1.00	0.87	n.s.		0.01	0.40	n.s.
Days between pre-test and follow-up[32]		0.32	0.20	1.38	n.s.		0.25	0.37	n.s.		1.59	0.42	***		-0.38	0.22	n.s.
High school graduate or GED (1=Yes)		0.62	0.50	1.86	n.s.		1.27	1.11	n.s.		2.30	1.15	*		0.15	0.57	n.s.
Pre-test value of outcome		N.A.	N.A.	N.A.	N.A.		0.49	0.16	**		0.62	0.06	***		0.37	0.09	***
Employed in the last 6 months*Treatment		1.46	0.76	4.32	†		2.26	1.55	n.s.		-3.56	1.72	*		0.54	0.74	n.s.

*** p<.001, ** p<.01, * p<.05, † p<.10, n.s. p>.10

[32] The variable measuring the number of days between pre-test and follow-up was standardized for all mixed effects models to improve model fit.

Table G8. Exploratory Civic Engagement Outcomes Predicted by Baseline Characteristics, Treatment, and Random Effects of Treatment by Program and Site for Treatment Participants Who Completed Their Target Number of Hours and Their Matches in The Comparison Group (N=184)

Student Level Fixed Effects (from Pre-test Characteristics)	***Currently Facing Charges or on Probation or Parole***	b	SE	Odds Ratio (Exp(B))	Sig?	***Mean Civic Engagement***	b	SE	Sig?
Treatment		-0.42	1.13	0.66	n.s.		-.03	0.21	n.s.
Criminal history (1=Yes)		4.14	1.41	62.94	**		0.17	0.34	n.s.
Employed in the last 6 months (1=Yes)		1.77	1.24	5.90	n.s.		0.03	0.16	n.s.
Age		0.18	0.23	1.20	n.s.		-0.01	0.04	n.s.
Female (1=Yes)		-2.11	1.27	0.12	†		-0.09	0.15	n.s.
Parent or Primary Caregiver of a child (1=Yes)		0.89	1.21	2.44	n.s.		0.74	0.27	**
Non-Hispanic White (1=Yes)		-20.68	1024.00	0.00	n.s.		0.45	0.16	**
Days between pre-test and follow-up[33]		-0.92	0.61	0.40	n.s.		0.06	0.08	n.s.
High school graduate or GED (1=Yes)		0.18	1.10	1.19	n.s.		0.34	0.22	n.s.
Pre-test value of outcome		N.A.	N.A.	N.A.	N.A.		0.52	0.07	***
Criminal History*Treatment		-3.31	1.95	0.04	†		-0.05	0.46	n.s.

*** p<.001, ** p<.01, * p<.05, † p<.10, n.s. p>.10

[33] The variable measuring the number of days between pre-test and follow-up was standardized for all mixed effects models to improve model fit.

G4. Dosage Analyses Using the full Unmatched Treatment-only Sample.

To maximize the utility of the data collected from treatment members and increase power, an exploratory analysis was conducted looking at dosage effects within the full treatment sample at follow-up. The results indicated that there were no statistically significant main effects of either AmeriCorps program completion or the total number of hours served. However, there was one marginally significant interaction between criminal history and total number of hours served on mean civic engagement (B=.001, SE=.001, p=.086), with youth who had a criminal history and served more hours showing slightly higher levels of civic engagement, see Table G9 below.

Table G9. Mean levels of Civic Engagement Predicted by Baseline Characteristics, Total Number of AmeriCorps Hours Served, and the Interaction of Criminal History and Total Number of AmeriCorps Hours Served (N=202)

Predictor Variables	Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
Constant	2.044	.919		2.223	.027
Total # of AC Hours Served	.000	.000	-.042	-.448	.655
Criminal History (either Charged or Convicted)	-.747	.561	-.265	-1.333	.184
Employed in the Past 6 months Y/N T1	-.175	.167	-.071	-1.050	.295
Age in Years as of 12/07/15	.004	.042	.007	.089	.929
Gender - Revised to Male/Female	-.033	.163	-.014	-.200	.841
Parent - Revised to Y/N	.360	.244	.109	1.476	.142
Non-Hispanic White	.093	.171	.039	.541	.589
Days From S1 to S3	-.001	.001	-.055	-.638	.524
High School Graduate at Pre-test	.228	.269	.072	.847	.398
Mean of Civic Engagement	.418	.072	.390	5.778	.000
Criminal History * Total Number of AC Hours Served	.001	.001	.341	1.725	.086

G5. Within-Subjects Analyses from Pre-Test to Post-Test and Follow-up Using the Unmatched Treatment-only Sample.

Table G10. Exploratory Analysis of Differences from Pre-test to Post-Test and Pre-Test to Follow-up for All Treatment Individuals in The Unmatched Study Sample (N=216)

Outcome Area	Outcome Variables in Treatment-only Sample	Pre-test Mean/ Percent	Pre-test SD	Post-test Mean/ Percent	Post-test SD	Follow-up Mean/ Percent	Follow-up SD	Pre-Post Sig?[34]	Pre-Follow-up Sig?
	Primary Outcome Variables								
Education	Has a high school diploma or GED	72%		83%		92%		***	***
Education	Completed a course in high school, college, or alternative school	33%		28%		26%		*	*
Education	Enrolled in college, trade school, or certification course	22%		21%		25%		n.s	n.s.
Employment	Employed currently	19%		25%		62%		*	***
Employment	Hourly income (in dollars)	9.61	3.04	9.64	2.93	11.00	3.28	n.s	***
Employment	Longest employment experience (in months)	9.81	8.49	11.72	8.78	13.34	8.37	***	***
Employment	Number of job search activities completed in the last 6 months (0-9)[1]	5.01	2.43	5.07	2.40	5.25	2.23	n.s.	†
Civic Engagement	Currently facing charges or on probation or parole	15%		15%		7%		n.s.	n.s.
Civic Engagement	Mean frequency of civic engagement activities (1-6)[2]	3.15	1.19	3.41	1.16	3.25	1.20	**	n.s.
	Secondary Outcome Variables								
Education	Took a GED test	4%		5%		4%		n.s.	*n.s.*
Education	Intends to pursue education beyond high school (1-5)[3]	4.28	0.92	4.31	0.84	4.19	0.99	n.s.	*
Education	Confidence in the ability to obtain postsecondary educ. (1-5)[4]	4.08	0.77	4.11	0.76	4.15	0.79	n.s.	n.s.
Education	Applied to college or trade school	13%		11%		14%		n.s.	n.s.
Employment	Employed in last 6 months	57%		59%		75%		n.s	**
Employment	Number of types of employment experiences (0-7)[5]	2.46	1.54	2.72	1.70	3.19	1.64	*	***

[34] Significance testing was conducted using paired Wilcoxin tests for continuous outcomes from pre-test to post-test within participants (since they were non-normally distributed) and using paired McNemar tests for dichotomous outcomes. *p<.05, **p<.01, ***p<.00, †p<.10, n.s. p>.10

Outcome Area	Outcome Variables in Treatment-only Sample	Pre-test Mean/ Percent	Pre-test SD	Post-test Mean/ Percent	Post-test SD	Follow-up Mean/ Percent	Follow-up SD	Pre-Post Sig?[34]	Pre-Follow-up Sig?
Employment	Confidence in job search skills (1-5)[4]	3.95	0.69	4.00	0.72	4.09	0.73	n.s.	**
Employment	Average perceptions of barriers to employment (1-3)[6]	1.60	0.38	1.52	0.38	1.45	0.37	**	**
Civic Engagement	Ever been convicted of a crime	24%		29%		25%		***	***
Civic Engagement	Confidence in ability to access community resources (1-5)[4]	3.70	0.92	3.77	0.95	3.75	0.98	†	n.s.
Civic Engagement	Average sense of community (1-5)[3]	3.82	0.76	3.89	0.70	3.85	0.82	n.s.	n.s.
Civic Engagement	Trust in community members (1-4)[7]	2.26	0.78	2.32	0.76	2.47	0.78	n.s.	*
Civic Engagement	Average self-efficacy to solve personal problems (1-5)[3]	4.02	0.61	4.04	0.68	4.02	0.75	n.s.	†
Civic Engagement	Average self-efficacy to solve community problems (1-5)[8]	3.92	0.71	3.94	0.80	3.99	0.85	n.s.	*
Civic Engagement	Received any form of government support[9]	46%		50%		43%		n.s.	n.s.
Civic Engagement	Received any form of community support	35%		32%		35%		n.s.	n.s.

Table Notes.

[1] Up to 9 of the following: Submitted job application, wrote/revised resume, interviewed, contacted employer, had informational interview, looked for full-time work, looked for part-time work, looked for internship/apprenticeship, looked for volunteer position.

[2] 1=Not at all, 2= Less than once a month, 3= Once a month, 4=A few times a month, 5=A few times a week, 6=Basically every day.

[3] 1=Strongly Agree, 2= Disagree, 3=Neither Agree nor Disagree, 4=Agree, 5=Strongly Agree

[4] 1= No confidence at all, 2=Very little confidence, 3=Moderate confidence, 4=Much confidence, 5=Complete Confidence

[5] Up to 7 of the following: paid part-time job, paid full-time job, paid internship, unpaid internship, volunteer work, program stipend, paid babysitting/yardwork

[6] 1=Not barriers, 2=Barriers that can be overcome, 3=Large barriers

[7] 4 = Can trust all of the people in my neighborhood, 3 = Most, 2 = Some, 1 = None of the people

[8] 1 = I definitely could not do this , 2 = I could not do this, 3 = Not sure, 4 = I probably could do this, 5 = I definitely could do this.

[9] Government supports included currently accessing: food assistance (such as WIC or SNAP), health care assistance (such as Medicaid or other health insurance), housing assistance (such as housing vouchers), or other financial or practical assistance (such as TANF or child care assistance programs).

[10] Community supports included using or visiting any of the following: local unemployment division, housing center (for help with finding housing), job center, crisis center, homeless shelter, food bank, community health clinic, adult school/community college extension programs, mutual support or other assistance programs (e.g., AA, NA, AlAnon, grief support groups).

Appendix H: Program-level Variability in Survey Implementation and Program Impacts

This appendix describes the programs that participated in the Opportunity Youth Evaluation Bundle, the comparison group recruitment strategies evaluators and programs devised during pre-implementation TA, and the choices programs made in consultation with evaluators and community partners about preferred survey modalities (paper, online, or phone).

H1. Description of Participating Programs

All the programs participating in the evaluation bundle are AmeriCorps State and National programs that actively enlist opportunity youth to serve as AmeriCorps members.[35]

The evaluators screened and selected programs for the evaluation bundle from a short list of candidate programs identified by CNCS. The final set of programs included in the bundle met a list of screening criteria that take into account each program's ability to effectively participate in an evaluation bundle with similar activities, outputs, and outcomes. Twelve of the 19 programs recruited for the bundle are affiliates ("Corps") of a national AmeriCorps program, The Corps Network. The parent organization was instrumental in facilitating the screening and selection process for these twelve sites.

All programs met screening criteria for minimum number of OY participants; prospects to supply comparison group members, either through internal or external recruitment; engaging members in service activities consistent with the evaluation's research questions; extant data collection systems that could support the evaluation; and availability of program staff to participate in TA and evaluation tasks.

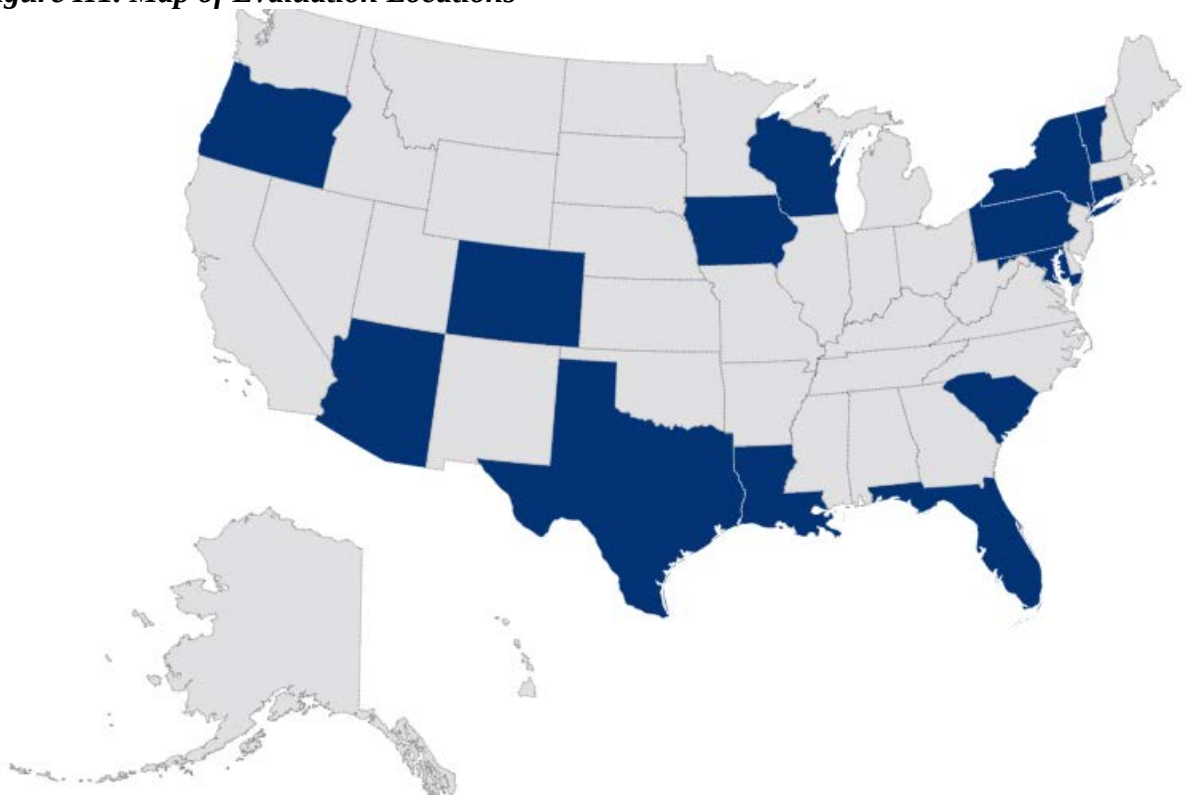
Table H1 and Figure H1 below provide information on the location of each program in the bundle and each program's focus of intended program outcomes. Seventeen of the 19 programs focus on intended outcomes in the areas of both education and employment. The other two programs focus on a single outcome area, one on education (Genesee County) and one on employment (UMBC Choice Program). All 19 programs were deemed compatible with the requirements of the bundle, including the outcomes to be measured in the evaluation. OY members enrolled in the programs served in positions of varying lengths, including full-time (1,700 hours), half time (900 hours), reduced half time (675 hours), quarter time (450 hours), and minimum time (300 hours). Members were generally required to complete these hours within a 12-month period. See Table H2 below for a more detailed breakdown of the AmeriCorps term lengths offered by site.

[35] In some cases, program participants consist entirely of OY. In other cases, the program has OY and non-OY program components.

Table H1. Evaluation Bundle: Location and Outcome Focus

Program Name	Location	Education Focus	Employment Focus
American Youthworks	Austin, TX	X	X
AZ Conservation Corps (AZCC)- Flagstaff	Flagstaff, AZ	X	X
AZ Conservation Corps (AZCC)- Tucson	Tucson, AZ	X	X
City of Davenport	Davenport, IA	X	X
Civic Works (GOSV and TCN)	Baltimore, MD	X	X
Genesee County	Batavia, NY	X	
Greater Miami Service Corps (GMSC)	Miami, FL	X	X
Green City Force (GCF)	Brooklyn, NY	X	X
Knox Inc.	Hartford, CT	X	X
Limitless Vistas	New Orleans, LA	X	X
Northwest Youth Corps (NWYC)	Eugene, OR	X	X
Operation Fresh Start (OFS)	Madison, WI	X	X
PowerCorps PHL	Philadelphia, PA	X	X
Renewal Unlimited	Portage, WI	X	X
Sustainability Institute	North Charleston, SC	X	X
UMBC Choice Program	Baltimore, MD		X
Vermont Youth Conservation Corps (VYCC)	Richmond, VT	X	X
Western Colorado Conservation (WCCC)	Grand Junction, CO	X	X
Western Dairyland	Independence, WI	X	X

Figure H1. Map of Evaluation Locations



Participant recruitment targets were 626 for treatment and 1,252 for comparison. In practice, programs recruited 482 treatment group members (76.9% of target) and 590 comparison group members (47.1% of target) in total (see Table H2 below). These statistics are indicative of the recruitment challenges many programs faced, particularly with regard to comparison group recruitment.

While programs came reasonably close to achieving treatment group targets, the inability to reach the treatment group recruitment target stems from a combination of program applicants who did not meet the study's eligibility requirements and programs simply not meeting enrollment targets.

Table H2. Evaluation Bundle AmeriCorps Term Lengths Offered and Anticipated and Obtained Pre-test Sample Sizes

Site #	**300 hrs. *(Minimum time)* Term offered (% within site)**	**450 hrs. *(Quarter-time)* Term offered (% within site)**	**675 hrs. *(Reduced Half-Time)* Term offered (% within site)**	**900 hrs. *(Half-Time)* Term offered (% within site)**	**1,700 hrs. *(Full-Time)* Term offered (% within site)**	**Targeted Treatment and Comparison Group Sample**	**Actual Treatment and Comparison Group Sample**
Site 1	X (100%)					40 / 80	37 / 67
Site 2	X (14%)			X (86%)		18 / 36	14 / 21
Site 3					X (100%)	15 / 30	7 / 0
Site 4		X (4%)	X (34%)	X (62%)		65 / 130	85 / 77
Site 5		X (11%)	X (16%)	X (74%)		24 / 48	19 / 4
Site 6					X (100%)	6 /12	5 / 4
Site 7	X (23%)	X (15%)	X (31%)	X (31%)		20 / 40	13 / 14
Site 8		X (50%)	X (8%)	X (35%)	X (8%)	42 / 104	26 / 81
Site 9				X (4%)	X (96%)	41 / 82	23 / 14
Site 10				X (100%)		42 / 84	15 / 50
Site 11		X (8%)	X (4%)	X (88%)		34 / 68	25 / 10
Site 13 [36]			X (91%)	X (9%)		46 / 92	48 / 78
Site 14					X (100%)	12 / 24	18 / 10
Site 15				X (100%)		24 / 48	10 / 6
Site 16		X (25%)		X (75%)		48 / 96	60 / 26
Site 17				X (100%)		50 / 100	22 / 73
Site 18				X (100%)		16 / 32	8 / 25
Site 19		X (26%)	X (42%)	X (32%)		59 / 118	31 / 27
Site 20			X (88%)	X (13%)		24 /48	16 / 3
Total	**9%**	**10%**	**23%**	**49%**	**10%**	**626 / 1,252**	**482 / 590**

[36] Please note there is no Site #12 in the current evaluation.

Challenges in meeting comparison group recruitment targets stem from the novelty of participating in an impact evaluation for many programs and community partners (i.e., they were doing it for the first time), overly optimistic estimates of over-recruitment, and the difficulty of finding eligible youth at community partners. Collectively, programs recruited less than half the comparison group youth needed to achieve a 2-to-1 ratio for statistical matching. Nine programs did not meet a 1-to-1 ratio, and one small program was unable to recruit any comparison group youth. Inability to reach comparison group recruitment targets meant that some youth in the treatment group could not be matched with a comparison youth, effectively excluding the treatment group youth from inferential analysis.

H2. Comparison Group Recruitment Strategies by Program

Almost every program in the bundle proposed (as documented in the IRP) to recruit comparison group members through a combination of internal ("over-recruitment") and external sources (community partners). The main difference amongst the programs was with regard to the proportion of comparison group youth that programs anticipated getting from internal versus external sources. Some programs acknowledged at the outset that they expected to get very few comparison group participants through over-recruitment and therefore anticipated needing to rely heavily on community partners to supply these participants. Also, programs frequently overestimated their ability to source comparison group participants internally, and then found it necessary to rely more heavily than anticipated on community partners (who were not always able to cover a shortfall in recruitment).

While all programs in the bundle were able to identify at least one community partner during the planning phase from which they proposed to recruit at least some OY for the comparison group, programs varied in their ability to secure comparison group youth from these sources during the implementation phase.

Programs often overestimated their ability to obtain comparison group youth through over-recruitment. In fact, some programs encountered unexpected challenges in meeting their member enrollment targets, and thus had relatively few or even no "extra" applicants to assign to the comparison group. In general, the majority of comparison group youth were recruited from community partners rather than through over-recruitment. Working with community partners to recruit comparison group youth also proved to be a valuable learning experience for many programs, as they attempted for the first time to work with these partners to obtain OY for participation in the comparison group.

Programs that had the greatest success recruiting comparison group youth from community partners typically either had existing relationships with these partners for the referral of youth to their programs as AmeriCorps applicants or had a formal relationship with the partner through a mutual parent organization. For example, City of Davenport, which recruits members from the local school district, was able to rely on the school district to identify non-member OY to participate in the comparison group. As another example, Baltimore-based Civic Works has close connections with the Youth Opportunity program (a community partner) through their mutual relationship with the Mayor's Office. Programs that were unable to obtain sufficient comparison group youth from the set of community partners identified at the planning stage reached out to additional community partners during the implementation phase to find more OY.

This strategy typically resulted in no more than ten percent additional comparison group recruitment.

Table H3. Planned and Actual Comparison Group Recruitment Strategies by Site

Program	Planned: Over-Recruitment	Planned: Community Partners	Planned: Schools	Actual: Over-Recruitment	Actual: Community Partners	Actual: Schools
American Youthworks	X	X		X	X	
AZCC Flagstaff	X			X		
AZCC Tucson	X	X		X	X	
Davenport	X	X	X	X	X	X
Civic Works	X	X		X	X	
Genesee	X	X		X	X	
GMSC	X	X	X	X		
GCF	X	X		X		
Knox Inc.		X				
Limitless Vistas	X	X		X	X	
NWYC	X			X		
OFS	X	X	X	X		
PowerCorps	X	X		X	X	
Renewal Unlimited	X	X			X	
Sustainability Institute	X	X			X	
UMBC	X	X	X		X	
VYCC	X	X		X		
WCCC	X			X		
Western Dairyland	X	X			X	

H3. Data Collection Modalities by Program

Table H4 summarizes each program's preferences for how to collect data from treatment and comparison group members. Decisions about data collection modalities were made collaboratively through TA to plan participant recruitment and survey implementation, and were recorded in the IRP and SIP, respectively.

Almost every program chose to rely on paper surveys to collect pre-test and post-test data from AmeriCorps members (treatment group). This decision is tied to the choice most programs made to integrate the pre- and post-surveys into member application and exit paperwork, respectively. Even in the case of American Youthworks, the program eventually opted to administer some pre- and post-surveys of members on paper.

Table H4. Preferred Data Collection Modality by Program

Program	Treatment: Paper	Treatment: Online	Treatment: Phone	Comparison: Paper	Comparison: Online	Comparison: Phone
American Youthworks		X	X		X	
AZCC Flagstaff	X			X	X	X
AZCC Tucson	X			X	X	X
Davenport	X			X	X	X
Civic Works	X			X		
Genesee	X			X	X	X
GMSC	X			X		
GCF	X			X		
Knox Inc.	X			X		
Limitless Vistas	X			X		
NWYC	X				X	X
OFS	X			X		
PowerCorps	X			X	X	
Renewal Unlimited	X			X		
Sustainability Institute	X			X		
UMBC	X			X	X	X
VYCC	X				X	
WCCC	X				X	
Western Dairyland	X			X		

The preferred survey modalities for comparison group youth obtained through over-recruitment typically corresponded to treatment group modalities, while comparison group youth recruited from community partners took the survey in a variety of modalities. This decision reflects a need to maximize outreach options for youth that typically were not directly accessible to program staff, who, in many cases, took on the task of administering the survey to comparison group youth.

As noted elsewhere in this report, paper surveys accounted for the vast majority of surveys collected, and yielded the highest response rates for treatment and comparison group youth.

Some programs found they could not recruit sufficient comparison group youth within the specified timeframe (within one month of the member recruitment period). Evaluators typically permitted these programs to extend the comparison group recruitment period a few weeks beyond the original window to maximize opportunities for programs to achieve the 2-to-1 ratio of treatment to comparison group members called for by the evaluation. The dates when pre-test OY took surveys were carefully tracked to identify comparison group youth recruited outside the standard window and to take this difference into account in the analysis.

Table H5. Number Surveys Completed by Modality and Program by Survey Wave

Site	Paper S1	Paper S2	Paper S3	Online S1	Online S2	Online S3	Phone S1	Phone S2	Phone S3
Overall	***993***	***246***	***45***	***70***	***164***	***295***	***9***	***24***	***15***
Site 1	103	30	12	1	31	22	0	7	0
Site 2	29	10	0	6	12	15	0	0	0
Site 3	7	3	0	0	2	2	0	0	0
Site 4	162	59	31	0	3	7	0	2	1
Site 5	23	10	0	0	1	4	0	0	3
Site 6	9	0	0	0	2	2	0	0	0
Site 7	27	5	0	0	3	6	0	0	0
Site 8	107	8	0	0	1	17	0	0	0
Site 9	19	18	0	18	11	25	0	0	0
Site 10	62	11	0	3	15	25	0	1	2
Site 11	35	17	0	0	2	21	0	1	2
Site 13	126	16	2	0	11	27	0	3	2
Site 14	28	10	0	0	3	9	0	0	2
Site 15	16	0	0	0	5	5	0	2	0
Site 16	61	22	0	21	29	49	4	2	0
Site 17	95	0	0	0	9	12	0	1	2
Site 18	33	3	0	0	1	3	0	2	1
Site 19	32	24	0	21	16	38	5	1	0
Site 20	19	0	0	0	7	6	0	2	0

H4. Response Rates by Program, by Condition, and by Survey Wave

Table H6. Response Rates by Program, by Condition, and by Survey Wave

Program	Survey 1 Treat	Survey 1 Comp	Survey 1 Total	Survey 2 Treat	Survey 2 Treat RR %	Survey 2 Comp	Survey 2 Comp RR %	Survey 3 Treat	Survey 3 Treat RR %	Survey 3 Comp	Survey 3 Comp RR %
Overall	***482***	***590***	***1,072***	***294***	***61%***	***141***	***24%***	***216***	***45%***	***139***	***24%***
Site 1	37	67	104	30	81%	38	57%	12	32%	22	33%
Site 2	14	21	35	8	57%	14	67%	6	43%	9	43%
Site 3	7	0	7	5	71%	0	N/A	2	29%	0	N/A
Site 4	85	77	162	59	69%	6	8%	37	44%	2	3%
Site 5	19	4	23	10	53%	1	25%	6	32%	1	25%
Site 6	5	4	9	2	40%	0	0%	1	20%	1	25%
Site 7	13	14	27	7	54%	1	7%	4	31%	2	14%
Site 8	26	81	107	8	31%	1	1%	9	35%	8	10%
Site 9	23	14	37	22	96%	7	50%	19	83%	6	43%
Site 10	25	10	35	19	76%	1	10%	18	72%	5	50%
Site 11	15	50	65	11	73%	16	32%	8	53%	19	38%
Site 13	48	78	126	17	35%	13	17%	15	31%	16	21%
Site 14	18	10	28	12	67%	1	10%	9	50%	2	20%
Site 15	10	6	16	5	50%	2	33%	4	40%	1	17%
Site 16	60	26	86	37	62%	16	62%	32	53%	17	65%
Site 17	22	73	95	3	14%	7	10%	5	23%	9	12%
Site 18	8	25	33	3	38%	3	12%	1	13%	3	12%
Site 19	31	27	58	27	87%	14	52%	22	71%	16	59%
Site 20	16	3	19	9	56%	0	0%	6	38%	0	0%

H5. Propensity Score Match Rate and Final Sample Sizes by Program

Table H7. Propensity Score Match Rate and Final Sample Sizes by Program

Program Name	Pre-match Sample Total	Pre-match Sample Treat	Pre-match Sample Comp	Post-match Sample Total	Post-match Sample Treat	Post-match Sample Comp	Percentage of Youth Retained Total	Percentage of Youth Retained Treat	Percentage of Youth Retained Comp
Site 1	34	12	22	24	12	12	71%	100%	55%
Site 2	15	6	9	10	5	5	67%	83%	56%
Site 3	2	2	0	0	0	0	0%	0%	N/A
Site 4	39	37	2	6	4	2	15%	11%	100%
Site 5	7	6	1	0	0	0	0%	0%	0%
Site 6	2	1	1	2	1	1	100%	100%	100%
Site 7	6	4	2	5	3	2	83%	75%	100%
Site 8	17	9	8	15	9	6	88%	100%	75%
Site 9	25	19	6	14	9	5	56%	47%	83%
Site 10	23	18	5	9	6	3	39%	33%	60%
Site 11	27	8	19	15	8	7	56%	100%	37%
Site 13	31	15	16	20	11	9	65%	73%	56%
Site 14	11	9	2	6	4	2	55%	44%	100%
Site 15	5	4	1	3	2	1	60%	50%	100%
Site 16	49	32	17	40	24	16	82%	75%	94%
Site 17	14	5	9	9	5	4	64%	100%	44%
Site 18	4	1	3	0	0	0	0%	0%	0%
Site 19	38	22	16	36	21	15	95%	96%	94%
Site 20	6	6	0	0	0	0	0%	0%	N/A
Total	**355**	**216**	**139**	**214**	**124**	**90**	**60%**	**57%**	**65%**

H6. Differences in Program Outcomes at Follow-up for Treatment Members, by Site

Table H8. Education Outcomes by Site within Treatment Youth

Site	Site N for S3	*Has a HS Diploma or GED*	Pre-Test Percent	Follow-up Percent	MSD	*Completed a course in high school or college in the last 6 months*	Pre-Test Percent	Follow-up Percent	MSD	*Enrolled in college, trade school, or certification course*	Pre-Test Percent	Follow-up Percent	MSD
Site 1	12		0%	22%	0.76		42%	50%	0.17		0%	17%	0.63
Site 2	6		50%	100%	1.41		33%	33%	0.00		0%	33%	1.00
Site 3 [37]	2												
Site 4	37		97%	100%	0.24		32%	22%	-0.25		30%	24%	-0.12
Site 5	6		0%	50%	1.41		17%	0%	-0.63		0%	0%	0.00
Site 6	1												
Site 7	4		100%	100%	0.00		50%	50%	0.00		50%	50%	0.00
Site 8	9		56%	78%	0.49		33%	22%	-0.25		33%	22%	-0.25
Site 9	19		100%	100%	0.00		16%	16%	0.00		11%	16%	0.16
Site 10	8		100%	100%	0.00		38%	38%	0.00		63%	38%	-0.52
Site 11	18		100%	100%	0.00		50%	0%	-1.41		17%	17%	0.00
Site 13	15		79%	86%	0.19		20%	27%	0.16		20%	27%	0.16
Site 14	9		100%	100%	0.00		33%	0%	-1.00		33%	11%	-0.55
Site 15	4		50%	100%	1.41		25%	25%	0.00		25%	25%	0.00
Site 16	32		97%	100%	0.25		41%	34%	-0.13		38%	34%	-0.07
Site 17	5		0%	80%	2.83		0%	60%	1.73		0%	0%	0.00
Site 18	1												
Site 19	22		95%	95%	0.00		55%	45%	-0.18		32%	36%	0.10
Site 20	6		100%	100%	0.00		33%	17%	-0.39		0%	33%	1.00

[37] Sites 3, 6, and 18 had follow-up sample sizes that were too small to be interpreted with confidence (1-2 youth) and thus are grayed out of the table.

Table H9. Employment Outcomes by Site within Treatment Youth

Site	**N for S3**	***Employed Currently***	S1 %	S3 %	MSD	***Hourly wage***	S1 Mean	S1 SD	S3 Mean	S3 SD	MSD
Site 1	12		0%	36%	1.07		7.00	1.80	8.10	1.13	0.73
Site 2	6		17%	83%	1.79		8.81	0.13	10.08	0.43	4.05
Site 3	2										
Site 4	37		24%	64%	0.88		10.19	4.02	11.54	2.66	0.40
Site 5	6		33%	33%	0.00		7.15	1.52	10.42	1.63	2.08
Site 6	1										
Site 7	4		100%	75%	-0.82		6.88	1.83	9.75	2.12	1.45
Site 8	9		0%	44%	1.26		9.63	1.32	8.53	1.54	-0.76
Site 9	19		16%	63%	1.11		10.00	3.30	12.26	4.88	0.54
Site 10	8		13%	75%	1.62		8.91	2.86	10.54	1.63	0.70
Site 11	18		18%	61%	0.99		10.01	2.51	12.36	4.76	0.62
Site 13	15		40%	50%	0.20		12.69	12.0	9.65	1.80	-0.35
Site 14	9		0%	67%	2.00		8.75	0.80	14.55	2.24	3.45
Site 15	4		50%	50%	0.00		9.38	1.69	10.35	3.29	0.37
Site 16	32		6%	61%	1.42		10.10	2.68	10.28	1.67	0.08
Site 17	5		80%	60%	-0.45		8.19	0.75	9.81	4.08	0.55
Site 18	1										
Site 19	22		0%	82%	3.00		10.49	1.87	11.75	4.24	0.38
Site 20	6		33%	50%	0.34		9.01	0.64	11.68	2.11	1.72

[38] Sites 3, 6, and 18 had follow-up sample sizes that were too small to be interpreted with confidence (1-2 youth) and thus are grayed out of the table.

Table H10. Employment Outcomes by Site within Treatment Youth

Site	**N for S3**	***Longest Employment***	S1 Mean	S1 SD	S3 Mean	S3 SD	MSD	***Number of Job Search Activities Completed***	S1 Mean	S1 SD	S3 Mean	S3 SD	MSD
Site 1	12		1.00	1.61	4.00	4.79	0.84		1.83	2.08	2.33	1.23	0.29
Site 2	6		4.67	6.53	12.17	6.55	1.15		3.33	2.50	3.83	2.79	0.19
Site 3 [39]	2												
Site 4	37		9.33	8.89	12.65	8.87	0.37		5.73	1.85	5.38	1.91	-0.19
Site 5	6		3.75	7.04	7.83	7.96	0.54		2.83	2.14	4.83	1.94	0.98
Site 6	1												
Site 7	4		19.25	7.59	23.25	3.50	0.68		5.50	1.29	3.75	2.75	-0.81
Site 8	9		13.78	6.57	15.06	6.22	0.20		6.33	1.66	5.56	2.24	-0.39
Site 9	19		15.00	7.66	18.39	6.46	0.48		6.21	1.87	6.32	1.53	0.06
Site 10	8		12.38	6.70	16.63	5.29	0.70		6.00	1.85	5.63	1.69	-0.21
Site 11	18		15.33	8.37	17.78	7.37	0.31		6.44	2.09	6.17	2.23	-0.13
Site 13	15		7.64	7.19	15.57	8.45	1.01		3.20	2.24	3.53	1.77	0.17
Site 14	9		6.00	5.34	8.89	5.69	0.52		6.44	1.51	5.11	3.02	-0.56
Site 15	4		17.50	6.56	17.50	6.56	0.00		5.25	2.22	6.25	0.96	0.59
Site 16	32		13.66	8.61	13.61	8.50	-0.01		6.47	1.52	5.59	2.30	-0.45
Site 17	5		3.30	3.73	9.40	8.53	0.93		4.60	1.67	4.20	1.64	-0.24
Site 18	1												
Site 19	22		8.50	6.50	7.73	7.16	-0.11		6.55	1.87	6.41	1.59	-0.08
Site 20	6		14.67	9.48	18.50	7.84	0.44		6.00	1.26	5.67	1.75	-0.22

[39] Sites 3, 6, and 18 had follow-up sample sizes that were too small to be interpreted with confidence (1-2 youth) and thus are grayed out of the table.

Table H11. Employment Outcomes by Site within Treatment Youth

Program	**Program N for S3**	***Currently Facing Charges or on Probation or Parole***	Pre-Test Percent	Follow-up Percent	MSD	***Mean Civic Engagement***	Pre-Test Mean(SD)	Follow-up Mean(SD)	MSD
Site 1	12		0%	0%	0.00		3.28 (1.0)	3.26 (1.08)	-0.01
Site 2	6		33%	15%	-0.43		3.07 (0.77)	3.03 (0.58)	-0.06
Site 3 [40]	2								
Site 4	37		32%	40%	0.16		3.35 (1.25)	3.40 (1.45)	0.04
Site 5	6		17%	33%	0.39		2.61 (0.85)	2.48 (1.04)	-0.13
Site 6	1								
Site 7	4		25%	23%	-0.05		2.50 (0.74)	2.96 (1.28)	0.44
Site 8	9		11%	4%	-0.28		3.42 (1.31)	3.28 (0.98)	-0.12
Site 9	19		0%	0%	0.00		3.29 (0.89)	3.40 (1.17)	0.11
Site 10	8		0%	7%	0.38		3.17 (0.87)	3.05 (0.76)	-0.15
Site 11	18		0%	4%	0.29		2.51 (1.09)	2.64 (1.18)	0.11
Site 13	15		7%	11%	0.14		3.60 (1.24)	3.12 (1.34)	-0.37
Site 14	9		11%	11%	0.00		2.71 (1.08)	2.85 (1.25)	0.12
Site 15	4		0%	11%	0.50		3.15 (1.34)	2.88 (0.46)	-0.27
Site 16	32		9%	8%	-0.04		3.06 (1.09)	3.17 (1.14)	0.10
Site 17	5		40%	27%	-0.27		2.40 (1.03)	2.84 (1.13)	0.40
Site 18	1								
Site 19	22		0%	0%	0.00		3.18 (1.04)	3.30 (0.96)	0.11
Site 20	6		0%	0%	0.00		3.27 (1.41)	3.16 (0.91)	-0.09

[40] Sites 3, 6, and 18 had follow-up sample sizes that were too small to be interpreted with confidence (1-2 youth) and thus are grayed out of the table.

Appendix I: Technical Assistance Strategies and Lessons Learned

I1. Introduction

Recruiting OY at each program to participate in the comparison group was critical to the success of the evaluation. This appendix describes strategies and lessons learned while providing technical assistance to the programs in the evaluation bundle.

I2. Lessons Learned from Comparison Group Recruitment and Retention

Evaluators worked with each program to develop and execute a program-specific Individual Recruitment Plan (IRP) for comparison group recruitment and retention. The IRP was developed to provide each grantee a roadmap of the practical steps and shared responsibilities for recruiting and retaining comparison group youth.

Components of the IRP include a description of the program and its key characteristics vis-à-vis member and comparison group recruitment; a list of key contacts at the local level; a description of member recruitment; methods of comparison group youth recruitment, including internally (through "over-recruitment" of members) and externally from community partners; the consent process for OY under and over 18 years of age; preferred survey modalities for members and comparison OY; plans for distributing survey incentives; an action plan for implementing the IRP, and a plan for maintaining updated contact information for comparison and treatment youth.

The IRP followed a standard format for all programs, while allowing flexibility to accommodate program-specific needs. The IRP also supported program-level evaluation capacity building (ECB) by providing each program a customized roadmap for comparison group recruitment that could be reused for future evaluations. Participating in developing and carrying out their own IRP afforded program staff an experiential opportunity to gain skills in key aspects of comparison group recruitment, including working with community partners to recruit youth, distinguishing eligible from ineligible youth, and adhering to a timeframe for comparison group recruitment corresponding to their member enrollment period.

Factors complicating comparison group *recruitment* included:

- Variation in the schedule programs followed to recruit and enroll AmeriCorps members, and the corresponding need to vary comparison group recruitment. These arrangements ranged from the simplest in which a program enrolls all members once per year within a well-defined timeframe (e.g., August 1 to September 30) to more complicated arrangements, such as enrolling multiple cohorts during the year and enrolling members throughout the year ("rolling enrollment"). These variations called for customization of the IRP and careful monitoring of comparison group recruitment to keep the process on track. Programs with rolling member enrollment faced the added challenge of keeping program staff and community partners engaged in a lengthy comparison group recruitment process.

- Challenges in meeting member enrollment targets, which obliged some programs to extend the period for member enrollment and comparison group recruitment. In some

cases, programs were not able to meet their member enrollment targets, thus reducing the number of study participants.

- Limited capacity of most programs to source comparison youth internally, often necessitating substantial reliance on community partners to attain the prescribed 2-to-1 ratio of comparison youth to members.

- Limited means to incentivize resource-strapped community partners to supply comparison youth for an evaluation in which these partners had no real stake.

The principle factor complicating comparison group *retention* is the fact that OY, as disconnected individuals, are inherently difficult to maintain contact with over an extended period of time. As one program stated in an email message, "We predicted this would be a problem, as the young people we serve are, by definition, disconnected." In anticipation of this challenge, the evaluators collected contact information from secondary contacts (family and friends) on the baseline survey. Evaluators also contacted OY and programs periodically to request updated contact information. Nonetheless, because OY frequently changed phone numbers, email addresses and physical addresses, it was difficult to locate them for the follow-up surveys.

These challenges point to several valuable lessons learned through the comparison group recruitment and retention process.

First, future evaluations would benefit from including programs that have more concise and well-defined timeframes (e.g., no more than 4 weeks) for enrolling opportunity youth. This would support establishing a clear timeframe for enrolling comparison youth in the evaluation, either through over-recruitment of program youth or through partner organizations. Selecting programs with the same service term (or length of service), while less essential, would also simplify data collection timeframes and achieve greater homogeneity across the bundle. Future evaluations would also benefit from identifying community partners with the resources (staff and financial resources), leadership, engagement, and similar population necessary to secure buy-in from youth to participate in the survey. Programs that have a memorandum of understanding with a community partner should consider updating that MOU to include language supportive of both organizations' future evaluation efforts. Community partner participation may often be crucial to the success of comparison group recruitment. To facilitate timely comparison group recruitment, engagement of community partners should commence several months before comparison youth are needed. An effort should also be made to provide community partners with meaningful near-term incentives for participation. These incentives would typically be in the form of in-kind or financial compensation, and would serve to acknowledge the time and effort partner agencies with limited resources must devote to the comparison group recruitment process.

Another valuable lesson came about when the evaluators and programs were working to pilot the development of the IRP. It was clear from this endeavor that the development of the IRP could be standardized and streamlined in future bundled evaluations. An option may be to create a condensed version of the IRP that focuses on the timing of the practical steps program staff should follow to successfully navigate the comparison group recruitment process. This would

result in two versions of the IPR: a condensed version to support programs' immediate needs for implementation and a fuller version to support ECB and future evaluations.

Finally, the feasibility and cost-savings benefits stemming from relying on local program staff to conduct key evaluation tasks (as opposed to placing evaluation staff in the field) did not always translate into efficiency, as shown by tardy delivery of data and attrition of study participants. One solution to this challenge would be to recruit local representatives (such as program participants with a track record of successful program participation) to serve as an extension of the evaluation team. These "surrogates" or "proxies" to the evaluation team could receive a small stipend in return for conducting outreach to community partners, supporting program staff and community partners to collect data from study participants, maintaining contact information for participants, and undertaking other evaluation-related tasks, as appropriate, that programs do not have the time or resources to conduct. Recruiting local representatives to serve as surrogates to the evaluation team may also be more cost-effective than, for example, sending evaluators on program visits to each program. The local representative's onsite presence may reinforce the study's importance to local stakeholders. Employing opportunity youth as surrogates would also build their resume and help them gain marketable skills.

I3. Lesson Learned from Survey Implementation and Management

Program staff turnover and changes in staff roles occurred from time to time during the more than three years of the evaluation. Program staff that worked directly with the evaluators and who were in the best position to understand and carry out local-level survey implementation tasks, sometimes assigned these tasks to other staff that were not as familiar with survey administration procedures. Details regarding study eligibility criteria or parental consent were sometimes lost in the transition, resulting in some ineligible youth taking the survey. Program staff with direct responsibility for survey administration were not always able to answer youths' questions about the survey or the study because they had not been fully trained in these details. Programs were often understaffed, so even staff assigned to work on the study had limited time to complete tasks such as checking eligibility, filling out tracking sheets, and responding to evaluators' information requests. Short of limiting bundled evaluations to "highly resourced" programs (which runs counter to one of the approach's stated purposes), these organizational capacity limitations imply the need for more substantial on-the-ground, program-level support for evaluation activities. This might take the form of a local volunteer with dedicated weekly time to support evaluation tasks. In the particular case of this bundled evaluation, second-term AmeriCorps members might have been recruited as incipient research assistants, possibly with the double incentive of earning an additional stipend (paid from the evaluation budget) while acquiring skills and experience to add to their resume.

Other implementation challenges stemmed from member recruitment that was slower than anticipated and challenges with comparison group recruitment. These factors resulted in an elongated timeline for pre-test survey administration. Several program sites experienced challenges reaching their recruitment goals and either delayed member onboarding until all slots were filled or administered the pre-test to the initial batch of newly recruited members and then surveyed other members as they came on board later. Additionally, in some cases, programs and evaluators agreed to extend comparison group data collection for several weeks beyond the original one-month window in response to difficulties with comparison group recruitment.

These challenges reinforced the need (already recognized by evaluators at the outset of the study) to include more points-of-contact in TA discussions related to survey implementation. While limited program staffing and busy schedules meant it was not always possible to include a "backup" staff member on TA calls, all things being equal, programs that did this more regularly tended to have fewer survey implementation challenges. Evaluators often included backup points-of-contact on pertinent email messages. Challenges with reaching member or comparison youth recruitment targets highlight the importance of careful screening of programs for inclusion in the evaluation bundle and in identifying a large pool of community partners and assessing their viability as potential source of comparison youth early on.

Paper-and-pencil surveys administered directly to members and program applicants have a much higher response rate than online or phone surveys, suggesting that paper surveys should be encouraged over other modalities. As a matter of necessity, the online survey was the primary modality for most data collection during the follow-up survey phase since most study participants were no longer accessible to sites and community partners. However, even in the few instances when multiple survey modalities were available to a program conducting the follow-up survey, the paper survey often proved most efficacious. Integrating the survey into existing program application and exit processes also produced the best response rates. Future studies should consider using paper surveys integrated into the entry and exit processes.

Evaluators provided program staff a baseline survey tracking sheet (BSTS) to track distribution of incentives and facilitate follow-up by collecting additional contact information and allowing evaluators to link a participant to successive surveys. While the BSTS was an invaluable tool, it was sometimes difficult to get the completed BSTS back from program sites. Evaluators found that including a self-addressed stamped envelope clearly marked "Baseline Survey Tracking Sheet" helped overcome this problem. Future studies may also benefit from providing a version of the BSTS that can be completed and submitted online.

Some programs saw members exit early, leading to challenges with retaining these members in the treatment group for follow-up surveys. Early exits occurred for a variety of reasons, including compelling personal circumstances, pursuing education or employment opportunities, and release for inappropriate conduct. Opportunity youth that exited early were less likely to take the first follow-up survey, which also meant a lost opportunity to obtain updated contact information. Frequent check-ins with the programs was necessary to minimize these lost survey opportunities.

Programs that sought to keep in touch with recent program graduates were most successful in providing evaluators updated contact information for former members participating in the study. A few programs and community partners were successful in maintaining contact with some recent program alumni, but most programs and partners that aspired to maintain contact with alumni lacked the resources to do so. The ability to offer participants a larger incentive ($20) for the final survey supported retention of study participants, suggesting the benefit of increasing incentives incrementally with each successive survey. In addition, working with programs prior to study implementation to set up or strengthen systems for keeping in contact with program alumni could boost retention of study participants while building sustainable program infrastructure.

I4. Lesson Learned from Study Participant Retention Efforts

Future bundled evaluations could benefit from more intensive efforts to maintain contact with treatment and comparison group members between surveys. In addition, offering modest financial incentives for study participants and community partners to provide updated contact information between surveys could improve retention while reducing the burden of outreach work by programs and evaluators. In addition to the obvious financial reward, providing incentives between surveys for sharing updated contact information with evaluators may habituate participants to the idea of staying connected to the study on an ongoing basis.

I5. Lesson Learned from Addressing Program Staff Turnover

While staff turnover became less of a liability to the evaluation as the primary responsibility for survey administration shifted from the programs to the evaluators, it still limited the ability of programs to participate effectively in evaluation-related tasks, such as supporting outreach for retention of study participants, submitting key paperwork, and enhancing evaluation capacity. Programs that engaged more than one staff member in evaluation TA were more likely to complete evaluation tasks and build evaluation capacity. Allowing generous lead time and flexible deadlines to fulfill information requests also reduced the perceived burden on programs.

The development of reusable systems, tools, and infrastructure offers the best chance for gains in grantee evaluation capacity to become sustainable even in the face of staff turnover. These systems, if properly designed and implemented, can remain useful even when current staff move on. The development of these systems should be an integral part of the TA provided to support an evaluation. Ideally, programs would have the opportunity to build infrastructure for data collection and maintenance and increase their capacity to conduct less intensive evaluation designs and become accustomed to using evaluation results prior to joining a bundled impact evaluation.

I6. Additional Lessons Learned

While economical and efficient, conducting the study remotely hindered establishing deep relationships with program staff and community partners. TA calls and emails gave evaluators a chance to understand the effects of changes in program staffing, cohesion between programs and community partners, and generally how the programs operate. However, because TA calls were often limited to one program staff person, an understanding of the complexities of working with opportunity youth was limited to this staff person's perspective. Remote TA limited evaluators' ability to witness aspects of program operations that program staff may have taken for granted and hence not mentioned during TA calls, as well as the dynamics of staff-participant interactions. Evaluators can only speculate on what may have transpired on the ground and how it might have informed the implementation of the evaluation.

It cannot be overstated that the majority of non-profit organizations do not have the resources to stretch their work beyond direct client services, and partner organizations—which are much more difficult than AmeriCorps programs to leverage—require more than a desire to assist opportunity youth to participate in evaluation studies. They require financial or human resources, or both. Future bundled evaluations performed in collaboration with small non-profits would benefit greatly from these additional resources.

Appendix J. References

Allen, L., Miles, M., and Steinberg, A. (2014). Achieving collective impact for opportunity youth: emerging lessons on using data and resources to improve the prospects of young people. Stanford Social Innovation Review. Fall 2014.

Aspen Institute (2014, December). Opportunity Youth Network. Retrieved from http://aspencommunitysolutions.org/press-release-aspen-institute-convening-explores-educational-and-employment-opportunities-for-opportunity-youth/.

Austin, P.C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research*, 46, pp.399-424.

Barley, Z.A. & Jenness, M. (1993). Cluster evaluation: A method to strengthen evaluation in smaller programs with similar purposes, Evaluation Practice, 14(2), 141-147.

Belfield, C.R., Levin, J. M., & Rosen, R. (2012). The economic value of opportunity youth. Corporation for National and Community Service. http://www.civicenterprises.net/MediaLibrary/Docs/econ_value_opportunity_youth.pdf

Bird, K. (2013) Opportunity says it all: how five communities are supporting the transformation, education, and employment success of young black men. Center for Law and Social Policy. http://www.clasp.org/resources-and-publications/publication-1/Opportunity-Says-It-All.pdf

Bird, K., Foster, M., and Ganzglass, E. (2014). New opportunities to improve economic and career success for low-income youth and adults: key provisions of the Workforce Innovation and Opportunity Act (WIOA). Center for Law and Social Policy. http://www.clasp.org/resources-and-publications/publication-1/KeyProvisionsofWIOA-Final.pdf

Bloom, D., Rich, S., Redcross, C., Jacobs, E., Yahner, J., & Pindus, N. (2009). Alternative Welfare-to-Work Strategies for the Hard-to-Employ: Testing transitional Jobs and Pre-Employment Services in Philadelphia." MDRC.

Bridgeland, H., & Milano, J. (2012). Opportunity Road: The Promise and Challenge of America's Forgotten Youth 2012, Civic Enterprises & America's Promise Alliance.

Bridgeland, J. & Mason-Elder, T. (2012). National roadmap for opportunity youth. Civic Enterprises.

Consolidated Appropriations Act (2014). Pub. L. 113-76, 128 Stat. 414. https://www.congress.gov/113/plaws/publ76/PLAW-113publ76.pdf

Corcoran, M., Hanleybrown, F., Steinberg, A., & Tallant, K. (2012). Collective impact for opportunity youth. FSG Social Impact Consultants.

Dong, N. and Maynard, R. A. (2013). *PowerUp!*: A tool for calculating minimum detectable effect sizes and sample size requirements for experimental and quasi-experimental

designs. *Journal of Research on Educational Effectiveness*, *6*(1), 24-67. doi: 10.1080/19345747.2012.673143

Duckworth, A. & Quinn, P.D. (2009). Development and Validation of the Short Grit Scale (Grit-S). *Journal of Personality Assessment, 91(2),* pp. 166-174. doi: 10.1080/00223890802634290

Forss, K., Kruse, S., Taut, S. & Tenden, E. (2006). Chasing a Ghost? An Essay on Participatory Evaluation and Capacity Development, Evaluation, 12(1), pp.128-144. http://journals.sagepub.com/doi/abs/10.1177/1356389006064190

Gennetian L., Sciandra M., Sanbonmatsu L., Ludwig J., Katz L., Duncan G., Kling J., & Kessler R. (2012). The Long-Term Effects of Moving to Opportunity on Youth Outcomes. Cityscape [Internet]. 14(2):137-68. http://scholar.harvard.edu/lkatz/publications/long-term-effects-moving-opportunity-youth-outcomes

Gewertz, C. (2011). GED revision opens path to higher education. Education Week, 31(12) 1, 16-17.

Hanleybrown, F., Kania, J., & Kramer, M. (2012). Channeling change: making collective impact work.

Julian, T.A., & Kominski, R.A. (2011). Education and Synthetic Work- Life Earnings Estimates. American Community Survey Reports, ACS-14. U.S. Census Bureau, Washington, DC.

Kemple J. & Willner C. (2008). Long-term impacts on labor market outcomes, educational attainment, and transitions to adulthood." New York, NY: MDRC.

Lochner, L., & Moretti, A. (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. American Economic Review, 94, 155-189.

Martin, V., & Broadus, J. (2013). Enhancing GED instruction to prepare students for college and careers, early success in Laguardia Community College's Bridge to Health and Business Program. http://www.mdrc.org/sites/default/files/Enhancing_GED_Instruction_brief.pdf

Mortimer, J.T. (2010). "The Benefits and Risks of Adolescent Employment." Prev Res 17(2): 8–11, http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2936460

Osypuk, T. L., Tchetgen, E. J. T., Acevedo-Garcia, D., Earls, F. J., Lincoln, A., Schmidt, N. M., & Glymour, M. M. (2012). Differential mental health effects of neighborhood relocation among youth in vulnerable families: results from a randomized trial. Archives of general psychiatry, 69(12), 1284-1294.

Powell, M. J. D. (2009), "The BOBYQA algorithm for bound constrained optimization without derivatives", Report No. DAMTP 2009/NA06, Centre for Mathematical Sciences, University of Cambridge, UK. http://www.damtp.cam.ac.uk/user/na/NA_papers/NA2009_06.pdf.

Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. American Journal of Evaluation, 29(4), 443-459.

Preskill, H., & Torres, R.T. (1999). Building capacity for organizational learning through evaluative inquiry, 5(1), pp. 42-60. http://www.stes-apes.med.ulg.ac.be/Documents_electroniques/EVA/EVA-GEN/ELE%20EVA-GEN%207458.pdf

Rutschow, E. Z., & Crary-Ross, S. (2014). Beyond the GED: promising models for moving high school dropouts to college.

Schochet, P. Z., Burghardt, J., & McConnell, S. (2008). Does job corps work? Impact findings from the National Job Corps Study. The American Economic Review, 1864-1886.

Snyder, T.D., & Dillow, S.D. (2011). Digest of Education Statistics 2010 (NCES 2011-015). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. http://nces.ed.gov/pubs2011/2011015.pdf

Stuart Foundation (2011). Promising Strategies from the California Connected by 25 Initiative: Tips and Resources to Improve Outcomes for Transition Age Foster Youth 2011.

Sum, A., Khatiwada, I., Trubskyy, M., & Palma, S. (2014). The plummeting labor market fortunes of teens and young adults. Brookings Institution.

Theodos, B., Pergamit, M., Edelstein, S., George, T., & Freiman, L. (2014). Preparing youth for college and career: a process evaluation of urban alliance. Urban Institute.

Vericker, T., Pergamit, M., Macomber, J., & Kuehn, D. (2009). Vulnerable Youth and the Transition to Adulthood. ASPE Policy Brief (Washington DC: U.S. Department of Health and Human Services).

Wallace, J. (Winter, 2014). Jobs for Disconnected Youth and Talent for Suppliers: Insights from an innovation lab dedicated to bridging the talent divide. Stanford Social Innovation Review.

Wahl, E., Kleinbard, P., Chavez Reilly, M., Beaton J. & Vazquez, V. (2012). The best of two worlds: Lessons from a community college – Community organization collaboration to increase Student Success, prepared by of the Youth Development Institute for Jobs for the Future http://www.jff.org/sites/default/files/publications/BestOfTwoWorlds_032112.pdf